

CORINNA COUPETTE

Juristische  
Netzwerkforschung

---

Mohr Siebeck

*Corinna Coupette*  
Juristische Netzwerkforschung





Corinna Coupette

# Juristische Netzwerkforschung

Modellierung, Quantifizierung und  
Visualisierung relationaler Daten im Recht

Mohr Siebeck

*Corinna Coupette*, geboren 1992; Studium der Rechtswissenschaft an der Bucerius Law School und der Stanford Law School; 2015 Erste Juristische Staatsprüfung; Wissenschaftliche Mitarbeiterin am Max-Planck-Institut für Steuerrecht und Öffentliche Finanzen; Bachelor of Science in Informatik, Ludwig-Maximilians-Universität München; 2018 Promotion (Dr. iur.), Bucerius Law School.

Dissertation, Bucerius Law School  
Datum der mündlichen Prüfung: 11.07.2018

Zur Arbeit gehört ein Online-Appendix, der abrufbar ist unter:  
DOI 10.1628/978-3-16-157012-4-appendix

ISBN 978-3-16-157011-7 / eISBN 978-3-16-157012-4  
DOI 10.1628/978-3-16-157012-4

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind über <http://dnb.dnb.de> abrufbar.

© 2019 Mohr Siebeck Tübingen. [www.mohrsiebeck.com](http://www.mohrsiebeck.com)

Dieses Werk ist lizenziert unter der Lizenz „Creative Commons Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International“ (CC-BY-NC-ND 4.0). Eine vollständige Version des Lizenztextes findet sich unter: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>.

Jede Verwendung, die nicht von der oben genannten Lizenz umfasst ist, ist ohne Zustimmung des Verlags unzulässig und strafbar.

Das Buch wurde von Gulde Druck in Tübingen auf alterungsbeständiges Werkdruckpapier gedruckt und von der Buchbinderei Spinner in Ottersweier gebunden.

Printed in Germany.

*Meinem Großvater*



## Vorwort

Die Arbeit wurde im Frühjahrstrimester 2018 von der Bucerius Law School als Dissertation angenommen; alle Auswertungen sind auf dem Stand von Januar 2018. Ich danke meinem Erstgutachter, *Christian Bumke*, und meiner Zweitgutachterin, *Katharina Anna Zweig*, für die Bereitschaft, mein interdisziplinäres Unterfangen mitzutragen, sowie *Wolfgang Schön* für die Unterstützung.

Besonders profitiert hat die Untersuchung von den Rückmeldungen der Otto-Hahn-Gruppe, deren Mitglieder alle Teile der Untersuchung gelesen, annotiert und mit mir diskutiert haben. Mein Dank gilt daher *Felix Bassier*, *Andreas M. Fleckner*, *Miguel Gimeno Ribes*, *Amin Kachabia*, *Philipp Aron Leimbach* und *Johannes Liefke*.

Ohne meine Eltern, *Regina Coupette* und *Andreas Coupette*, und meinen Bruder, *Fabian Coupette*, wäre die Arbeit nicht möglich gewesen; für ihre bedingungslose Unterstützung bin ich ihnen sehr dankbar. Gewidmet ist die Untersuchung meinem Großvater, *Claus Coupette*. Dass er die Fertigstellung der Arbeit noch miterleben konnte, war das größte Geschenk.

München, im Dezember 2018

*Corinna Coupette*



# Inhaltsverzeichnis

Vorwort .....	VII
Tabellenverzeichnis .....	XIV
Abbildungsverzeichnis .....	XV
Kapitel 1. Einleitung .....	1
A. <i>Annäherung</i> .....	5
B. <i>Aufbau</i> .....	10
Kapitel 2. Theorie .....	11
A. <i>Literatur</i> .....	11
I. Juristische Texte .....	13
II. Juristische Akteure .....	18
III. Recht und Realität .....	22
B. <i>Gegenstand</i> .....	28
I. Objekte .....	28
1. Rechtssubjekte .....	29
a) Individuen .....	29
b) Kollektive .....	31
2. Rechtstexte .....	32
a) Gerichtsentscheidungen .....	32
b) Gesetzestexte .....	33
c) Verwaltungsentscheidungen .....	34
d) Verträge .....	35
e) Literatur .....	36
3. Rechtsbegriffe .....	36

II.	Beziehungen .....	37
1.	Beziehungsgrad .....	38
2.	Beziehungsprofil .....	39
3.	Beziehungsform .....	51
III.	Probleme .....	55
1.	Selektion .....	56
2.	Inklusion .....	57
3.	Rekursion .....	59
C.	<i>Perspektiven</i> .....	61
I.	Untersuchungsziele .....	62
1.	Exploration .....	62
a)	Charakteristika .....	62
b)	Möglichkeiten .....	63
c)	Verantwortung .....	64
2.	Explikation .....	65
a)	Charakteristika .....	65
b)	Möglichkeiten .....	67
c)	Verantwortung .....	68
3.	Extrapolation .....	68
a)	Charakteristika .....	68
b)	Möglichkeiten .....	69
c)	Verantwortung .....	70
II.	Untersuchungsebenen .....	71
1.	Mikroebene .....	72
2.	Mesoebene .....	73
3.	Makroebene .....	73
Kapitel 3. Methodik .....		74
A.	<i>Arbeitsmaterial: Daten</i> .....	74
I.	Reale Daten .....	74
1.	Quellen juristischer Daten .....	75
a)	Inhaltliche Schwierigkeiten .....	76
b)	Technische Schwierigkeiten .....	79
c)	Rechtliche Schwierigkeiten .....	82
2.	Fehlerquellen juristischer Daten .....	85
a)	Abdeckungsfehler .....	85
b)	Aufzeichnungsfehler .....	86
c)	Aufbereitungsfehler .....	87

II.	Fiktive Daten .....	88
1.	Datensatz 1: Rechtsprechungsdaten .....	88
2.	Datensatz 2: Unternehmensdaten .....	91
III.	Beschreibung von Daten .....	93
<i>B.</i>	<i>Visualisierung</i> .....	99
I.	Funktionen von Visualisierung .....	100
II.	Prozess der Visualisierung .....	103
III.	Visualisierung von Graphen .....	109
1.	Graphdiagramme .....	109
a)	Merkmalsbasierte Positionierung .....	110
b)	Kräftebasierte Positionierung .....	112
c)	Distanzbasierte Positionierung .....	116
2.	Graphmatrizen .....	118
<i>C.</i>	<i>Quantifizierung</i> .....	122
I.	Basis: Bewegung auf Graphen .....	123
II.	Mikroebene .....	129
1.	Problemstellung: Was ist wichtig? .....	129
2.	Lösungsansätze .....	130
a)	Gradzentralitäten .....	131
b)	Rückkopplungszentralitäten .....	135
c)	Positionszentralitäten .....	139
III.	Mesoebene .....	142
1.	Problemstellung: Was gehört zusammen? .....	142
2.	Lösungsansätze .....	145
a)	Gemeinschaften .....	147
b)	Positionen und Rollen .....	159
IV.	Makroebene .....	165
1.	Problemstellung: Was ist charakteristisch? .....	165
2.	Lösungsansätze .....	166
a)	Statistische Kennzahlen .....	167
b)	Eindimensionale Verteilungen .....	174
c)	Zweidimensionale Verteilungen .....	181
V.	Zusatzfaktor: Zeit .....	187
<i>D.</i>	<i>Modellierung</i> .....	193
I.	Strukturen .....	195
1.	Verdrahtungsmodelle .....	196
2.	Wachstumsmodelle .....	204

II. Prozesse .....	211
1. Modelle für Suchprozesse .....	212
2. Modelle für Diffusionsprozesse .....	216
<i>E. Herausforderung: Mehrschichtigkeit</i> .....	224
Kapitel 4. Praxis .....	227
A. Daten .....	228
I. Netzwerkdefinition .....	229
1. Objekte: Entscheidungen in BVerfGE 1–140 .....	230
2. Beziehungen: Zitate im Format „{Band}, {Startseite}“ .....	232
3. Zusatzinformationen: Merkmale .....	234
II. Netzwerkkonstruktion .....	235
1. Sammlung der Entscheidungen .....	236
2. Extraktion der Zitate .....	240
3. Aufbereitung der Daten .....	248
B. Experimente .....	251
I. Makroebene: Was ist charakteristisch? .....	252
1. Die Sammlung sichten: Graphiken und Kennzahlen .....	252
a) Visuelle Repräsentation .....	253
b) Quantitative Zusammenfassung .....	258
2. Das Zitiernetzwerk beschreiben: Verteilungen .....	263
II. Mikroebene: Was ist wichtig? .....	269
1. Zentrale Entscheidungen finden: Abstrakte Wichtigkeit .....	269
a) Statische Betrachtung: Meistzitierte Entscheidungen ..	270
b) Dynamische Betrachtung: Juristische Zitierkurven .....	273
2. Leseempfehlungen abgeben: Konkrete Wichtigkeit .....	278
III. Mesoebene: Was gehört zusammen? .....	282
1. Rechtsprechung ordnen: Gemeinschaften in BVerfGE- Entscheidungen .....	282
a) Problemfelder: Gemeinschaften von $k$ -Cliques .....	282
b) Rechtsbereiche: Algorithmische Gemeinschafts- erkennung .....	289
2. Begriffskontexte erkunden: Textbasierte Subgraphen .....	301

C. <i>Evaluation</i> .....	313
I. Inhaltliche Erkenntnisse .....	313
II. Methodische Erkenntnisse .....	315
III. Praktische Erkenntnisse .....	316
Kapitel 5. <i>Fazit</i> .....	320
A. <i>Rückblick</i> .....	320
B. <i>Ausblick</i> .....	324
Appendix .....	329
A. <i>Struktur des Online-Appendix</i> .....	330
B. <i>Ergebnisse der Stichprobenüberprüfung</i> .....	331
C. <i>Registerzeichen des Bundesverfassungsgerichts</i> .....	332
D. <i>Namen der Gemeinschaften in BVerfGE</i> .....	333
Literatur .....	337
Glossar .....	362
Sachregister .....	371

## Tabellenverzeichnis

1.1	Annäherung: Grundbegriffe der Netzwerkforschung . . . . .	9
2.1	Beziehungen: Zusammenfassung der Beziehungsprofile . . . . .	51
3.1	Fiktive Daten: Zitierdaten . . . . .	90
3.2	Fiktive Daten: Unternehmensdaten . . . . .	92
3.3	Beschreibung: Skalen von Merkmalen . . . . .	94
3.4	Beschreibung: Merkmalstypen (Rechtsprechungsdaten) . . . . .	96
3.5	Beschreibung: Merkmalstypen (Unternehmensdaten) . . . . .	97
3.6	Beschreibung: Zusammenfassende Statistiken (Unternehmensdaten) . . . . .	97
3.7	Beschreibung: Zusammenfassende Statistiken (Rechtsprechungsdaten) . . . . .	98
3.8	Charakterisierung: Netzwerkstatistiken . . . . .	170
3.9	Zeit: Veränderungsprozesse . . . . .	188
4.1	Mikroebene: Meistzitierte Entscheidungen (binäre Zählung) . . . . .	270
4.2	Mikroebene: Meistzitierte Entscheidungen (gewichtete Zählung) . . . . .	271
4.3	Mikroebene: Meistzitierte Entscheidungen (Vergleich der Zählungen) . . . . .	272
4.4	Mikroebene: Meistzitierte Entscheidungen (temporale Betrachtung – binär) . . . . .	274
4.5	Mikroebene: Meistzitierte Entscheidungen (temporale Betrachtung – gewichtet) . . . . .	275
4.6	Mikroebene: Leseempfehlungen (BVerfGE 7, 198 – Lüth) . . . . .	280
4.7	Mikroebene: Leseempfehlungen (BVerfGE 120, 274 – Online-Durchsuchungen) . . . . .	280
4.8	Mikroebene: Leseempfehlungen (BVerfGE 51, 222 – 5%-Klausel) . . . . .	281
4.9	Mesoebene: Gemeinschaftserkennung ( <i>Beruf</i> ) . . . . .	292
4.10	Mesoebene: Gemeinschaftserkennung ( <i>Eigentum</i> ) . . . . .	292
4.11	Mesoebene: Gemeinschaftserkennung ( <i>Europa</i> ) . . . . .	293
4.12	Mesoebene: Gemeinschaftserkennung ( <i>Meinung/Presse/APR</i> ) . . . . .	293

# Abbildungsverzeichnis

Eine hochauflösende Ansicht aller Graphiken findet sich unter:  
DOI 10.1628/978-3-16-157012-4

1.1	Einleitung: Königsberg in der frühen Neuzeit	1
1.2	Einleitung: Königsberger Brückenproblem – Karte	2
1.3	Einleitung: Königsberger Brückenproblem – Graph	4
1.4	Annäherung: <i>Arbor Consanguinitatis</i> (1582)	6
1.5	Annäherung: Struktur des Bürgerlichen Gesetzbuchs	7
1.6	Annäherung: Struktur des Grundgesetzes	8
2.1	Beziehungen: Beziehungsgrade	38
2.2	Beziehungen: Beziehungsprofile (Grad 1)	40
2.3	Beziehungen: Beziehungsprofile (Grad 2)	45
2.4	Beziehungen: Multilaterale Beziehungen	54
2.5	Beziehungen: Beziehungsformen	55
2.6	Untersuchungsziele: Soziogramm (1934)	66
2.7	Untersuchungsebenen: Überblick	72
3.1	Reale Daten: Möglichkeiten der Zitaterfassung	77
3.2	Fiktive Daten: Gerichtslandschaft	89
3.3	Visualisierung: Rechtsprechungszitrate im Kontext	101
3.4	Visualisierung: Visualisierungsprozess	103
3.5	Visualisierung: Visuelle Variablen	106
3.6	Visualisierung: Histogramme	107
3.7	Visualisierung: Streudiagramm	108
3.8	Graphdiagramme: Merkmalsbasierte Positionierung	111
3.9	Graphdiagramme: Kräftebasierte Positionierung (Reproduktionsproblem)	113
3.10	Graphdiagramme: Kräftebasierte Positionierung (Skalierungsproblem)	115
3.11	Graphdiagramme: Distanzbasierte Positionierung (Multidimensionale Skalierung)	117
3.12	Graphmatrizen: Königsberger Brückenproblem	119
3.13	Graphmatrizen: Rechtsprechungs- und Unternehmensdaten	120
3.14	Graphmatrizen: Relevanz der Zeilen- und Spaltenanordnung	121

3.15	Bewegung auf Graphen: Königsberger Brücken	123
3.16	Bewegung auf Graphen: Ermittlung der Wegzahl	125
3.17	Bewegung auf Graphen: Kürzeste-Wege-Baum	126
3.18	Bewegung auf Graphen: Juristische Recherche	128
3.19	Wichtigkeit: Zitierprofile	133
3.20	Wichtigkeit: Gradzentralitäten	134
3.21	Wichtigkeit: Rückkopplungszentralitäten	139
3.22	Zusammengehörigkeit: Konzepte	146
3.23	Zusammengehörigkeit: Direkte Definitionen von Gemeinschaft	150
3.24	Zusammengehörigkeit: Hierarchisches Clustering	154
3.25	Zusammengehörigkeit: Indirekte Definitionen von Gemeinschaft	158
3.26	Zusammengehörigkeit: Positionen	161
3.27	Charakterisierung: Komponenten	169
3.28	Charakterisierung: Transitivität	173
3.29	Charakterisierung: Gradverteilungen	175
3.30	Charakterisierung: Verteilungstypen	177
3.31	Charakterisierung: Schiefe Verteilungen	179
3.32	Charakterisierung: Selbstähnlichkeit	181
3.33	Charakterisierung: Zweidimensionale Verteilungen	182
3.34	Charakterisierung: Gradassortativitäten	185
3.35	Zeit: Co-Zitation auf mehreren Ebenen	191
3.36	Modellierung von Netzwerkstrukturen: Vergleichsgrundlage (Unternehmensgraph)	196
3.37	Modellierung von Netzwerkstrukturen: Zufallsgraph	198
3.38	Modellierung von Netzwerkstrukturen: Regulärer Graph	200
3.39	Modellierung von Netzwerkstrukturen: Kleine-Welt-Modell	201
3.40	Modellierung von Netzwerkstrukturen: Konfigurationsmodell (ungerichtet)	202
3.41	Modellierung von Netzwerkstrukturen: Konfigurationsmodell (gerichtet)	204
3.42	Modellierung von Netzwerkstrukturen: Einfaches Wachstumsmodell	206
3.43	Modellierung von Netzwerkstrukturen: Kumulationsmodell und Kopiermodell	208
3.44	Modellierung von Suchprozessen: Zitate und Annotationen	213
3.45	Modellierung von Suchprozessen: Breitensuche und Tiefensuche	215
3.46	Modellierung von Diffusionsprozessen: SI-Modell (Graphdiagramme)	217
3.47	Modellierung von Diffusionsprozessen: SI-Modell (Liniendiagramm)	219
4.1	Daten: Zitierte Quellen in Entscheidungen des BVerfG	230
4.2	Daten: Schritte bei der Netzwerkkonstruktion	236

4.3	Daten: Sammlung von BVerfGE-Entscheidungen	239
4.4	Daten: Verhältnis von gefundenen zu gesuchten Zitaten	242
4.5	Daten: Zufallsstichprobe	246
4.6	Daten: Extraktion von Zitaten aus BVerfGE-Entscheidungen	247
4.7	Daten: Aufbereitung der Daten zu den BVerfGE-Entscheidungen	250
4.8	Experimente – Makroebene: Zeitliche Struktur	254
4.9	Experimente – Makroebene: Bandstruktur	256
4.10	Experimente – Makroebene: Senatsstruktur	257
4.11	Experimente – Makroebene: Entwicklung der Entscheidungslänge	259
4.12	Experimente – Makroebene: Seitenumfang der Senate	260
4.13	Experimente – Makroebene: Entscheidungsumfang nach Registerzeichen	261
4.14	Experimente – Makroebene: Entscheidungslänge nach Registerzeichen	262
4.15	Experimente – Makroebene: Schiefe Verteilung der Entscheidungslängen	263
4.16	Experimente – Makroebene: Gradverteilungen (Histogramme)	264
4.17	Experimente – Makroebene: Gradverteilungen (Anpassungen)	266
4.18	Experimente – Makroebene: Vermischung (nach Entscheidungsjahr)	267
4.19	Experimente – Makroebene: Vermischung (nach Entscheidungsband)	268
4.20	Experimente – Mikroebene: Zitierkurven	276
4.21	Experimente – Mesoebene: Juristische Problemfelder ( <i>Gleichheit bei Wahlen I</i> )	284
4.22	Experimente – Mesoebene: Juristische Problemfelder ( <i>Gleichheit bei Wahlen II</i> )	285
4.23	Experimente – Mesoebene: Juristische Problemfelder ( <i>Kommunale Selbstverwaltung</i> )	286
4.24	Experimente – Mesoebene: Juristische Problemfelder ( <i>Datenschutz</i> )	287
4.25	Experimente – Mesoebene: Juristische Problemfelder ( <i>Rundfunkordnung</i> )	288
4.26	Experimente – Mesoebene: Gemeinschaftserkennung (Gemeinschaftsgrößen)	291
4.27	Experimente – Mesoebene: Gemeinschaftserkennung (Gemeinschaftsgraph)	297
4.28	Experimente – Mesoebene: Gemeinschaftserkennung (Konsensgemeinschaften)	298
4.29	Experimente – Mesoebene: Begriffskontexte ( <i>Europa – zeitlich</i> )	302

4.30	Experimente – Mesoebene: Begriffskontexte ( <i>Europa</i> – thematisch) . . . . .	303
4.31	Experimente – Mesoebene: Begriffskontexte ( <i>Datenschutz</i> – zeitlich) . . . . .	304
4.32	Experimente – Mesoebene: Begriffskontexte ( <i>Datenschutz</i> – thematisch) . . . . .	305
4.33	Experimente – Mesoebene: Begriffskontexte ( <i>Religion</i> – zeitlich) . . . . .	306
4.34	Experimente – Mesoebene: Begriffskontexte ( <i>Christentum</i> – zeitlich) . . . . .	307
4.35	Experimente – Mesoebene: Begriffskontexte ( <i>Judentum</i> – zeitlich) . . . . .	307
4.36	Experimente – Mesoebene: Begriffskontexte ( <i>Islam</i> – zeitlich) . . . . .	308
4.37	Experimente – Mesoebene: Begriffskontexte ( <i>Religion</i> – thematisch) . . . . .	309
4.38	Experimente – Mesoebene: Begriffskontexte ( <i>Christentum</i> – thematisch) . . . . .	310
4.39	Experimente – Mesoebene: Begriffskontexte ( <i>Judentum</i> – thematisch) . . . . .	311
4.40	Experimente – Mesoebene: Begriffskontexte ( <i>Islam</i> – thematisch) . . . . .	312

## Kapitel 1

# Einleitung



Abbildung 1.1: Königsberg in der frühen Neuzeit<sup>1</sup>

Die Untersuchung beginnt in Königsberg (Abb. 1.1, S. 1). Zu Anfang des 18. Jahrhunderts vertrieb man sich dort die Zeit mit einem Rätsel: War es möglich, alle sieben Brücken über den Fluss Pregel zu queren, ohne eine einzige Brücke zweimal gehen zu müssen?

Den Beweis dafür, dass die Antwort negativ lautet, lieferte in den 1730er Jahren der Mathematiker *Leonhard Euler*;<sup>2</sup> die Frage selbst ist heute als *Königsberger*

<sup>1</sup> Kolorierter Kupferstich von Matthäus Merian (1650).

<sup>2</sup> *Euler*, *Commentarii Academiae Scientiarum Petropolitanae* 8 (1741), S. 128–140. *Euler* präsentierte seine Lösung 1735 in St. Petersburg und schrieb sie im Folgejahr auf, der Artikel wurde aber erst 1741 veröffentlicht; *Wilson*, *Journal of Graph Theory* 10 (1986), S. 265 (265); *Alexanderson*, *Bulletin of the American Mathematical Society* 43 (2006), S. 567 (567–568).

Brückenproblem bekannt.<sup>3</sup> Grundlage von *Eulers* Lösung war die in Abbildung 1.2 (S. 2) reproduzierte Darstellung der Stadt. In ihr zeigt sich ein Prinzip, das auch für das Rechtsdenken charakteristisch ist: die *Abstraktion*. *Euler* erkannte, dass es zur Lösung des Problems nicht auf die Länge der Brücken oder die räumliche Distanz zwischen den Stadtteilen ankam. Stattdessen konzentrierte er sich darauf, wie die Stadtteile über die Brücken miteinander verbunden waren.

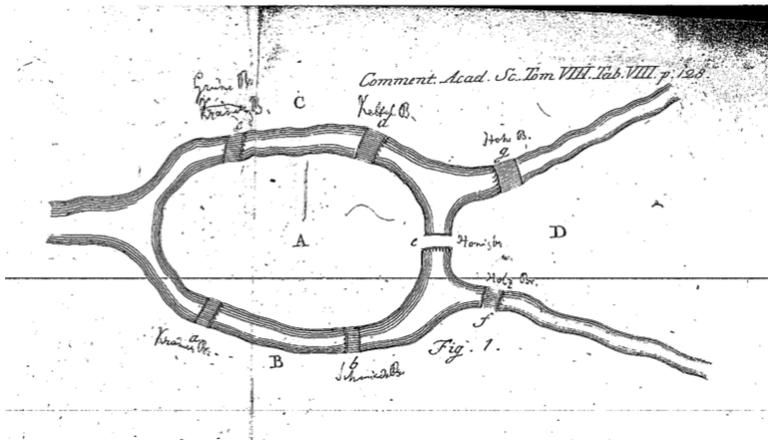


Abbildung 1.2: *Eulers* Darstellung des Königsberger Brückenproblems

Vereinfacht man *Eulers* Darstellung weiter, so erhält man ein mathematisches Objekt, genannt *Graph* (*graph*).<sup>4</sup> Ein Graph besteht aus einer Menge von *Punk-*

<sup>3</sup> Wann sich diese Bezeichnung eingebürgert hat, lässt sich nicht genau feststellen. Sie findet sich ähnlich schon in einem auf den 9. März 1736 datierten Brief des Bürgermeisters von Danzig, *Carl Leonhard Gottlieb Ehler* (partiell reproduziert in *Sachs, Stiebitz und Wilson*, *Journal of Graph Theory* 12 (1988), S. 133 (134–135, Figure 1 und Figure 2)), mit dem dieser *Euler* um „Solutionem Problematis [...] 7 pontium Regiomontanorum“ bittet. *Eulers* Arbeit wurde in den ersten 150 Jahren nach ihrer Veröffentlichung kaum rezipiert, dann aber in einigen Werken zur Freizeitmathematik aufgegriffen; *Wilson*, *Journal of Graph Theory* 10 (1986), S. 265 (272). Unter diesen diskutiert *Lucas*, *Récréations Mathématiques*, 1891, S. 21 „un fameux Mémoire d’Euler, connu sous le nom de *Problème des Ponts de Kœnigsberg*“ (und zitiert für das „Mémoire“ die falsche Fundstelle), während *Rouse*, *Mathematical Recreations and Problems of Past and Present Times*, 1892, S. 122 und *Ahrens*, *Mathematische Unterhaltungen und Spiele*, 1901, S. 317 von „Euler’s problem“ und „Euler’sche[m] Brückenproblem“ sprechen (und so die Person gegenüber dem Ort in den Vordergrund rücken); bei *König*, *Theorie der endlichen und unendlichen Graphen*, 1936, S. 24 heißt es wiederum „*Problem der Königsberger Brücken*“. Arbeiten, die sich mit der Geschichte des Problems beschäftigen (etwa *Wilson*, *Journal of Graph Theory* 10 (1986), S. 265–275, *Hopkins und Wilson*, *The College Mathematics Journal* 35 (2004), S. 198–207 oder *Alexanderson*, *Bulletin of the American Mathematical Society* 43 (2006), S. 567–573), gehen auf die Namensfrage nicht näher ein.

<sup>4</sup> In anderem Kontext mögen dem Leser *Graphen* zur Veranschaulichung mathematischer Funktionen begegnet sein. Die hier interessierenden Graphen kommen ohne Funktionsvorschrift,

ten, auch genannt Knoten (*points, nodes, vertices*), und einer Menge von *Linien*, auch genannt Kanten (*lines, links, arcs, edges*).<sup>5</sup> Den Graphen zum Königsberger Brückenproblem erhält man, indem man für jedes Landstück einen Punkt und für jede Brücke eine Linie definiert.<sup>6</sup> Das lässt sich unmittelbar in eine Zeichnung übersetzen (Abb. 1.3 (a), S. 4).<sup>7</sup> Es kann aber auch in tabellarischer Form festge-

---

x-Achse und y-Achse aus und haben mit den in der Schulmathematik verbreiteten Zeichnungen wenig gemein. Das Sprachproblem des Graphen wiegt im Englischen noch schwerer: Dort kann ein *graph* (unter anderem) den Graphen als mathematisches Objekt, die graphische Darstellung dieses mathematischen Objekts oder auch allgemein jede Form von Graphik (*graph* im Sinne von *diagram* oder *chart*) bezeichnen – der Ausdruck ist mit Bedeutungen überladen. Was gemeint ist, ergibt sich aber in der Regel aus dem Kontext.

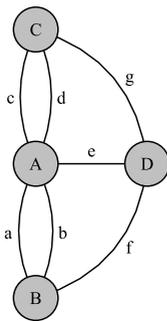
<sup>5</sup> Die Terminologie ist uneinheitlich; wie hier (*points* und *lines*) Harary, *Graph Theory*, 1969, S. 9. Im Deutschen liest man oft von *Knoten* und *Kanten*; diese Wörter haben allerdings zwei Nachteile: Erstens rufen sie beim (nicht in der Mathematik oder der Informatik sozialisierten) Leser Bilder hervor, die zur Bedeutung von *Knoten* und *Kanten* in der Graphentheorie nicht passen. Zweitens unterscheiden sich die Wörter nur um einen Buchstaben (*o* bzw. *a*) und eine Vertauschung von Vokal und Konsonant (*no* bzw. *an*), sodass für den graphentheoretisch wenig bewanderten Leser bei zügiger Lektüre große Verwechslungsgefahr besteht. Das hier gewählte Begriffspaar *Punkte* und *Linien* orientiert sich an der natürlichen Sprache, mit der man die Elemente typischer visueller Repräsentationen von Graphen beschreiben könnte (dazu sogleich).

<sup>6</sup> In manchen Fußnoten werden – wie im Folgenden – mathematische Ausführungen gemacht. Diese können von Lesern, die sich der juristischen Netzwerkforschung zunächst über die Intuition nähern wollen, auch übersprungen werden. Wo mathematische Einzelheiten für das Verständnis des Haupttextes erforderlich sind, werden sie im Haupttext (ggf. nochmals) eingeführt. — *Hintergrundinformationen zur Konstruktion des Graphen zum Königsberger Brückenproblem:* Die Menge der Punkte eines Graphen wird als  $V$  (für *Vertices*), die Menge seiner Linien als  $E$  (für *Edges*) abgekürzt, der Graph selbst ist dann gegeben als  $G = (V, E)$ . Diese Notation ist unabhängig davon, wie die einzelnen Bestandteile von  $V$  und  $E$  in der natürlichen Sprache bezeichnet werden. Die Anzahl der Punkte wird typischerweise mit  $n$ , die Anzahl der Linien mit  $m$  abgekürzt (einzelne Punkte bzw. Linien werden dann abstrakt bezeichnet mit  $v_i$  bzw.  $e_i$  für eine natürliche Zahl  $i$  [*Index*] zwischen 1 und  $n$  bzw.  $m$ ). Formal lässt sich der Graph zum Königsberger Brückenproblem also beschreiben als  $G = (V, E)$  mit Punkten  $V = \{A, B, C, D\}$  und Linien  $E = \{a, b, c, d, e, f, g\}$ , wobei die einzelnen Linien gegeben sind durch  $a = \{A, B\}$ ,  $b = \{A, B\}$ ,  $c = \{A, C\}$ ,  $d = \{A, C\}$ ,  $e = \{A, D\}$ ,  $f = \{B, D\}$ ,  $g = \{C, D\}$ . Die Anzahl der Punkte ist  $n = 4$ , die Anzahl der Linien ist  $m = 7$ . *Geschweifte Klammern* „ $\{ \dots \}$ “ signalisieren, dass es sich um mathematische *Mengen* (*sets*) oder auch *Multimengen* (*multisets*) handelt, bei denen die Reihenfolge der Elemente keine Rolle spielt (z.B. ist  $V = \{A, B, C, D\} = \{D, C, B, A\}$ ). Wo hingegen in mathematischer Notation *runde Klammern* auftauchen „ $( \dots )$ “, handelt es sich in der Regel um *Tupel* (*tuples*), bei denen die Reihenfolge eine Rolle spielt (z.B. ist  $G = (V, E) \neq (E, V)$ ); allerdings werden zum Teil auch runde Klammern zur Angabe von Linien eingesetzt, bei denen die Reihenfolge keine Rolle spielt (Beispiel: *Newman, Networks*, 2010, S. 110). Tupel mit zwei Elementen heißen *Paare* (*pairs*); ein Graph ist also ein Paar von (Multi-)Mengen.

<sup>7</sup> Ähnliche Darstellungen sind in der Lehrbuchliteratur weit verbreitet, so bereits bei *König, Theorie der endlichen und unendlichen Graphen*, 1936, S. 24 Fig. 7; *Harary, Graph Theory*, 1969, S. 2 Fig. 1.2; *Agnarsson und Greenlaw, Graph Theory*, 2007, S. 3. Das laut *Wilson, Journal of Graph Theory* 10 (1986), S. 265 (272), älteste Beispiel findet sich bei *Rouse, Mathematical Recreations and Problems of Past and Present Times*, 1892, S. 123.

halten werden, indem man für jeden Punkt eine Zeile und eine Spalte vorsieht und in den resultierenden Feldern die Anzahl der Linien einträgt, die zwischen den einzelnen Punkten verlaufen (Abb. 1.3 (b), S. 4). Beide Darstellungen machen den Graphen als mathematisches Objekt visuell zugänglich, sind aber vom Graphen selbst zu unterscheiden.

Der Untersuchung von Graphen widmet sich heute ein eigenständiger Zweig der Mathematik, der *Euler* als seinen Vater anerkennt: die *Graphentheorie* (*graph theory*).<sup>8</sup> Auf der Graphentheorie wiederum baut die *Netzwerkforschung* (*network science, science of networks*) und mit ihr die *juristische Netzwerkforschung* (*legal network science*) auf.<sup>9</sup>



(a) Zeichnung

	A	B	C	D
A	0	2	2	1
B	2	0	0	1
C	2	0	0	1
D	1	1	1	0

(b) Tabelle

Abbildung 1.3: Graphdarstellungen des Königsberger Brückenproblems<sup>10</sup>

<sup>8</sup> Siehe etwa *Harary*, *Graph Theory*, 1969, S. 1; *Volkman*, *Fundamente der Graphentheorie*, 1996, S. xvi; *Agnarsson und Greenlaw*, *Graph Theory*, 2007, S. 2; *Krischke und Röpcke*, *Graphen und Netzwerktheorie*, 2015, S. 15. Das Vaterschaftsnarrativ ist wissenschaftshistorisch und wissenschaftssoziologisch interessant, da *Euler* selbst Graphen in seiner Lösung gar nicht erwähnte, sondern das Problem der *geometria situs* im Leibnizschen Sinne zuordnete; dazu *Hopkins und Wilson*, *The College Mathematics Journal* 35 (2004), S. 198–207. Das älteste bekannte Lehrbuch zur Graphentheorie ist *König*, *Theorie der endlichen und unendlichen Graphen*, 1936.

<sup>9</sup> Zum Begriff *network science* *Estrada u.a.* (Hrsg.), *Network Science*, 2010, S. 5. Der Begriff *Netzwerkforschung* hat gegenüber jenem der *Netzwerkanalyse* (*network analysis*) den Vorteil, theoretische, methodologische und anwendungsorientierte Forschungsbeiträge gleichermaßen erfassen zu können; im soziologischen Kontext *Stegbauer und Häußling*, Einleitung in das Handbuch *Netzwerkforschung*, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 13 (13). Andere verstehen den Begriff der Analyse selbst umfassend, so etwa *Marin und Wellman*, *Social Network Analysis: An Introduction*, in: *Scott und Carrington* (Hrsg.), *The SAGE Handbook of Social Network Analysis*, 2011, S. 11 (22), für die Soziale Netzwerkanalyse (*social network analysis, SNA*).

## A. Annäherung

Die Netzwerkforschung trägt ihren Untersuchungsgegenstand im Namen: Sie erforscht Netzwerke. Ein Netzwerk ist dabei die Zusammenfassung einer Menge von Objekten und einer Menge von Beziehungen (zwischen diesen Objekten) zu einer Einheit.<sup>11</sup> Objekte und Beziehungen sind die *Elemente (elements)*, aus denen sich ein Netzwerk zusammensetzt. Generell können diese Elemente ganz unterschiedlicher Natur sein: Bahnhöfe und Gleisabschnitte (*Infrastrukturnetzwerke*), Webseiten und Hyperlinks (*Informationsnetzwerke*),<sup>12</sup> Menschen und ihre sozialen Beziehungen (*soziale Netzwerke*).

Wer nun unter den Beziehungen zwischen Menschen speziell die Verwandtschaftsbeziehungen betrachtet, ist schon ganz nah am Recht: Die Objekte und Beziehungen, die Abbildung 1.4 (S. 6) veranschaulicht und die sich auch im heutigen Familienrecht wiederfinden, konstruiert das Recht, indem es an Lebenssachverhalte anknüpft und biologische Beziehungen nachzeichnet.

Doch die juristische Netzwerkforschung interessiert sich für mehr als die bildliche Darstellung von Verwandtschaftsverhältnissen. Denn Recht strukturiert nicht nur Beziehungen, sondern es wird auch selbst durch Beziehungen strukturiert. Das können Beziehungen zwischen den Personen sein, die miteinander im Rechtssystem interagieren, die Recht setzen, Recht sprechen, *über* Recht sprechen. Es können Beziehungen zwischen den Texten sein, die diese Personen verfassen – Gesetzestexte verweisen auf andere Gesetzestexte, Gerichtsentscheidungen zitieren neben Gesetzestexten auch andere Gerichtsentscheidungen oder (jedenfalls in Deutschland) Beiträge aus der Fachliteratur. Es können sogar Beziehungen innerhalb einzelner Texte sein.

---

<sup>10</sup> Ausführlich zur Darstellung von Graphen noch im dritten Kapitel (S. 109–122), wo auch *Graphdiagramm* und *Adjazenzmatrix* als genauere Bezeichnungen für die hier gezeigten Konstruktionen eingeführt werden.

<sup>11</sup> Newman, Networks, 2010, S. 1; Easley und Kleinberg, Networks, Crowds, and Markets, 2010, S. 2.

<sup>12</sup> Webseiten (*web pages*) sind zu unterscheiden von Websites (*websites*). Eine Webseite (*web page*) ist ein einzelnes Dokument, das über das *World Wide Web* abrufbar und anhand seiner URL (*uniform resource locator*) eindeutig identifizierbar ist (z.B. [www.bundesgerichtshof.de/DE/DasGericht/dasGericht\\_node.html](http://www.bundesgerichtshof.de/DE/DasGericht/dasGericht_node.html)). Eine Website (*website*) ist eine Menge von Webseiten, die unter einer Domäne (*domain*) (z.B. [www.bundesgerichtshof.de](http://www.bundesgerichtshof.de)) abrufbar sind.

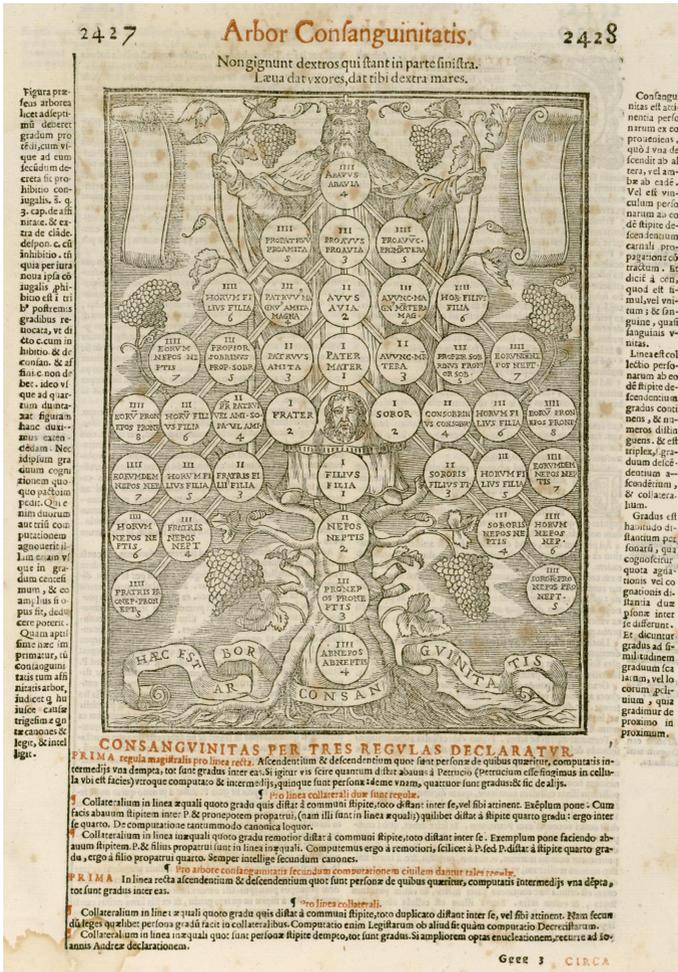


Abbildung 1.4: *Arbor Consanguinitatis* (Baum der Blutsverwandtschaft) im *Corpus Iuris Canonici* (1582)<sup>13</sup>

<sup>13</sup> Decretum Gratiani emendatum et notis illustratum. Gregorii XIII. pont. max. iussu editum. Romae: In aedibus Populi Romani, 1582, Decreti Pars Secunda C. 35 q. 5 col. 2427–2428. Die Objekte sind hier Positionen in der Abstammungshierarchie, zwischen denen das Recht unterscheidet. Die Beziehungen sind konzeptueller Natur; sie ermöglichen die systematische Trennung der Bezeichnungen nach Geschlecht. Die Darstellung hilft, in Kenntnis einer Bezeichnung den Grad der Verwandtschaft zu ermitteln, der durch die Anzahl der Bezeichnungen zwischen dem Gesicht in der Mitte des Bildes und der gesuchten Bezeichnung bestimmt wird und hier bereits eingetragen ist. *Arbores Consanguinitatis* und andere Baumdarstellungen wurden im Mittelalter eingesetzt, um die Strukturen von Rechtsverhältnissen übersichtlich darzustellen. Dazu *Schadt*, Die Darstellungen der Arbores Consanguinitatis und der Arbores Affinitatis, 1982.

Betrachtet man beispielsweise die Gliederungseinheiten, so stellt sich das Bürgerliche Gesetzbuch (BGB) als Baum dar (Abb. 1.5, S. 7).<sup>14</sup> Dieser Baum erweist sich als sehr aussagekräftig, insbesondere dann, wenn man die Paragraphen, die Buchstaben enthalten (ein Indiz für ihre nachträgliche Ergänzung ohne grundlegende Reformen), sowie die direkt zu diesen führenden Zweige rot markiert. Stellt man mehrere Bäume nebeneinander und kontrastiert etwa das BGB mit dem Grundgesetz (GG) (Abb. 1.6, S. 8), so werden auch einige Unterschiede zwischen den dargestellten Gesetzen hinsichtlich ihres Umfangs, ihrer Struktur und ihrer Änderungsanfälligkeit deutlich sichtbar.<sup>15</sup>

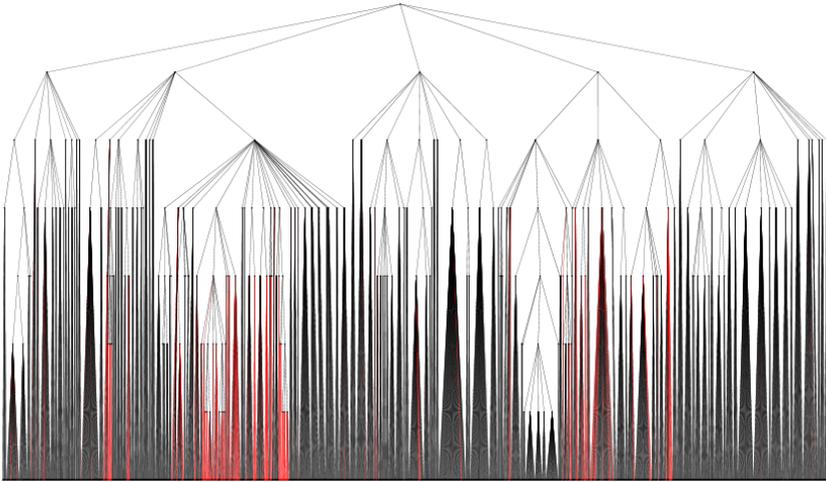
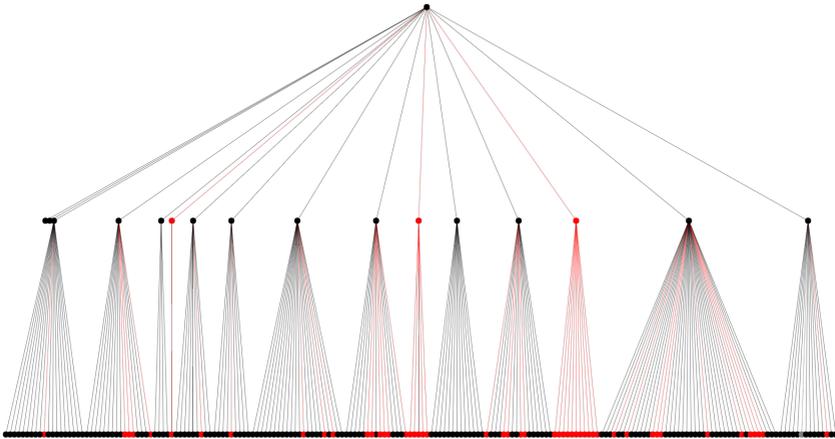


Abbildung 1.5: Bürgerliches Gesetzbuch als Baum<sup>16</sup>

<sup>14</sup> Die nachfolgend abgebildeten Bäume wachsen, ausgehend von genau einem Punkt, der die Wurzel darstellt, *von oben nach unten*. Das ist für Bäume im biologischen Sinn ungewöhnlich, aber für das Verständnis hilfreich, da so das Allgemeine über dem Speziellen steht. Auch der Stammbaum in Abbildung 1.4 (S. 6) wächst sachlich mit der Zeit von oben nach unten (und damit der Wachstumsrichtung des biologischen Baumes, auf den er gezeichnet ist, genau entgegen): Ganz oben stehen mit *abavus* und *abavia* Urugroßvater und Urugroßmutter, ganz unten mit *abnepos* und *abneptis* Ururenkel und Ururenkelin.

<sup>15</sup> Die Bäume in Abbildung 1.5 (S. 7) und Abbildung 1.6 (S. 8) heben die Gliederungsstruktur der Gesetze hervor, stellen allerdings alle Gliederungseinheiten gleich groß dar. Dadurch wird etwa ein Paragraph mit sieben Absätzen ebenso behandelt wie ein Paragraph mit lediglich einem Satz und es gehen viele Informationen verloren, die für eine umfassende Beurteilung des Gesetzesumfangs nützlich sind. Will man die Verteilung des Gesetzesumfangs auf die einzelnen Gliederungselemente visuell hervorheben, so kann man beispielsweise die Größe der einzelnen Gliederungselemente vom Umfang des Textes, der ihnen untergeordnet ist, abhängig machen.

<sup>16</sup> Der Baum wächst hier von der Gesetzesbezeichnung als Wurzel über die fünf von links nach rechts angeordneten Bücher nach unten bis hin zu den einzelnen Paragraphen, die den Text

Abbildung 1.6: Grundgesetz als (kleiner) Baum<sup>17</sup>

Die juristische Netzwerkforschung interessiert sich für die Beziehungen, die Recht schafft, vor allem aber für die Beziehungen, aus denen Recht sich zusammensetzt. Mit diesen Beziehungen und den Objekten, zwischen denen sie angenommen werden, lassen sich Netzwerke definieren. Die juristische Netzwerkforschung untersucht, wie diese Netzwerke strukturiert sind. Sie ergründet, wie die Netzwerke im Recht sich zwischen einzelnen Jurisdiktionen unterscheiden und wie sie sich über die Zeit verändern. An den Gesetzesbäumen wird deutlich, dass die juristische Netzwerkforschung dabei auf die Werkzeuge der Graphentheorie zurückgreifen kann: Abbildung 1.5 (S. 7) und Abbildung 1.6 (S. 8) zeigen Graphen, in denen jeder Punkt für ein Gliederungselement (z.B. Buch, Abschnitt oder Paragraph) und jede Linie für einen Einschluss steht (z.B. enthält Abschnitt I des Grundgesetzes die Artikel 1–19).<sup>18</sup> Zu jedem Netzwerk als Konstrukt in der

---

enthalten. Weggefallene Elemente sind grau markiert, aber so selten, dass sie in der Darstellung kaum sichtbar sind. Jede Gliederungseinheit ist ein Punkt und zwei Gliederungseinheiten sind genau dann durch eine Linie miteinander verbunden, wenn die tieferstehende in der höherstehenden enthalten ist. Viel Bewegung war offenbar im mittleren Teil des besonderen Schuldrechts und in einzelnen Bereichen des Familienrechts. Bei näherer Betrachtung entpuppen sich die gefärbten Stellen im Schuldrecht überwiegend als Richtlinienumsetzungen, im Familienrecht als Reformbemühungen in Unterhaltssachen.

<sup>17</sup> Die nach der Darstellung besonders änderungsanfälligen Bereiche betreffen die Bundesverwaltung, die Verwaltungszusammenarbeit und den Verteidigungsfall; als – in ihrer äußeren Struktur – besonders stabil erweisen sich die Abschnitte über die Grundrechte und die Rechtsprechung.

<sup>18</sup> Gesetzesbäume sind sogar *Bäume (trees)* im graphentheoretischen Sinn: Sie sind *verbunden (connected)* und enthalten *keine Zyklen*, d.h. keine Folgen von Linien, durch deren Nachzeichnen in eine Richtung man den Ausgangspunkt wieder erreichen könnte. Graphen ohne Zyklen heißen *azyklisch*; Harary, Graph Theory, 1969, S. 32.

Realität existiert als Gegenstück (mindestens) ein Graph als Konstrukt in der Mathematik (Tab. 1.1, S. 9).

<i>Realität</i>	<i>Mathematik</i>
Netzwerk	Graph
Objekt	Punkt
Beziehung	Linie

Tabelle 1.1: Grundbegriffe der Netzwerkforschung<sup>19</sup>

Man kann daher Netzwerke im Recht definieren (*Definition*) und Forschungsfragen, die diese betreffen, graphentheoretisch formulieren (*Abstraktion*). Auch können – einen geeigneten Graphen vorausgesetzt – abstrakte Fragen auf Graphen untersucht (*Analyse*) und die Ergebnisse für das Verständnis von Recht in der Realität fruchtbar gemacht werden (*Interpretation*).

Das klingt kompliziert, sodass sich die Frage aufdrängt: Wozu juristische Netzwerkforschung? Eine zufriedenstellende Antwort hierauf setzt voraus, dass man sich zuvor einen Überblick darüber verschafft hat, wie juristische Netzwerkforschung aussehen kann.<sup>20</sup> Dieser Überblick ist Ziel der vorliegenden Arbeit.

<sup>19</sup> Eine ähnliche Übersicht findet sich bei *Barabási*, *Network Science*, 2016, S. 45 Box 2.1. Die hier gegenübergestellten Begriffe werden in der Netzwerkforschung oft synonym verwendet. Das hat den Nachteil, dass der Übersetzungsvorgang von der Realität in die Mathematik aus der Sprache nicht mehr deutlich wird. Da gerade dieser Übersetzungsvorgang oft problematisch ist (dazu später mehr), versucht diese Arbeit, die Begriffswelten getrennt zu halten.

<sup>20</sup> Vermutlich lässt sich die juristische Netzwerkforschung auch rechtstheoretisch fundieren. Überlegungen hierzu könnten in den Bereich der *komplexen adaptiven Systeme* (*complex adaptive systems*, CAS) führen, zu deren Erforschung nicht selten auf Netzwerke zurückgegriffen wird. Leicht zugängliche Einführungen bieten etwa *Miller und Page*, *Complex Adaptive Systems*, 2007, und *Mitchell*, *Complexity*, 2011. Erwägungen zur Übertragung von Konzepten aus der Theorie komplexer adaptiver Systeme auf das Recht, allerdings überwiegend metaphorischen Charakters, finden sich prominent in den Arbeiten von *Ruhl*, etwa *Ruhl*, *Duke Law Journal* 45 (1996), S. 849–928; *Ruhl*, *Vanderbilt Law Review* 49 (1996), S. 1406–1490; *Ruhl*, *Houston Law Review* 34 (1997), S. 933–1002; *Ruhl*, *BYU Law Review* 1997, S. 777–801; *Ruhl und Ruhl*, *U.C. Davis Law Review* 30 (1997), S. 405–482; *Ruhl*, *Georgia State University Law Review* 24 (2008), S. 885–911; *Ruhl und Katz*, *Iowa Law Review* 101 (2015), S. 191–244; *Ruhl, Katz und Bommarito*, *Science* 355 (2017), S. 1377–1378; in deutscher Sprache und insgesamt deutlich zurückhaltender bei *Zollner*, *Komplexität und Recht*, 2014. Weitere Beispiele aus dem amerikanischen Diskurs: *Balkin*, *Rutgers Law Review* 39 (1986), S. 1–103; *Reynolds*, *Columbia Law Review* 91 (1991), S. 110–117; *Scott*, *William & Mary Law Review* 35 (1993), S. 329–351; *Di Lorenzo*, *Yale Law & Policy Review* 12 (1994), S. 425–485; *Geu*, *Tennessee Law Review* 61 (1994), S. 933–990; *Roe*, *Harvard Law Review* 109 (1996), S. 641–668; *Di Lorenzo*, *William & Mary Law Review* 38 (1997), S. 1729–1815; *Post und Johnson*, *Chicago-Kent Law Review* 1998, S. 1055–1099; *Post und Eisen*, *The Journal of Legal Studies* 29 (2000), S. 545–584; *Jones*, *Georgia State University Law Review* 24 (2008), S. 873–883; *Gewirtzman*, *American University Law Review* 61 (2012), S. 457–522; *Shur-Ofry*, *The Intellectual Property Law Review* 54 (2013), S. 55–102.

## B. Aufbau

Die Arbeit dient der Einführung in die juristische Netzwerkforschung. Sie beginnt abstrakt und wird konkret. Den Kern der Untersuchung bilden drei Kapitel:

*Kapitel 2 — Theorie* öffnet mit einem Streifzug durch die existierende Literatur zur Netzwerkforschung im Recht. Dieser Streifzug weckt das Bedürfnis nach Strukturierung, sowohl mit Blick auf den Gegenstand der juristischen Netzwerkforschung als auch mit Blick auf die Herangehensweise derer, die sie betreiben. Im Anschluss werden daher einige Unterscheidungen eingeführt, die eine systematische Erschließung der juristischen Netzwerkforschung als rechtswissenschaftliches Forschungsfeld erleichtern.

*Kapitel 3 — Methodik* behandelt mögliche Werkzeuge der juristischen Netzwerkforschung. Nach einigen Überlegungen zur Arbeit mit juristischen Daten werden hier Visualisierung, Quantifizierung und Modellierung als charakteristische Herangehensweisen vorgestellt, bevor Mehrschichtigkeit als methodische Herausforderung adressiert wird. Anhand fiktiver juristischer Daten wird erläutert, wo die grundlegenden Konzepte der Netzwerkforschung in der *juristischen* Netzwerkforschung eingesetzt werden können.

*Kapitel 4 — Praxis* illustriert die Möglichkeiten und Schwierigkeiten der juristischen Netzwerkforschung anhand einer Fallstudie zur Rechtsprechung des Bundesverfassungsgerichts (BVerfG). Hier wird der Forschungsprozess der juristischen Netzwerkforschung nachgezeichnet, von der Zusammenstellung realer juristischer Daten über ihre Analyse bis hin zur Evaluation der Ergebnisse. Im Rahmen einer Reihe von Experimenten werden mehrere der im zweiten Kapitel eingeführten Differenzierungen wieder aufgegriffen und einige der im dritten Kapitel vorgestellten Methoden erprobt.

*Kapitel 5 — Fazit* fasst die Ergebnisse der Arbeit zusammen und zeigt einige Potenziale für zukünftige Forschung auf. Die Daten und Computerprogramme, die den Analysen in Teilen der Arbeit zugrunde liegen, sind in einem *Online-Appendix* zusammengestellt, dessen Struktur im *Appendix* wiedergegeben ist. Zum Nachschlagen von Fachbegriffen enthält die Arbeit außerdem ein *Glossar*.

## Kapitel 2

# Theorie

In diesem Kapitel werden die theoretischen Grundlagen der juristischen Netzwerkforschung skizziert. Mit welchen Objekten und Beziehungen lassen sich Netzwerke im Recht definieren und welche Probleme können dabei auftreten (B.)? Aus welchen Blickwinkeln und mit welchen Zielen kann man Netzwerke im Recht betrachten (C.)? Die Literatur zur Netzwerkforschung im Recht liefert anschauliche Teilantworten auf beide Fragen; um als *juristische* Netzwerkforschung bezeichnet werden zu können, fehlt es ihr aber an einem rechtswissenschaftlichen Rahmen (A.).

### A. Literatur

Was sind die wichtigsten Entscheidungen des Europäischen Gerichtshofs? Wie komplex ist der *U.S. Code*? Welche Staaten schließen ähnliche bilaterale Investitionsabkommen? — Wie organisieren sich konservative Anwälte in den USA? Wer sorgt für den Erfolg von Gesetzen im *U.S. Congress*? Woher nehmen europäische Richter ihre Informationen? — Wie verbreiten sich *Corporate Governance*-Praktiken in Unternehmen oder finanzielle Schocks auf Finanzmärkten? Wodurch lässt sich Kriminalität effektiv bekämpfen? Welcher Zusammenhang besteht zwischen Patentrecht und Innovation?

Diese Fragen haben zweierlei gemeinsam: Erstens betreffen sie im weitesten Sinne das Recht. Zweitens sind sie bereits mit Werkzeugen aus der Netzwerkforschung untersucht worden. Damit ist auch klar, was gemeint ist, wenn hier von Literatur zur Netzwerkforschung im Recht gesprochen wird: *Literatur, die mit Werkzeugen aus der Netzwerkforschung Fragen untersucht, die das Recht betreffen*.<sup>1</sup> Davon gibt es mittlerweile mehr, als man vermuten mag. Denn die Definition erfasst Beiträge, die neben einem grob umrissenen Gegenstand (Recht) vor allem eine Herangehensweise (Netzwerkforschung) teilen. Diese Beiträge stammen aus

---

<sup>1</sup> Im Einzelfall mag zweifelhaft sein, ob eine untersuchte Frage das Recht betrifft oder ob ein Instrument der Netzwerkforschung zuzurechnen ist. Hier werden die kritischen Begriffe weit ausgelegt.

den verschiedensten Disziplinen und werden in den unterschiedlichsten Medien publiziert.<sup>2</sup>

Die Verfasserin dieser Arbeit ist durch einen Aufsatz in der Zeitschrift *Political Analysis* auf die Netzwerkforschung im Recht aufmerksam geworden.<sup>3</sup> Der Beitrag hat fünf Autoren, allesamt Politikwissenschaftler; der erste Eintrag im Literaturverzeichnis entstammt den *Reviews of Modern Physics*.<sup>4</sup> Gegenstand des Artikels sind die Entscheidungen des *U.S. Supreme Court*. Die geschilderte Konstellation ist für die Netzwerkforschung im Recht typisch. Diese ist häufig *doppelt interdisziplinär*: Einmal *als Netzwerkforschung*, die *per se* Wissenschaftler aus vielen verschiedenen Disziplinen anzieht,<sup>5</sup> und dann *als Forschung im Recht*,

<sup>2</sup> Ein Einblick in die Vielfalt der Publikationsmedien: **Klassische Law Reviews und Law Journals:** *Smith*, San Diego Law Review 44 (2007), S. 309–354; *Derlén und Lindholm*, European Law Journal 20 (2014), S. 667–687; *Whalen*, Michigan State Law Review 2016, S. 539–565. — **Auf Empirical Legal Studies spezialisierte Zeitschriften:** *Paik, Southworth und Heinz*, Law & Social Inquiry 32 (2007), S. 883–917; *Lupu und Fowler*, The Journal of Legal Studies 42 (2013), S. 151–186; *Black und Spriggs*, Journal of Empirical Legal Studies 10 (2013), S. 325–358. — **Rechtinformatik:** *Boulet, Mazzega und Bourcier*, Artificial Intelligence and Law 19 (2011), S. 333–355; *Winkels und de Ruyter*, Survival of the Fittest: Network Analysis of Dutch Supreme Court Cases, in: *Palmirani u.a.* (Hrsg.), AI Approaches to the Complexity of Legal Systems, 2012, S. 106–115; *Panagis und Šadl*, The Force of EU Case Law: A Multi-Dimensional Study of Case Citations, in: *Rotolo* (Hrsg.), Legal Knowledge and Information Systems, 2015, S. 71–80. — **Politikwissenschaft:** *Tam Cho und Fowler*, The Journal of Politics 72 (2010), S. 124–135; *Lupu und Voeten*, British Journal of Political Science 42 (2012), S. 413–439; *Pelc*, American Political Science Review 108 (2014), S. 547–564. — **Sozialwissenschaften:** *Caldeira*, Social Networks 10 (1988), S. 29–55; *Kogut und Walker*, American Sociological Review 66 (2001), S. 317–335; *Fowler und Jeon*, Social Networks 30 (2008), S. 16–30. — **Wirtschaftswissenschaften:** *Davis*, Administrative Science Quarterly 36 (1991), S. 583–613; *Battiston u.a.*, Journal of Economic Dynamics & Control 36 (2012), S. 1121–1141; *Siems*, Journal of Institutional Economics 12 (2016), S. 579–602. — **Naturwissenschaften:** *Bommarito und Katz*, Physica A 389 (2010), S. 4195–4200; *Tarissan und Nollez-Goldbach*, Journal of Complex Networks 4 (2016), S. 616–634; *Ruhl, Katz und Bommarito*, Science 355 (2017), S. 1377–1378.

<sup>3</sup> *Fowler u.a.*, Political Analysis 15 (2007), S. 324–346.

<sup>4</sup> *Albert und Barabási*, Reviews of Modern Physics 74 (2002), S. 47–97 (Statistical Mechanics of Complex Networks); im Literaturverzeichnis bei *Fowler u.a.*, Political Analysis 15 (2007), S. 324 (344).

<sup>5</sup> Eine lange Tradition hat die Netzwerkforschung in der Soziologie, wo sie als Soziale Netzwerkanalyse (*social network analysis*, SNA) bezeichnet wird; zur Geschichte *Freeman*, The Development of Social Network Analysis, 2004; *Freeman*, The Development of Social Network Analysis – With an Emphasis on Recent Events, in: *Scott und Carrington* (Hrsg.), The SAGE Handbook of Social Network Analysis, 2011, S. 26–39. Pionierarbeit leistete insbesondere *Moreno*, ein studierter Mediziner, dessen Hauptwerk, *Moreno*, Who Shall Survive?, 1934, auch heute noch überaus lesenswert ist. Seit den 1990er Jahren treiben verstärkt auch Naturwissenschaftler, Mathematiker und Informatiker die Netzwerkforschung voran, die sich seitdem einer regen wissenschaftsjournalistischen Berichterstattung erfreut; siehe etwa *Watts und Strogatz*, Nature 393 (1998), S. 440–442; *Barabási und Albert*, Science 286 (1999), S. 509–512; *Kleinberg*, Journal of the ACM 46 (1999), S. 604–632; *Albert, Jeong und Barabási*, Nature 406 (2000), S. 378–382; *Kleinberg*, Nature 406 (2000), S. 845; *Albert und Barabási*, Reviews of Modern

die durch Wissenschaftler betrieben wird, welche oft nicht juristisch ausgebildet sind. Vor diesem Hintergrund ist es schwer, den Forschungsstand vollständig zu erfassen. Zur Motivation der juristischen Netzwerkforschung genügt es aber, ein Verständnis für die Kontexte zu entwickeln, in denen Netzwerkforschung im Recht auftaucht. Viele dieser Kontexte sind in den Fragen zu Beginn dieses Abschnitts bereits angeklungen. Sie betreffen juristische Texte (I.), juristische Akteure (II.) und die Wechselwirkungen zwischen Recht und Realität (III.).

### I. Juristische Texte

Studien, die sich mit juristischen Texten auseinandersetzen, interessieren sich oft für *Gerichtsentscheidungen*. Viel Aufmerksamkeit erfahren dabei traditionell die Entscheidungen des *U.S. Supreme Court*.<sup>6</sup> Daneben existiert eine Reihe von Arbeiten zu anderen US-amerikanischen Gerichten und zu Gerichten außerhalb der USA, insbesondere zu supranationalen und internationalen Gerichten.<sup>7</sup> Die

---

Physics 74 (2002), S. 47–97; Newman, SIAM Review 45 (2003), S. 167–256; zum Verhältnis dieses Stranges der Netzwerkforschung zur Sozialen Netzwerkanalyse Watts, Annual Review of Sociology 30 (2004), S. 243–270; Stegbauer, „Die Invasion der Physiker“: Naturwissenschaft und Soziologie in der Netzwerkanalyse, in: *Deutsche Gesellschaft für Soziologie* (Hrsg.), Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006, 2008, S. 1060–1077. Zur Netzwerkforschung in der Politikwissenschaft Lazer, PS: Political Science & Politics 44 (2011), S. 61–68; Ward, Stovel und Sacks, Annual Review of Political Science 14 (2011), S. 246–264; in der Ökonomie Kenett und Havlin, Mind & Society 14 (2015), S. 155–167; sowie die Beiträge in Bramoullé, Galeotti und Rogers (Hrsg.), The Oxford Handbook of the Economics of Networks, 2016. Auch Archäologen und Historiker betreiben Netzwerkforschung, siehe etwa Düring und Eumann, Geschichte und Gesellschaft 39 (2013), S. 369–390; Knappett, Network Analysis in Archaeology, 2013; Morrissey, Historical Methods 48 (2015), S. 67–79; sowie die Beiträge in Brughmans, Collar und Coward (Hrsg.), The Connected Past, 2016.

<sup>6</sup> Beispiele **in sozial- und politikwissenschaftlichen Publikationsmedien**: Fowler u.a., Political Analysis 15 (2007), S. 324–346; Fowler und Jeon, Social Networks 30 (2008), S. 16–30; Bommarito, Katz und Zelner, Procedia Social and Behavioral Sciences 2010, S. 26–37; Clark, Lauderdale und Katz, Political Analysis 20 (2012), S. 329–350. — **In naturwissenschaftlichen Publikationsmedien**: Chandler, The Mathematica Journal 10 (2007), S. 501–526; Leicht u.a., The European Physical Journal B 59 (2007), S. 75–83; Karrer und Newman, Physical Review E 80 (2009), S. 046110-1–046110-12; Bommarito u.a., Physica A 389 (2010), S. 4201–4208; Whalen, Modeling Annual Supreme Court Influence: The Role of Citation Practices and Judicial Tenure in Determining Precedent Network Growth, in: Menezes, Evsukoff und González (Hrsg.), Complex Networks, 2013, S. 169–176; Clough u.a., Journal of Complex Networks 3 (2015), S. 189–203. — **In rechtswissenschaftlichen Publikationsmedien**: Cross, Smith und Tomarchio, Emory Law Journal 57 (2008), S. 1227–1257; Cross u.a., University of Illinois Law Review 2010, S. 489–575; Black und Spriggs, Journal of Empirical Legal Studies 10 (2013), S. 325–358; Lupu und Fowler, The Journal of Legal Studies 42 (2013), S. 151–186; Whalen, Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Supreme Court Precedent, in: Winkels, Lettieri und Faro (Hrsg.), Network Analysis in Law, 2014, S. 1–21.

<sup>7</sup> **Verschiedene US-amerikanische Gerichte**: Post und Eisen, The Journal of Legal Studies 29 (2000), S. 545–584; Smith, San Diego Law Review 44 (2007), S. 309–354; Dibadj, San

Arbeiten versuchen häufig, wichtige Entscheidungen zu identifizieren, die globale Struktur des Fallrechts zu beschreiben oder dessen sachliche Organisation aufzudecken. Dabei verlassen sie sich nicht auf die persönliche Lektüre einzelner Entscheidungen oder den Eindruck, der sich aus dem Schrifttum ergibt. Stattdessen sammeln sie möglichst *alle* potenziell relevanten Entscheidungen und extrahieren aus diesen Zitate. So können sie Netzwerke konstruieren, in denen jede Entscheidung ein Objekt und jedes Zitat eine Beziehung darstellt, und die ihnen entsprechenden Graphen mathematisch analysieren.<sup>8</sup>

---

Diego Law Review 45 (2008), S. 1–32; *Dibadj*, San Diego Law Review 46 (2009), S. 1–26; *Bommarito, Katz und Isaacs-See*, Virginia Tax Review 30 (2011), S. 523–557. — **Europäischer Gerichtshof:** *Derlén und Lindholm*, Europarättslig Tidskrift 2012, S. 462–481; *Mirshahvalad u.a.*, PLoS ONE 7 (2012), e33721-1–e33721-7; *Derlén u.a.*, Europarättslig Tidskrift 2013, S. 517–535; *Schaper*, The Structure and Organization of EU Law in the Field of Direct Taxes, 2013, S. 143–272; *Derlén und Lindholm*, European Law Journal 20 (2014), S. 667–687; *Derlén und Lindholm*, German Law Journal 16 (2015), S. 1073–1098; *Šadl*, European Journal of Legal Studies 8 (2015), S. 18–45; *Panagis und Šadl*, The Force of EU Case Law: A Multi-Dimensional Study of Case Citations, in: *Rotolo* (Hrsg.), Legal Knowledge and Information Systems, 2015, S. 71–80; *Tarissan, Panagis und Šadl*, IEEE/ACM ASONAM 2016, S. 661–668; *Derlén und Lindholm*, German Law Journal 18 (2017), S. 647–686 (mit Vergleich zum *U.S. Supreme Court*); *Šadl und Olsen*, Leiden Journal of International Law 30 (2017), S. 327–349. — **Europäischer Gerichtshof für Menschenrechte:** *Lupu und Voeten*, British Journal of Political Science 42 (2012), S. 413–439; *Christensen, Olsen und Tarissan*, Identification of Case Content with Quantitative Network Analysis: An Example from the ECtHR, in: *Bex und Villata* (Hrsg.), Legal Knowledge and Information Systems, 2016, S. 53–62; *Olsen und Küçüksu*, International Journal of Discrimination and the Law 17 (2017), S. 4–22; *Šadl und Olsen*, Leiden Journal of International Law 30 (2017), S. 327–349. — **Verschiedene oberste nationale Gerichte in Europa:** *Gelter und Siems*, Utrecht Law Review 8 (2005), S. 88–99; *Siems*, King’s Law Journal 21 (2010), S. 152–171; *Gelter und Siems*, Supreme Court Economic Review 21 (2013), S. 215–270; *Gelter und Siems*, The American Journal of Comparative Law 62 (2014), S. 35–86. — **Internationaler Strafgerichtshof:** *Tarissan und Nollez-Goldbach*, The Network of the International Criminal Court Decisions as a Complex System, in: *Sanayei, Zelinka und Rossler* (Hrsg.), Emergence, Complexity and Computation, 2013, S. 255–264; *Tarissan und Nollez-Goldbach*, Temporal Properties of Legal Decision Networks: A Case Study from the International Criminal Court, in: *Rotolo* (Hrsg.), Legal Knowledge and Information Systems, 2015, S. 111–120; *Tarissan und Nollez-Goldbach*, Journal of Complex Networks 4 (2016), S. 616–634. — **Italien:** *Agnoloni und Pagallo*, The Power Laws of the Italian Constitutional Court, and Their Relevance for Legal Scholars, in: *Rotolo* (Hrsg.), Legal Knowledge and Information Systems, 2015, S. 1–10. — **Kanada:** *Neale*, Journal of Open Access to Law 1 (2013), S. 1–60. — **Niederlande:** *Winkels, de Ruyter und Kroese*, Determining Authority of Dutch Case Law, in: *Atkinson* (Hrsg.), Frontiers in Artificial Intelligence and Applications, 2011, S. 103–112; *Winkels und de Ruyter*, Survival of the Fittest: Network Analysis of Dutch Supreme Court Cases, in: *Palmirani u.a.* (Hrsg.), AI Approaches to the Complexity of Legal Systems, 2012, S. 106–115. — **WTO-Streitschlichtung:** *Pelc*, American Political Science Review 108 (2014), S. 547–564.

<sup>8</sup> Bei näherer Betrachtung erweist sich dieser Strang der Netzwerkforschung im Recht als eine Variante der juristischen Zitationsanalyse. Dies zeigt sich insbesondere in einigen jüngeren Studien, die klassische Statistik mit Netzwerkmaßen kombinieren, wie etwa *Cross u.a.*, University of Illinois Law Review 2010, S. 489–575, und *Black und Spriggs*, Journal of Empirical

Eine weitere Gruppe von Studien untersucht *Gesetzestexte*.<sup>9</sup> Dabei werden verschiedene Interessen verfolgt: Manchen geht es darum, die Komplexität von Gesetzestexten sichtbar zu machen und in Zahlen auszudrücken;<sup>10</sup> anderen geht es darum, das Recht in der Recherche besser navigierbar zu machen;<sup>11</sup> wiederum andere wollen schlicht sehen, was man über die Texte mit den Instrumenten der Netzwerkforschung lernen kann.<sup>12</sup> Dazu werden zum Einen die hierarchischen Strukturen einzelner Gesetzestexte (Inklusionsbeziehungen) und zum Anderen die Verweise einzelner Normen aufeinander (Verweisungsbeziehungen) betrachtet.<sup>13</sup>

---

Legal Studies 10 (2013), S. 325–358, aber auch schon bei *Marx*, Jurimetrics Journal 10 (1970), S. 121–137, dessen Arbeit das Potenzial der Netzwerkperspektive für die juristische Recherche hervorhebt. Zur Tradition der Zitationsanalyse in den USA *Shapiro*, Journal of the American Society for Information Science 43 (1992), S. 337–339; ältere Beispiele für juristische Zitationsanalysen sind *Merryman*, Stanford Law Review 6 (1954), S. 613–673; *Caldeira*, The American Political Science Review 79 (1985), S. 178–194; *Landes, Lessig und Solimine*, The Journal of Legal Studies 27 (1998), S. 271–332. In Deutschland ist die juristische Zitationsanalyse kaum verbreitet und die Arbeiten, die speziell die Rechtsprechung betreffen, stammen aus den 1990er Jahren: *Wagner-Döbler und Philipps*, Rechtstheorie 23 (1992), S. 228–241; *Wagner-Döbler*, Scientometrics 29 (1994), S. 15–26; *Wagner-Döbler*, RabelsZ 59 (1995), S. 113–127.

<sup>9</sup> So unter anderem *Bourcier und Mazzega*, Codification, Law Article and Graphs, in: *Lodder und Mommers* (Hrsg.), Legal Knowledge and Information Systems, 2007, S. 29–38; *Bommarito und Katz*, Physica A 389 (2010), S. 4195–4200; *Boulet, Mazzega und Bourcier*, Network Analysis of the French Environmental Code, in: *Casanovas u.a.* (Hrsg.), AI Approaches to the Complexity of Legal Systems, 2010, S. 39–53; *Boulet, Mazzega und Bourcier*, Artificial Intelligence and Law 19 (2011), S. 333–355; *Guclu u.a.*, Public Health Reports 129 (2014), S. 154–165; *Katz und Bommarito*, Artificial Intelligence and Law 22 (2014), S. 337–374; *Sweeney u.a.*, Network Analysis of Manually-Encoded State Laws and Prospects for Automation, in: *Winkels, Lettieri und Faro* (Hrsg.), Network Analysis in Law, 2014, S. 53–77; *Winkels, Boer und Plantevin*, Creating Context Networks in Dutch Legislation, in: *Ashley* (Hrsg.), Legal Knowledge and Information Systems, 2014, S. 155–164; *Li u.a.*, Journal of Business and Technology Law 10 (2015), S. 297–374; *Ruhl und Katz*, Iowa Law Review 101 (2015), S. 191–244.

<sup>10</sup> *Bourcier und Mazzega*, ICAIL Proceedings 2007, S. 211–215; *Bourcier und Mazzega*, Codification, Law Article and Graphs, in: *Lodder und Mommers* (Hrsg.), Legal Knowledge and Information Systems, 2007, S. 29–38; *Katz und Bommarito*, Artificial Intelligence and Law 22 (2014), S. 337–374; *Li u.a.*, Journal of Business and Technology Law 10 (2015), S. 297–374; *Ruhl und Katz*, Iowa Law Review 101 (2015), S. 191–244.

<sup>11</sup> *Liiv, Vedeshin und Täks*, ICAIL Proceedings 2007, S. 189–190; *Winkels, Boer und Plantevin*, Creating Context Networks in Dutch Legislation, in: *Ashley* (Hrsg.), Legal Knowledge and Information Systems, 2014, S. 155–164.

<sup>12</sup> *Boulet, Mazzega und Bourcier*, Network Analysis of the French Environmental Code, in: *Casanovas u.a.* (Hrsg.), AI Approaches to the Complexity of Legal Systems, 2010, S. 39–53; *Boulet, Mazzega und Bourcier*, Artificial Intelligence and Law 19 (2011), S. 333–355.

<sup>13</sup> Wer sich auf die hierarchischen Strukturen von Gesetzestexten konzentriert, erhält Bäume, die jenen aus Abbildung 1.5 (S. 7) und Abbildung 1.6 (S. 8) ähneln; siehe etwa *Katz und Bommarito*, Artificial Intelligence and Law 22 (2014), S. 337 (343 Fig. 1, 348 Fig. 2). Für eine kombinierte Darstellung von hierarchischen Strukturen und Verweisstrukturen siehe *Bommarito und Katz*, Physica A 389 (2010), S. 4195 (4197 Fig. 1).

Die Netzwerke, die diese Untersuchungen definieren, bestehen dementsprechend aus den Gliederungselementen der Texte (Objekte) und ihren Inklusions- oder Verweisungsbeziehungen (Beziehungen).<sup>14</sup>

Schließlich gibt es einige Arbeiten, die sich mit internationalen Verträgen auseinandersetzen.<sup>15</sup> Sie versuchen unter anderem, die Entwicklung einzelner Bereiche des internationalen Rechts nachzuzeichnen,<sup>16</sup> die Zusammenhänge zwischen einzelnen Verträgen aufzuzeigen<sup>17</sup> oder die wirtschaftlichen Konsequenzen der Verträge zu modellieren.<sup>18</sup> Dazu untersuchen die Autoren, welche Verträge zwischen einzelnen Staaten geschlossen werden, wie diese einander ähneln oder welche Klauseln sie enthalten.<sup>19</sup> Auch erkunden sie, welche multilateralen Abkommen von welchen Staaten in welcher Reihenfolge ratifiziert werden<sup>20</sup> oder welche dieser Abkommen aufeinander Bezug nehmen.<sup>21</sup> Die Netzwerke, die hier definiert werden, bestehen aus Staaten (Objekte) und Verträgen oder Ratifikationen (Beziehungen) – oder aber aus Verträgen (Objekte) und staatlichen Ratifikationshandlungen (Beziehungen).

Die skizzierten Studien unterscheiden sich in Stil und Detailgrad der Darstellung zum Teil sehr stark, ihre grundlegenden Herangehensweisen sind einander jedoch ähnlich. Als Beispiel sei noch einmal die eingangs erwähnte Studie zur Rechtsprechung des *U.S. Supreme Court* betrachtet. In ihr gehen die Autoren vor allem einer Frage nach: „how can we identify the most legally central cases at the U.S. Supreme Court at a given point in time in an empirically rigorous, reliable, and valid manner?“<sup>22</sup> Um diese Frage zu beantworten, erstellen sie „the complete

<sup>14</sup> Einen etwas anderen Ansatz verfolgen *Guclu u.a.*, *Public Health Reports* 129 (2014), S. 154–165, und *Sweeney u.a.*, *Network Analysis of Manually-Encoded State Laws and Prospects for Automation*, in: *Winkels, Lettieri und Faro* (Hrsg.), *Network Analysis in Law*, 2014, S. 53–77, die aus dem Wortlaut von Gesetzestexten die rechtlich vorgesehenen Beziehungen zwischen juristischen Akteuren ableiten.

<sup>15</sup> Beispiele: *Saban, Bonomo und Stier-Moses*, *Physica A* 389 (2010), S. 3661–3673; *Kim*, *Global Environmental Change* 23 (2013), S. 980–991; *Polak*, *Algorithms for the Network Analysis of Bilateral Tax Treaties*, 2014; *van't Riet und Lejour*, *CPB Discussion Paper* 2014, S. 1–68; *Boulet, Barros-Platiau und Mazzega*, *Artificial Intelligence and Law* 2016, S. 133–148; *Alschner und Skougarevskiy*, *Journal of International Economic Law* 19 (2016), S. 561–588.

<sup>16</sup> *Saban, Bonomo und Stier-Moses*, *Physica A* 389 (2010), S. 3661–3673; *Kim*, *Global Environmental Change* 23 (2013), S. 980–991; *Boulet, Barros-Platiau und Mazzega*, *Artificial Intelligence and Law* 2016, S. 133–148.

<sup>17</sup> *Alschner und Skougarevskiy*, *Journal of International Economic Law* 19 (2016), S. 561–588.

<sup>18</sup> *Polak*, *Algorithms for the Network Analysis of Bilateral Tax Treaties*, 2014; *van't Riet und Lejour*, *CPB Discussion Paper* 2014, S. 1–68.

<sup>19</sup> *Polak*, *Algorithms for the Network Analysis of Bilateral Tax Treaties*, 2014; *van't Riet und Lejour*, *CPB Discussion Paper* 2014, S. 1–68; *Saban, Bonomo und Stier-Moses*, *Physica A* 389 (2010), S. 3661–3673; *Alschner und Skougarevskiy*, *Journal of International Economic Law* 19 (2016), S. 561–588.

<sup>20</sup> *Boulet, Barros-Platiau und Mazzega*, *Artificial Intelligence and Law* 2016, S. 133–148.

<sup>21</sup> *Kim*, *Global Environmental Change* 23 (2013), S. 980–991.

<sup>22</sup> *Fowler u.a.*, *Political Analysis* 15 (2007), S. 324 (325).

network of 26,681 majority opinions written by the U.S. Supreme Court and the cases that cite them from 1791 to 2005<sup>23</sup>. Dabei wird jede *Court Opinion* zum Objekt, und eine Beziehung zwischen zwei Entscheidungen wird genau dann angenommen, wenn eine Entscheidung die andere in ihrer *Majority Opinion* zitiert.<sup>24</sup> Die so definierten Beziehungen haben eine Richtung (sie verlaufen jeweils von der zitierenden zur zitierten Entscheidung) und werden im Graphen als *Pfeile* bzw. gerichtete Linien (*directed lines, directed links, directed edges*) dargestellt.

Die Autoren nutzen nun Werkzeuge der Graphentheorie, um das Konzept der rechtlichen Relevanz (*legal relevance*)<sup>25</sup> quantitativ handhabbar zu machen. Ausgehend von einigen Standardkonzepten aus der Netzwerkforschung, auf die im dritten Kapitel eingegangen wird,<sup>26</sup> entwickeln sie neue Maße, welche den rechtlichen Einfluss einzelner Entscheidungen des *U.S. Supreme Court* einerseits und deren Verankerung im Fallrecht andererseits besser als bisherige Maße quantifizieren sollen.<sup>27</sup> Diese Maße werden an ausgewählten Entscheidungen zur Abtreibung (*landmark abortion decisions*) getestet und genutzt, um herauszufinden, welche Entscheidungen im Jahr 2005 die größte rechtliche Relevanz haben.<sup>28</sup> Mithilfe statistischer Verfahren wird außerdem geprüft, inwieweit die entwickelten Maße in der Lage sind, zu erklären, wie eine Entscheidung nach ihrer Veröffentlichung zitiert wird.<sup>29</sup> Die Arbeit enthält fünf Graphiken, zwei Gleichungen und vier Tabellen. Ihre Verfasser sehen sie als quantitativen Beitrag zur Rechtstheorie, insbesondere zur Theorie der Rechtsentwicklung, und gleichzeitig als Beitrag zum allgemeinen Netzwerkforschungsdiskurs.<sup>30</sup>

Das Vorgehen der geschilderten Studie ist in vielerlei Hinsicht charakteristisch für die Studien, die sich mit Texten im Recht befassen. Diese beginnen nicht selten mit der Sammlung und Aufbereitung von Daten und enden in Publikationen mit vielen visuellen und quantitativen Elementen. Die Datenquellen, auf die zugegriffen werden soll, sind bereits digital verfügbar; Datensammlung und Datenaufbereitung verlaufen ganz oder teilweise automatisiert; die Datenanalyse erfolgt computergestützt. Man fragt, wie das Recht in bestimmter Hinsicht beschaffen *ist*, nicht, wie es beschaffen sein *soll*. Man antwortet, indem man Standardwerkzeuge aus der Netzwerkforschung auf die gesammelten Daten anwendet oder neue Werkzeuge entwickelt, mit denen man aus den Daten dasjenige lernen kann, was man zu lernen hofft. Manchmal steht das Recht dabei im Vordergrund, manchmal die Methode.<sup>31</sup>

<sup>23</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (324).

<sup>24</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (328).

<sup>25</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (327).

<sup>26</sup> Dazu S. 130–142.

<sup>27</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (330–332).

<sup>28</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (333–336).

<sup>29</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (336–343).

<sup>30</sup> Fowler u.a., *Political Analysis* 15 (2007), S. 324 (325–326, 343–344).

<sup>31</sup> Letzteres trifft insbesondere auf Studien zu, welche die Anwendung neu entwickelter Methoden anhand eines leicht zugänglichen Datensatzes demonstrieren, der das Zitiernetzwerk des *U.S.*

## II. Juristische Akteure

Unter den juristischen Akteuren haben *Anwälte* und die Organisationen, für die sie aktiv sind, in den USA schon früh besondere Aufmerksamkeit erfahren.<sup>32</sup> Vielen Studien in diesem Bereich geht es darum, ein Bild von der Struktur der Anwaltslandschaft oder der rechtspolitischen Landschaft in einem sachlich oder geographisch abgegrenzten Bereich zu zeichnen. Hierzu werden einerseits Interviews mit ausgewählten Anwälten geführt, welche darauf ausgerichtet sind, offenzulegen, mit wem die Anwälte kommunizieren, woher sie ihre Informationen erhalten oder wie sie ihre Zeit einteilen.<sup>33</sup> Andererseits werden Daten zu den Gemeinsamkeiten zwischen Anwälten gesammelt, etwa zum Einsatz für dieselben Organisationen.<sup>34</sup>

---

*Supreme Court* enthält; etwa *Leicht u.a.*, *The European Physical Journal B* 59 (2007), S. 75–83; *Karrer und Newman*, *Physical Review E* 80 (2009), S. 046110-1–046110-12; *Bommarito, Katz und Zelner*, *Procedia Social and Behavioral Sciences* 2010, S. 26–37; *Clough u.a.*, *Journal of Complex Networks* 3 (2015), S. 189–203; *Clough und Evans*, *Physica A* 448 (2016), S. 235–247. Diese Arbeiten werden hier nicht von der Betrachtung ausgenommen, da sich auch aus ihnen zum Teil juristisch interessante Erkenntnisse gewinnen lassen.

<sup>32</sup> *Laumann und Heinz*, *American Bar Foundation Research Journal* 2 (1977), S. 155–216; *Laumann und Heinz*, *American Bar Foundation Research Journal* 4 (1979), S. 217–246; *Laumann u.a.*, *Stanford Law Review* 37 (1985), S. 465–502; *Nelson u.a.*, *Law & Society Review* 22 (1988), S. 237–300; *Lazega*, *Revue Française de Sociologie* 33 (1992), S. 559–589; *Heinz u.a.*, *The Hollow Core*, 1993; *Heinz und Laumann*, *Chicago Lawyers*, 1994; *Lazega*, *Revue Suisse de Sociologie* 21 (1995), S. 61–85; *Lazega und van Duijn*, *Social Networks* 19 (1997), S. 375–397; *Heinz u.a.*, *Law & Society Review* 31 (1997), S. 441–472; *Heinz und Laumann*, *Law & Society Review* 32 (1998), S. 751–775; *Lazega und Pattison*, *Social Networks* 21 (1999), S. 67–90; *Heinz, Southworth und Paik*, *Law & Society Review* 37 (2003), S. 5–50; *Uzzi und Lancaster*, *American Sociological Review* 69 (2004), S. 319–344; *Heinz u.a.*, *Urban Lawyers*, 2005; *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883–917; *Heinz*, *Arizona Law Review* 53 (2011), S. 455–492; *Paik, Heinz und Southworth*, *Law & Social Inquiry* 36 (2011), S. 892–918; *Southworth, Paik und Heinz*, *Lawyers in National Policymaking*, in: *Cummings* (Hrsg.), *The Paradox of Professionalism: Lawyers and the Possibility of Justice*, 2011, S. 220–242.

<sup>33</sup> Schilderungen zu Interviews finden sich etwa bei *Laumann und Heinz*, *American Bar Foundation Research Journal* 2 (1977), S. 155 (164–177); *Laumann und Heinz*, *American Bar Foundation Research Journal* 4 (1979), S. 217 (222–224); *Nelson u.a.*, *Law & Society Review* 22 (1988), S. 237 (241–244); *Lazega*, *Revue Française de Sociologie* 33 (1992), S. 559 (564–565); *Lazega*, *Revue Suisse de Sociologie* 21 (1995), S. 61 (63–66); *Heinz u.a.*, *Law & Society Review* 31 (1997), S. 441 (446–449); *Lazega und Pattison*, *Social Networks* 21 (1999), S. 67 (75–76, 88); *Lazega und van Duijn*, *Social Networks* 19 (1997), S. 375 (381–382); *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (887–890). Oft resultieren aus einem einzigen Interviewprozess mehrere Veröffentlichungen.

<sup>34</sup> *Heinz, Southworth und Paik*, *Law & Society Review* 37 (2003), S. 5–50; *Paik, Heinz und Southworth*, *Law & Social Inquiry* 36 (2011), S. 892 (895–897).

Die hier definierten Netzwerke setzen sich typischerweise aus Anwälten (Objekte) und ihren Kommunikations- oder Anerkennungsbeziehungen (Beziehungen), aus Anwälten (Objekte) und ihren Gemeinsamkeiten (Beziehungen) oder aus Organisationen (Objekte) und ihren – unter anderem über die Kooperation mit denselben Anwälte vermittelten – Gemeinsamkeiten (Beziehungen) zusammen; ihre Analyse ist auch an den Eigenschaften einzelner Objekte interessiert.

Einen etwas anderen Fokus als die Arbeiten zu Anwälten haben jene, die *Mitglieder von Gesetzgebungsorganen* betreffen.<sup>35</sup> Sie versuchen, die übergreifenden Machtstrukturen in den betrachteten Kollektiven offenzulegen. Dazu ergründen sie, welche Mandatsträger in denselben Gremien sitzen (z.B. in den *Committees* des *U.S. House of Representatives*),<sup>36</sup> eine Gesetzesvorlage gemeinsam einbringen (*co-sponsorship*)<sup>37</sup> oder an gemeinsamen Presseveranstaltungen (*press events*) teilnehmen,<sup>38</sup> ohne sich aber mit einzelnen Organmitgliedern im Detail auseinanderzusetzen. Die Netzwerke, die in diesen Studien definiert werden, bestehen entweder aus Gremien (Objekte) und gemeinsamen Mandatsträgern (Beziehungen) oder aus Mandatsträgern (Objekte) und gemeinsamen Aktionen (*co-sponsorship* oder *press events*) (Beziehungen).

Schließlich gibt es noch eine Reihe von Arbeiten zu *anderen Akteuren im Rechtssystem*, insbesondere zu *Richtern*.<sup>39</sup> Diese Arbeiten bemühen sich, Struktur

<sup>35</sup> Porter u.a., PNAS 102 (2005), S. 7057–7062; Fowler, Political Analysis 14 (2006), S. 454–465; Fowler, Social Networks 28 (2006), S. 454–465; Porter u.a., Physica A 386 (2007), S. 414–438; Zhang u.a., Physica A 387 (2008), S. 1705–1712; Desmarais u.a., Social Networks 40 (2015), S. 43–54; Schildknecht, California Law Review 103 (2015), S. 433–466. Alle genannten Studien beschäftigen sich mit US-amerikanischen Gesetzgebungsorganen.

<sup>36</sup> Porter u.a., PNAS 102 (2005), S. 7057–7062; Porter u.a., Physica A 386 (2007), S. 414–438.

<sup>37</sup> Fowler, Political Analysis 14 (2006), S. 454–465; Fowler, Social Networks 28 (2006), S. 454–465; Zhang u.a., Physica A 387 (2008), S. 1705–1712.

<sup>38</sup> Desmarais u.a., Social Networks 40 (2015), S. 43–54.

<sup>39</sup> **Richter und Gerichte in den USA:** Altfeld und Spaeth, Jurimetrics 24 (1984), S. 236–247; Spaeth und Altfeld, The Western Political Quarterly 38 (1985), S. 70–83; Heinz und Manikas, Law & Society Review 26 (1992), S. 831–862; Hook, Constitutional Commentary 24 (2007), S. 221–264; Katz, Stafford und Provins, Georgia State University Law Review 24 (2008), S. 977–1001; Shomade und Hartley, Justice System Journal 31 (2010), S. 144–163; Katz und Stafford, Ohio State Law Journal 71 (2010), S. 457–509; — **Richter und Gerichte in Frankreich:** Lazega und Mounier, Interlocking Judges: On Joint Exogenous and Self-Governance of Markets, in: Buskens, Raub und Snijders (Hrsg.), The Governance of Relations in Markets and Organizations, 2003, S. 267–296; Lazega, Lemerrier und Mounier, European Management Review 3 (2006), S. 113–122; Rouchier, Lazega und Mounier, Articulation of Hierarchy and Networks as an Evolving Social Structure, in: Terano u.a. (Hrsg.), Agent-Based Approaches in Economic and Social Complex Systems IV, 2007, S. 97–104; Lazega u.a., Social Networks 34 (2012), S. 323–332; Rouchier und Tubaro, Procedia Social and Behavioral Sciences 10 (2011), S. 123–131. — **Richter auf europäischer Ebene:** Lazega, Utrecht Law Review 8 (2012), S. 115–128; Lazega, Quitane und Casenaz, Social Networks 48 (2017), S. 10–22. — **Schiedsrichter in ICSID-Verfahren:** Puig, The European Journal of International Law 25 (2014), S. 387–424. — **Amerikanische Law Schools und Professoren:** Katz u.a., Journal of Legal Education 61 (2011), S. 76–103.

und Dynamik sozialer Prozesse im Recht zu erfassen, an denen die betrachteten Akteure beteiligt sind. In Fokus und Herangehensweise ähneln sie zum Teil den Studien zu Anwälten,<sup>40</sup> zum Teil den Studien zu Mitgliedern von Gesetzgebungsorganen.<sup>41</sup>

Zur Illustration der Vorgehensweisen, die in Untersuchungen zu Menschen im Recht verbreitet sind, seien zwei Beispiele betrachtet – eines zu Anwälten, das andere zu Mitgliedern von Gesetzgebungsorganen. Die Studie „Lawyers of the Right: Networks and Organization“ untersucht die Strukturen der Beziehungen zwischen Anwälten, die sich in den USA für konservative und libertäre Interessen einsetzen.<sup>42</sup> Als originäre Datengrundlage dienen 72 semi-strukturierte Interviews, die zwischen 2001 und 2002 mit ausgewählten Anwälten geführt wurden;<sup>43</sup> die Analyse konzentriert sich dann allerdings auf die Beziehungen zwischen 26 besonders prominenten Anwälten.<sup>44</sup> In den Interviews hatten diese Anwälte Angaben zur Häufigkeit ihres Kontakts mit den anderen Anwälten gemacht, auf deren Basis die Studie ein Kontaktnetzwerk mit Anwälten als Objekten und Kontaktangaben als Pfeilen definiert.<sup>45</sup> Die Autoren machen die Struktur dieses Kontaktnetzwerks sichtbar, indem sie den ihm entsprechenden Graphen mithilfe eines Computerprogramms zeichnen, das die Kommunikationsintensität zwischen einzelnen Anwälten in visuelle Nähe zwischen den sie repräsentierenden Punkten übersetzt.<sup>46</sup> Die resultierende Struktur wird mit Attributen der Anwälte (etwa die Organisationen, für die sie tätig sind, ihre Religion und ihr Engagement in der *Federalist Society*) in Bezug gesetzt.<sup>47</sup> Die Studie, die insgesamt vier Tabellen und sechs Graphiken enthält, stellt unter anderem fest, dass die Kontaktnetzwerke der Anwälte sich grundsätzlich nach deren *constituencies* (*religious, libertarian* und *business*) in Blöcke einteilen lassen, dass es aber gleichzeitig eine Reihe von Anwälten gibt, die zwischen diesen Blöcken vermitteln (*core*).<sup>48</sup>

---

<sup>40</sup> Beispiele: *Heinz und Manikas*, *Law & Society Review* 26 (1992), S. 831–862; *Shomade und Hartley*, *Justice System Journal* 31 (2010), S. 144–163; *Lazega*, *Utrecht Law Review* 8 (2012), S. 115–128.

<sup>41</sup> Beispiele: *Spaeth und Altfeld*, *The Western Political Quarterly* 38 (1985), S. 70–83; *Hook*, *Constitutional Commentary* 24 (2007), S. 221–264; *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457–509.

<sup>42</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (883).

<sup>43</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (887). Die Anwälte waren anhand ihrer Beteiligung an für das konservative Spektrum wichtigen *legislative events* und ihrer Verbindungen zu konservativen *key organizations* ausgewählt worden; dazu *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (887–889).

<sup>44</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (889–891).

<sup>45</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (893–894).

<sup>46</sup> Zur verwendeten Zeichenmethode, die auf einer Simulation physikalischer Kräfte aufbaut, *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (893 Fn. 15 und begleitender Text); außerdem unten S. 112–115.

<sup>47</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (893–910).

<sup>48</sup> *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (897–899).

Die Arbeit „A Network Analysis of Committees in the U.S. House of Representatives“ analysiert die Beziehungen zwischen den einzelnen *committees* und *subcommittees* des *U.S. House of Representatives*, die für einen Großteil der Detailarbeit in der US-amerikanischen Gesetzgebung verantwortlich sind, in einem Zeitraum von 1989 bis 2004.<sup>49</sup> Dabei greift sie auf Daten zurück, aus denen sich die Zuordnung von *Representatives* zu den einzelnen Gremien ergibt. Auf Grundlage dieser Daten definieren die Autoren zunächst ein Netzwerk mit *Representatives* und *committees* bzw. *subcommittees* als Objekten und *committee assignments* als Beziehungen.<sup>50</sup> Hieraus leiten sie ein weiteres Netzwerk ab, das nur aus *committees* und *subcommittees* als Objekten besteht. In diesem Netzwerk existiert zwischen zwei Gremien genau dann eine Beziehung, wenn es mindestens einen *Representative* gibt, der in beiden Gremien sitzt, und die Stärke der Beziehung ist proportional zur Anzahl der gemeinsamen Mitglieder beider Gremien.<sup>51</sup> Die Arbeit untersucht, wie das Gremienetzwerk strukturiert ist, indem sie die Punkte des korrespondierenden Graphen mithilfe eines Computerprogramms so in Gruppen einteilt, dass Gremien mit dem höchsten Anteil an gemeinsamen Mitgliedern Schritt für Schritt zu größeren Gruppen zusammengefasst werden.<sup>52</sup> Diese Analyse wird ergänzt durch eine mathematische Untersuchung von Abstimmungsmustern in den *roll-call votes* der *Representatives*, die es den Autoren unter anderem ermöglicht, Aussagen über die Polarisierung einzelner *committees* zu treffen.<sup>53</sup> Die Arbeit enthält drei Tabellen, drei Formeln und vier Graphiken.

An den geschilderten Beispielen zeigt sich, wie verschieden die Herangehensweisen der Studien sein können, die sich mit Menschen im Recht beschäftigen – und das auch dann, wenn sie ihre Forschungsergebnisse gleichermaßen mithilfe von Tabellen und Graphiken kommunizieren. Ebenso verschieden sind die wissenschaftlichen Ausbildungshintergründe der Menschen, die diese Arbeiten verfassen: Die Autoren der ersten Studie sind Juristen und Soziologen, die Autoren der zweiten Studie Mathematiker und Physiker – eine fachliche Vielfalt, die sich auch in vielen weiteren Studien zu Menschen im Recht finden lässt.<sup>54</sup> Gemeinsam ist

<sup>49</sup> Porter u.a., PNAS 102 (2005), S. 7057 (7057).

<sup>50</sup> Porter u.a., PNAS 102 (2005), S. 7057 (7057). Dieses Netzwerk hat zwei verschiedene Mengen von Objekten, *Representatives* und *committees* bzw. *subcommittees*, und Beziehungen bestehen nur zwischen Objekten aus verschiedenen Mengen. Ein solches Netzwerk bzw. den ihm korrespondierenden Graphen nennt man *bipartit* oder auch *bimodal* (*bipartite* bzw. *bimodal*); siehe Wasserman und Faust, *Social Network Analysis*, 1994, S. 120–121; Barabási, *Network Science*, 2016, S. 56 Fig. 2.9.

<sup>51</sup> Porter u.a., PNAS 102 (2005), S. 7057 (7057). Diese Abbildung eines bipartiten Netzwerks auf ein unipartites Netzwerk nennt man *one-mode projection*; dazu Newman, *Networks*, 2010, S. 124–125; Barabási, *Network Science*, 2016, S. 56.

<sup>52</sup> Porter u.a., PNAS 102 (2005), S. 7057 (7059).

<sup>53</sup> Porter u.a., PNAS 102 (2005), S. 7057 (7060–7062).

<sup>54</sup> Das hängt auch damit zusammen, dass ein Großteil der vorhandenen Studien auf einige wenige Wissenschaftler zurückzuführen ist, die über längere Zeit Netzwerkforschung zu Menschen im Recht betrieben haben. Drei Beispiele seien hervorgehoben: John P. Heinz, Jurist, hat in wechselnden Kooperationen mit Soziologen (allen voran Edward O. Laumann) und einem

diesen Studien, dass die untersuchten Netzwerke im Vergleich zu den Netzwerken aus Texten relativ klein sind:<sup>55</sup> mehrere Dutzend bis mehrere Hundert, oft eng miteinander verbundene Objekte hier; mehrere Hundert bis mehrere Tausend, oft lose miteinander verbundene Objekte dort. Auch ist das dominierende Erkenntnisinteresse hier nicht auf Texte als (mehr oder weniger) direkte Quellen von Recht, sondern auf die sozialen und politischen Kontexte gerichtet, in denen es geformt wird. Mit den Untersuchungen zu Texten im Recht gemein haben die Arbeiten zu Menschen im Recht allerdings, dass hier rechtssysteminterne Strukturen und Prozesse betrachtet werden. Anders verhält es sich mit Studien, die Recht und Realität mithilfe der Netzwerkforschung in Verbindung setzen.

### III. Recht und Realität

Untersuchungen, welche die Wechselwirkungen zwischen Recht und Realität in den Blick nehmen, beschäftigen sich unter anderem mit *Beziehungen zwischen Unternehmen*. Seit den 1970er Jahren erfahren dabei die *Unternehmensverflechtungen* in Form von *Personalverflechtungen (interlocking directorates)* besondere Aufmerksamkeit.<sup>56</sup> Eine Personalverflechtung zwischen zwei Unternehmen liegt immer dann vor, wenn mindestens eine Person in Führungsgremien beider Unternehmen sitzt.<sup>57</sup> Unter Juristen stoßen Personalverflechtungen traditionell

---

Politikwissenschaftler (*Robert H. Salisbury*) eine Reihe von Studien zu Anwälten verfasst. Zahlreiche Veröffentlichungen zu Mitgliedern von Gesetzgebungsorganen gehen auf Forscher um *Mason A. Porter* zurück, der Mathematiker ist. *Emmanuel Lazega*, beteiligt an einer Vielzahl von Studien zu Anwälten oder Richtern, die sich auf Dynamiken in Organisationen konzentrieren, ist Soziologe.

<sup>55</sup> Zu Netzwerken aus Texten bereits S. 13–17.

<sup>56</sup> Überblicke über die empirische Forschung zu Personalverflechtungen seit den 1970er Jahren finden sich bei *Mizruchi*, *Annual Review of Sociology* 22 (1996), S. 271–298, sowie in jüngerer Zeit bei *Carroll und Sapinski*, *Corporate Elites and Intercorporate Networks*, in: *Scott und Carrington* (Hrsg.), *The SAGE Handbook of Social Network Analysis*, 2011, S. 180–195. Der Beginn der Forschung zu Personalverflechtungen liegt früher und wird oft auf eine Arbeit von *Otto Jeidels* zurückgeführt: *Jeidels*, *Das Verhältnis der deutschen Großbanken zur Industrie mit besonderer Berücksichtigung der Eisenindustrie*, 1905; dazu *Fennema und Schijf*, *Social Networks* 1 (1978), S. 297 (298–300). In den USA wurden *interlocking directorates* für bestimmte Konstellationen erstmals 1914 gesetzlich untersagt, nachdem das *Pujo Committee* diese 1913 als wettbewerbseinschränkend identifiziert und ihr Verbot empfohlen hatte: *Clayton Act*, Pub.L. 63–212 (1914), Sec. 8; *Pujo Committee: Report of the Committee Appointed Pursuant to House Resolutions 428 and 504 to Investigate the Concentration of Control of Money and Credit*. House Report No. 1593. 62nd Congress, 3rd Session, 28. Feb. 1913, S. 139–140. Auch im Schrifttum waren zuvor bereits Rufe nach einem Verbot laut geworden, siehe nur *Pam*, *Harvard Law Review* 26 (1913), S. 467 (481).

<sup>57</sup> *Mizruchi*, *Annual Review of Sociology* 22 (1996), S. 271 (271). Welche Führungsgremien in die Analyse einbezogen werden können, hängt insbesondere davon ab, ob den betrachteten Unternehmen (typischerweise Kapitalgesellschaften) ein *One-Tier Board System* (nur Vorstand bzw. *Board of Directors*) oder ein *Two-Tier Board System* (Vorstand und Aufsichtsrat bzw. *Board of Directors* und *Supervisory Board*) zugrunde liegt.

auf Skepsis, insbesondere aus gesellschaftsrechtlichen, kartellrechtlichen und ethischen Gründen.<sup>58</sup> Arbeiten, die sich dem Phänomen mit Instrumenten der Netzwerkforschung nähern, definieren Netzwerke aus Unternehmen (Objekte) und ihren Personalverflechtungen (Beziehungen).<sup>59</sup> Diese oft soziologisch oder organisationswissenschaftlich motivierten Studien interessieren sich zum Einen für die Struktur der Verflechtungsnetzwerke und deren Veränderung über die Zeit,<sup>60</sup> zum Anderen für den Einfluss von Personalverflechtungen auf die Ertragskraft und die Führungspraktiken (*Corporate Governance*) der betroffenen Unternehmen.<sup>61</sup> Einige Arbeiten versuchen auch, die Erkenntnisse der Netzwerkforschung

---

<sup>58</sup> So heißt es etwa bei *Louis D. Brandeis*, der in diesem Kontext gern zitiert wird (z.B. von *Dooley*, *American Economic Review* 59 (1969), S. 314 (314), und von *Barzuza und Curtis*, *Delaware Journal of Corporate Law* 39 (2015), S. 669 (678–679): „The practice of interlocking directorates is the root of many evils. It offends laws human and divine. Applied to rival corporations, it tends to the suppression of competition. [...] Applied to corporations which deal with each other, it tends to disloyalty and to violation of the fundamental law that no man can serve two masters.“; *Brandeis*, *Other People’s Money*, 1914, S. 51. *Brandeis* legt dort allerdings eine Definition von *interlocking directorates* zugrunde, die sämtliche Unternehmensverflechtungen umfasst und damit weiter ist als die hier verwendete, heute übliche Definition; siehe *Brandeis*, *Other People’s Money*, 1914, S. 51–52.

<sup>59</sup> Die Ähnlichkeit mit der oben als Beispiel geschilderten Studie zu Mitgliedern von Gesetzgebungsorganen, „A Network Analysis of Committees in the U.S. House of Representatives“, ist kein Zufall. Vielmehr haben sich die Autoren dieser Studie von der Literatur zu *interlocking directorates* inspirieren lassen; siehe *Porter u.a.*, *PNAS* 102 (2005), S. 7057 (7057).

<sup>60</sup> Beispiele: *Sonquist und Koenig*, *Insurgent Sociologist* 5 (1975), S. 196–229 (Personalverflechtungen zwischen Unternehmen der US-amerikanischen *Fortune 500* im Jahr 1969); *Davis und Mizruchi*, *Administrative Science Quarterly* 44 (1999), S. 215–239 (Personalverflechtungen zwischen den 50 größten *bank holding companies* und einer Auswahl an anderen großen Unternehmen in den USA für die Jahre 1982, 1986, 1990 und 1994); *Fohlin*, *The Economic History Review* 52 (1999), S. 307–333 (Personalverflechtungen zwischen börsennotierten Unternehmen in Deutschland zwischen 1880 und 1913); *Heinze*, *Zeitschrift für Soziologie* 31 (2002), S. 391–410 (Personalverflechtungen von 69 großen Unternehmen in Deutschland in den Jahren 1989 und 2001); *Windolf*, *European Sociological Review* 25 (2009), S. 443–457 (Personalverflechtungen in Deutschland und den USA in ausgewählten Jahren zwischen 1896 und 1938); *van Veen und Kratzer*, *Economy & Society* 40 (2011), S. 1–25 (nationale und internationale Personalverflechtungen zwischen Unternehmen aus fünfzehn EU-Staaten am Anfang des Jahres 2006); *Takes und Heemskerck*, *Social Network Analysis and Mining* 6 (2016), S. 97–1–97–18 (Personalverflechtungen zwischen 971.891 Unternehmen aus verschiedenen Ländern).

<sup>61</sup> Beispiele: *Davis*, *Administrative Science Quarterly* 36 (1991), S. 583–613 (*Poison Pills*); *Haunschild*, *Administrative Science Quarterly* 38 (1993), S. 564–592 (*Corporate Acquisition Activity*); *Stearns und Mizruchi*, *The Academy of Management Journal* 36 (1993), S. 603–618 (*Corporate Financing*); *Gulati und Westphal*, *Administrative Science Quarterly* 44 (1999), S. 473–506 (*Joint Ventures*); *Bizjak, Lemmon und Whitby*, *The Review of Financial Studies* 22 (2009), S. 4821–4847 (*Option Backdating*); *Brown*, *The Accounting Review* 86 (2011), S. 23–57 (*Corporate-Owned Life Insurance Shelter*); *Buch-Hansen*, *International Sociology* 29 (2014), S. 249–267 (*Collusion*).

unmittelbar in die juristische Diskussion zu Notwendigkeit der Regulierung von Personalverflechtungen einfließen zu lassen.<sup>62</sup>

Neben den Studien zu Personalverflechtungen hat sich – unter anderem als Reaktion auf Krisenerfahrungen zu Beginn des 21. Jahrhunderts – ein weiterer, ökonomisch und finanzwissenschaftlich ausgerichteter Forschungszweig etabliert, der *finanzielle Beziehungen* zwischen Unternehmen im Finanzsektor speziell mit Blick auf die *Stabilität von Finanzmärkten* analysiert (*financial networks*).<sup>63</sup> Dazu werden Netzwerke mit Banken oder anderen finanziellen Institutionen als Objekte definiert, die in der Regel durch Kreditbeziehungen, zum Teil aber auch durch mehrere verschiedene Arten von finanziellen Beziehungen (etwa Unternehmensbeteiligungen, Kredite und Kreditsicherheiten) miteinander verbunden sind.<sup>64</sup> Die Studien fragen, wie sich finanzielle Schocks durch unterschiedlich strukturierte Netzwerke von Finanzbeziehungen ausbreiten und unter welchen Umständen der Ausfall einzelner Akteure durch Ansteckungseffekte (*contagion*) zum Ausfall weiterer Akteure führt (*cascading failures*).<sup>65</sup> Sie bemühen sich, systemisches

<sup>62</sup> Barzuza und Curtis, Delaware Journal of Corporate Law 39 (2015), S. 669–701 (Rezeption vorhandener Literatur); Fischer, The Journal of Corporation Law 41 (2015), S. 313–341 (mit eigener Studie); Petersen, European Business Law Review 27 (2016), S. 821–864 (Rezeption vorhandener Literatur); Barzuza und Curtis, The Journal of Legal Studies 46 (2017), S. 129–160 (mit eigener Studie); Ansätze auch bei Gerber, Yale Journal on Regulation 24 (2007), S. 107 (116–117), und bei Falce, Journal of Competition Law & Economics 9 (2013), S. 457 (458–459).

<sup>63</sup> Beispiele mit expliziter Bezugnahme auf vorangegangene Krisen: Gai und Kapadia, Proceedings of the Royal Society A 2010, S. 1–23; Haldane und May, Nature 469 (2011), S. 351–355; Battiston u.a., Journal of Economic Dynamics & Control 36 (2012), S. 1121–1141; Markose, Journal of Banking Regulation 14 (2013), S. 285–305; Elliott, Golub und Jackson, American Economic Review 104 (2014), S. 3115–3153; Acemoglu, Ozdaglar und Tahbaz-Salehi, American Economic Review 105 (2015), S. 564–608; Heath u.a., Journal of Financial Stability 27 (2016), S. 217–233.

<sup>64</sup> Beispiele: Peltonen, Scheicher und Vuilleme, Journal of Financial Stability 13 (2014), S. 118–133 (verschiedene Finanzinstitutionen und Credit Default Swap (CDS) Exposures); Acemoglu, Ozdaglar und Tahbaz-Salehi, American Economic Review 105 (2015), S. 564–608 (Banken und Kredite); Heath u.a., Journal of Financial Stability 27 (2016), S. 217–233 (Banken, Central Counterparties und Clearing-Beziehungen); Battiston, Caldarelli und D’Errico, The Financial System as a Nexus of Interconnected Networks, in: Garas (Hrsg.), Interconnected Networks, 2016, S. 195–229 (verschiedene Akteure im Finanzsystem und verschiedene Arten von Finanzbeziehungen). Manchmal werden die Daten auch auf die Ebene von Staaten aggregiert, so etwa bei Elliott, Golub und Jackson, American Economic Review 104 (2014), S. 3115 (3144–3147) (Kreditbeziehungen zwischen Banken in sechs EU-Ländern).

<sup>65</sup> Literaturüberblicke finden sich bei Upper, Journal of Financial Stability 7 (2011), S. 111–125; Cabrales, Gale und Gottardi, Financial Contagion in Networks, in: Bramoullé, Galeotti und Rogers (Hrsg.), The Oxford Handbook of the Economics of Networks, 2016, S. 543–568; und Acemoglu, Ozdaglar und Tahbaz-Salehi, Networks, Shocks, and Systemic Risk, in: Bramoullé, Galeotti und Rogers (Hrsg.), The Oxford Handbook of the Economics of Networks, 2016, S. 569–607. In manchen Arbeiten werden juristische Konzepte explizit ökonomisch modelliert. So heißt es etwa in einer Studie zur Wirkung von Central Counterparties (CCPs): „By novation, a third party steps into a bilateral contract, guaranteeing performance of both legs of the trade.

Risiko (*systemic risk*) in Finanzmärkten messbar zu machen, die Voraussetzungen und Konsequenzen seiner Realisierung zu verstehen und mit ihren Erkenntnissen auf die Finanzmarktregulierung Einfluss zu nehmen.<sup>66</sup>

Ein distanzierteres Verhältnis zu rechtlicher Regulierung bewahren Studien, die sich aus der Netzwerkperspektive mit *Kriminalität* beschäftigen (*criminal networks*). Strukturell den Arbeiten zu juristischen Akteuren ähnlich, konstruieren diese Untersuchungen ihre Netzwerke häufig aus tatsächlichen oder potenziellen Straftätern (Objekte), die durch soziale Beziehungen miteinander verbunden sind (Beziehungen). Dabei lassen sich zwei Schwerpunkte des Erkenntnisinteresses ausmachen: Zum Einen bemüht man sich, die sozialen Kontexte und Dynamiken von Kriminalität zu verstehen und herauszufinden, wie die kriminellen Karrieren Einzelner durch ihr soziales Umfeld beeinflusst werden.<sup>67</sup> Zum Anderen sucht man mit Werkzeugen aus der Netzwerkforschung nach Möglichkeiten, Kriminalität durch gezielte Eingriffe in kriminelle Netzwerke effektiv zu bekämpfen – sowohl repressiv als auch präventiv.<sup>68</sup> Beide Problemstellungen werden insbeson-

[...] Novation is modelled here as a function which transforms (trading) exposures into (cleared) exposures.“; *Galbati und Soramäki*, *Journal of Economic Behavior & Organization* 83 (2012), S. 609 (609).

<sup>66</sup> Die Einflussbemühungen sind von unterschiedlicher Intensität. Vergangenes betonend etwa *Hautsch, Schaumburg und Schienle*, *Review of Finance* 19 (2015), S. 685–738 (731): „[...] [W]e highlight that our approach could have served as a solid basis for a sensible forward-looking monitoring tool before the start of the financial crisis in 2007.“ Zukunftsgerichtet hingegen *Gai, Haldane und Kapadia*, *Journal of Monetary Economics* 58 (2011), S. 453 (469): „The paper has also identified some public policy measures which could mitigate these fragilities, many of which are currently being designed or implemented internationally. These include tougher micro-prudential liquidity regulation, countercyclical liquidity requirements, and surcharges for systemically important financial institutions.“ Besonders präzise *Markose, Giansante und Shaghghi*, *Journal of Economic Behavior & Organization* 83 (2012), S. 627 (644): „We recommend that the Basel II provision for capital reduction on bank assets that have CDS cover should be discontinued.“ Ebenso *Battiston und Caldarelli*, *Financial Networks*, in: *D’Agostino und Scala* (Hrsg.), *Networks of Networks: The Last Frontier of Complexity*, 2014, S. 311 (319): „This result supports the proposal of the Basel III Committee to introduce a 30-day liquidity coverage ratio, and suggests the monthly scale as a reasonable time window for observing the system.“ *Haldane und May* widmen den „Implications for public policy“ gleich drei Spalten: *Haldane und May*, *Nature* 469 (2011), S. 351 (354–355) (*Haldane* ist zum Publikationszeitpunkt mit der *Bank of England* assoziiert). Eine regulatorische Agenda bereits im Titel trägt der Beitrag von *Poledna und Thurner*, *Quantitative Finance* 16 (2016), S. 1599–1613: „Elimination of Systemic Risk in Financial Networks by Means of a Systemic Risk Transaction Tax“.

<sup>67</sup> Überblicke zu dieser Forschungsrichtung bieten *Carrington*, *Crime and Social Network Analysis*, in: *Scott und Carrington* (Hrsg.), *The SAGE Handbook of Social Network Analysis*, 2011, S. 236–255, und *Papachristos*, *Sociology Compass* 8 (2014), S. 347–357; Beispiele für frühe theoretische Erwägungen finden sich etwa bei *Krohn*, *Social Problems* 33 (1986), S. S81–S93, und *McIlhwan*, *Crime, Law & Social Change* 32 (1999), S. 301–323; Fallstudien unter anderem bei *Morselli*, *Inside Criminal Networks*, 2009.

<sup>68</sup> Frühe Beispiele dieser Untersuchungsrichtung finden sich bei *Davis*, *FBI Law Enforcement Bulletin* 50 (1981), S. 11–19, und *Sparrow*, *Social Networks* 13 (1991), S. 251–274. Aus der

dere für die organisierte Kriminalität untersucht.<sup>69</sup> Die Daten, auf denen Studien zu kriminellen Netzwerken aufbauen, werden mit klassischen soziologischen Methoden (z.B. semistrukturierten Interviews) erhoben<sup>70</sup> oder aus behördlichen oder gerichtlichen Dokumenten (z.B. Vernehmungsprotokollen oder Verfahrensakten) extrahiert.<sup>71</sup>

Neben den Beziehungen zwischen Menschen oder Unternehmen, die in rechtlich interessanten Kontexten interagieren, nehmen Studien zu den Wechselwirkungen von Recht und Realität auch *Patente* in den Blick. Die Arbeiten in diesem Bereich ähneln in ihrer Herangehensweise oft den Studien zu Gerichtsentscheidungen: Sie konstruieren Netzwerke, in denen einzelne Patente oder Patentfamilien<sup>72</sup> die Objekte darstellen, deren Beziehungen zueinander durch Zitate im Rahmen der Angaben zum Stand der Technik (*prior art*) hergestellt werden (*patent citation networks*)<sup>73, 74</sup>. Als Grundlage der Netzwerkdefinition dienen oft Daten aus

---

jüngerer Literatur *Carley*, Computational & Mathematical Organization Theory 12 (2006), S. 51–66; *Schwartz und Rouselle*, Trends in Organized Crime 12 (2009), S. 188–207; *Everton*, Disrupting Dark Networks, 2012; *Strang*, Network Analysis in Criminal Intelligence, in: *Masys* (Hrsg.), Networks and Network Analysis for Defence and Security, 2014, S. 1–26; *Duijn, Kashirin und Sloat*, Scientific Reports 4 (2014), S. 4238-1–4238-15; *Tayebi und Glässer*, Social Network Analysis in Predictive Policing, 2016; *Berlusconi*, Social Network Analysis and Crime Prevention, in: *Leclerc und Savona* (Hrsg.), Crime Prevention in the 21st Century, 2017, S. 129–141.

<sup>69</sup> Beispiele: *Krebs*, Connections 24 (2002), S. 43–52 (Terrorismus); *Bright, Hughes und Chalmers*, Crime, Law & Social Change 57 (2012), S. 151–176; *Calderoni*, Crime, Law & Social Change 58 (2012), S. 321–349 (jeweils Drogenschmuggel); *Papachristos*, American Journal of Sociology 115 (2009), S. 74–128; *Tita und Radil*, Journal of Quantitative Criminology 27 (2011), S. 521–545 (jeweils gang violence).

<sup>70</sup> Beispiele: *Papachristos, Meares und Fagan*, The Journal of Criminal Law & Criminology 102 (2012), S. 397–440; *Haynie, Doogan und Soller*, Criminology 52 (2014), S. 688–722; *Kreager u.a.*, Justice Quarterly 33 (2016), S. 1000–1028.

<sup>71</sup> Beispiele: *Baker und Faulkner*, American Sociological Review 58 (1993), S. 837–860; *Calderoni*, Crime, Law & Social Change 58 (2012), S. 321–349; *Papachristos, Hureau und Braga*, American Sociological Review 78 (2013), S. 417–447.

<sup>72</sup> Eine Patentfamilie ist eine Gruppe von Patentveröffentlichungen, die von denselben Erfindern in verschiedenen Ländern zum Schutz derselben Erfindung eingereicht wurden; *Simmons*, World Patent Information 31 (2009), S. 11 (11). Speziell zum Einfluss von Informationen zu Patentfamilien auf die Analyse von Patentnetzwerken *Nakamura u.a.*, Scientometrics 104 (2015), S. 437–452.

<sup>73</sup> Die Parallelen zwischen *judicial citations* und *patent citations* führten in den USA schon in den späten 1940er Jahren zu Vorschlägen, eine von *Shepard's Citations* inspirierte Patentindizierung einzuführen; siehe *Seidel*, Journal of the Patent Office Society 31 (1949), S. 554; *Hart*, Journal of the Patent Office Society 31 (1949), S. 714; *Garfield*, Journal of the Patent Office Society 39 (1957), S. 583 (584). *Whalen*, Michigan State Law Review 2016, S. 539 (548–551), bespricht Patentnetzwerke gemeinsam mit gerichtlichen Zitiernetzwerken unter dem Oberbegriff *Legal Citation Networks*.

<sup>74</sup> Beispiele für verschiedene alternative Netzwerkdefinitionen finden sich unter anderem bei *Balconi, Breschi und Lissoni*, Research Policy 33 (2004), S. 127–145 (Erfinder); *Sternitzke, Bartkowski und Schramm*, World Patent Information 30 (2008), S. 115–131 (Erfinder und

den USA, die sich auf Patente beschränken, welche beim *United States Patent and Trademark Office (USPTO)* registriert sind.<sup>75</sup> Viele der Studien versuchen, Innovationsprozesse nachzuvollziehen,<sup>76</sup> Technologietrends zu erkennen,<sup>77</sup> Patentierungsverhalten zu beschreiben<sup>78</sup> oder Maßstäbe für die Patentbewertung zu entwickeln.<sup>79</sup> Damit nutzen sie die Strukturen, die sich aus dem Recht (in Form von Patenten) ergeben, vor allem als Proxy für die Entwicklung der dahinterstehenden Technologien.<sup>80</sup>

Der kurze Streifzug durch die Literatur zur Netzwerkforschung im Recht verdeutlicht vor allem ihre Vielfalt. Wissenschaftler aus den verschiedensten Disziplinen nutzen Instrumente der Netzwerkforschung, um Fragen zu untersuchen, die das Recht betreffen. Die Gegenstände, die dabei in den Blick genommen werden,

---

Patentanmelder); *Kim und Song*, *Technological Forecasting and Social Change* 80 (2013), S. 944–955 (Unternehmen mit Patentverletzungsklagen gegeneinander); *Choi und Hwang*, *Technological Forecasting and Social Change* 83 (2014), S. 170–182 (*patent keywords*).

<sup>75</sup> Der derzeit noch immer dominierende Datensatz ist *Hall, Jaffe und Trajtenberg*, *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools*, NBER Working Paper No. 8498, 2001; dazu auch *Jaffe und de Rassenfosse*, *Journal of the Association for Information Science and Technology* 68 (2017), S. 1360 (1374), im Rahmen eines (nicht auf netzwerkanalytische Ansätze beschränkten) Überblicks zur sozialwissenschaftlichen Nutzung von *patent citation data*. In den USA werden die Zitate von *prior art*, auf die sich die Patentevaluation stützt, seit dem 4. Februar 1947 auf dem Patent selbst aufgeführt; bis dahin waren sie lediglich aus dem *file wrapper* ersichtlich gewesen; dazu *Garfield*, *Journal of the Patent Office Society* 39 (1957), S. 583 (586); *Alcácer, Gittelman und Sampat*, *Research Policy* 38 (2009), S. 415 (416 Fn. 1). Beide Quellen nehmen auf eine Anordnung des *USPTO* vom 19. Dezember 1946 Bezug, deren (jeweils nicht angegebene) Fundstelle die Verfasserin nicht aufspüren konnte. Dass es diese Anordnung gegeben hat, ergibt sich allerdings aus *Wolcott*, *Journal of the Patent Office Society* 29 (1947), S. 214.

<sup>76</sup> Beispiele: *Fontana, Nuvolari und Verspagen*, *Economics of Innovation and New Technology* 18 (2009), S. 311–336; *Kim, Cho und Kim*, *Scientometrics* 98 (2014), S. 975–998; *Martinelli und Nomaler*, *Journal of Evolutionary Economics* 24 (2014), S. 623–652; *Whalen*, *Michigan State Law Review* 2016, S. 539–565.

<sup>77</sup> Beispiele: *Érdi u.a.*, *Scientometrics* 95 (2013), S. 225–242; *Bruck u.a.*, *Scientometrics* 107 (2016), S. 1465–1475.

<sup>78</sup> Beispiele: *Strandburg u.a.*, *Berkeley Technology Law Journal* 21 (2006), S. 1293–1362; *Strandburg u.a.*, *North Carolina Law Review* 87 (2009), S. 1657–1698.

<sup>79</sup> Beispiele: *Yang u.a.*, *Scientometrics* 105 (2015), S. 1319–1346; *Torrance und West*, *Virginia Journal of Law & Technology* 20 (2017), S. 466–504.

<sup>80</sup> Neben dem Schwerpunkt des Erkenntnisinteresses der Netzwerkstudien, nach dem sich die hiesige Einordnung richtet, bestehen zahlreiche Unterschiede zwischen den Kontexten, in denen Patentnetzwerke und gerichtliche Zitiernetzwerke zustande kommen. Die Zitierung von Patenten als *prior art* erfolgt durch Patentprüfer oder Patentanmelder nach bestimmten Kriterien in einem geordneten Verfahren, dessen Einzelheiten vom jeweiligen Patentamt abhängen; die Zitierung anderer Gerichtsentscheidungen durch Richter unterliegt weit weniger strengen Regeln. Während die wahrgenommene Legitimation gerichtlicher Entscheidungen durch Zitate tendenziell wächst, sinkt die Patentierbarkeit einer Erfindung tendenziell mit jedem Zitat von *prior art*. Diese Unterschiede könnten auch die festgestellten Differenzen in der Struktur beider Zitiernetzwerkformen erklären; dazu *Clough u.a.*, *Journal of Complex Networks* 3 (2015), S. 189–203.

und die Perspektiven, aus denen diese betrachtet werden, sind so unterschiedlich, dass jeder Jurist sich angesprochen fühlen kann – unabhängig davon, welcher Fachsäule oder Teildisziplin innerhalb der Rechtswissenschaft er sich zuordnet. Gleichzeitig fällt auf, dass die dargestellten Forschungsströmungen bisher wenig miteinander interagieren und dass das Recht selbst selten im Zentrum der Untersuchungen steht. Dieser Befund motiviert die *juristische Netzwerkforschung*. Die juristische Netzwerkforschung ist Netzwerkforschung, die das Recht in den Fokus rückt. Sie interessiert sich für die bestehende Netzwerkforschung im Recht genau so weit, wie sie daraus rechtswissenschaftliche Erkenntnis generieren kann. Sie stellt neue Bezüge zwischen existierenden Forschungsfeldern her, soweit es den rechtswissenschaftlichen Erkenntnisgewinn fördert. Und sie entwickelt die Instrumente der Netzwerkforschung genau dort weiter, wo dies die Rechtswissenschaft voranbringt.

## B. Gegenstand

Die Entfaltung der juristischen Netzwerkforschung beginnt mit einem Definitionsproblem: Was sind die Objekte und Beziehungen, deren gemeinsame Betrachtung die Rechtswissenschaft voranbringen könnte? Die Beispiele im Literaturüberblick decken nur einen Bruchteil der Konstellationen ab, die von rechtswissenschaftlichem Interesse sein können, und ihre zusammenfassende Umschreibung verdeckt einen Großteil der Probleme, die sich bei der Netzwerkdefinition stellen. Sowohl die Ergebnisse aller Analysen als auch ihre Interpretation und Plausibilität hängen stark von der gewählten Netzwerkdefinition ab. Im Folgenden wird daher eine Taxonomie der Objekte (I.) und Beziehungen (II.) entwickelt, deren Untersuchung sich rechtswissenschaftlich lohnen kann. Dabei werden insbesondere die Schwierigkeiten aufgezeigt, die bei der Definition von Objekt- und Beziehungsmengen auftreten können. Anschließend lassen sich einige typische Probleme hervorheben, die speziell aus dem Zusammenspiel von Objekt- und Beziehungsdefinition resultieren (III.).

### I. Objekte

Die Vielzahl der Objekte, welche die juristische Netzwerkforschung in den Fokus rücken kann, lässt sich am besten überblicken, wenn man sie zunächst anhand juristisch sinnvoller Kategorien in drei Mengen einteilt:<sup>81</sup> Rechtssubjekte (1.), Rechtstexte (2.) und Rechtsbegriffe (3.).

---

<sup>81</sup> Zu den Schwierigkeiten der Objektdefinition allgemein *Zweig*, *Network Analysis Literacy*, 2016, S. 121–127, 289–293.

### 1. Rechtssubjekte

Wissend, dass sich zum Begriff des Rechtssubjekts ganze Bücher schreiben lassen,<sup>82</sup> wird hier folgende Definition zugrunde gelegt: *Rechtssubjekt ist, wer fähig ist, Inhaber von Rechten und Pflichten zu sein.* Die Einheiten, die durch diese Definition erfasst werden, können wiederum in Individuen (a) und Kollektive (b) eingeteilt werden.<sup>83</sup>

#### a) Individuen

Die Möglichkeit, einzelne Individuen als Objekte in einem Netzwerk aufzufassen, entspricht der Alltagserfahrung des Menschen (z.B. als Akteur in *sozialen Netzwerken*). Es drängt sich aber sofort die Frage auf: Welche Individuen *genau* sollen betrachtet werden? Der Literaturüberblick hat gezeigt, dass die einzelnen Arbeiten zur Netzwerkforschung im Recht diese Frage zum Teil sehr unterschiedlich beantworten.<sup>84</sup> Man betrachtet Anwälte in einer bestimmten Region, Mitglieder eines bestimmten Gesetzgebungsorgans, Richter eines bestimmten Gerichts. Um die Menge der untersuchten Objekte eindeutig festzulegen, sind aber weit mehr Kriterien erforderlich, als der vorstehenden Beschreibung zugrunde liegen. Diese Kriterien lassen sich als *Variablen* begreifen, deren *Ausprägungen* (Werte im konkreten Fall) die Objektdefinition anleiten können. Beispielfhaft seien genannt:

- *Zeit.* Meist interessiert man sich für Individuen, die zu einem bestimmten Zeitpunkt oder in einem bestimmten Zeitraum lebendig waren bzw. sind und die darüber hinaus noch mindestens ein weiteres Eingrenzungskriterium erfüllen. Das mag trivial erscheinen, wird in vielen Netzwerkstudien allerdings nur implizit kommuniziert – etwa, weil Interviews in einem bestimmten Zeitraum durchgeführt und die Interviewpartner auf Basis von Informationen ausgewählt werden, die wiederum nur einen bestimmten Zeitraum abdecken.<sup>85</sup> In der älteren Netzwerkforschung hat die Zeit kaum eine Rolle gespielt; in der jüngeren Literatur erfährt sie aber zunehmende Aufmerksamkeit.<sup>86</sup>
- *Ort.* Individuen lassen sich auch danach auswählen, dass sie (zu einer bestimmten Zeit) an einem bestimmten Ort leben oder ein anderes besonderes Verhältnis zu diesem Ort aufweisen. Für die juristische Netzwerkforschung ist das Staats-

<sup>82</sup> Klassiker: Ehrlich, Die Rechtsfähigkeit, 1909.

<sup>83</sup> Manche Autoren unterscheiden zwischen *basic units* und *aggregate units*, so etwa Börner, Atlas of Science, 2010, S. 54–55. Diese Differenzierung wird hier bewusst nicht aufgegriffen, da ein Kollektiv nicht notwendig (nur) die Summe seiner Teile ist, wie es die Aggregationsterminologie suggeriert.

<sup>84</sup> Dazu bereits S. 18–22.

<sup>85</sup> Siehe etwa die Studie von Paik, Southworth und Heinz, Law & Social Inquiry 32 (2007), S. 883–917 (ausführlich als Beispiel behandelt oben auf S. 20–20).

<sup>86</sup> Dabei interessiert man sich insbesondere für die Veränderung von Netzwerken über die Zeit; siehe etwa Holme und Saramäki, Physics Reports 519 (2012), S. 97–125, mit zahlreichen weiteren Nachweisen.

gebiet ein natürlicher Bezugspunkt, aber auch kleinere Anknüpfungspunkte (z.B. Bundesland, Gemeinde) oder größere Anknüpfungspunkte (z.B. Europäische Union) sind denkbar. Die Werte der Ortsvariablen werden allerdings häufig von den Werten anderer Variablen abhängig sein – etwa, weil Menschen, die einen ortsgebundenen Beruf ausüben, häufig im Land der Berufsausübung leben.

- *Beruf*. Angesichts der Vielzahl spezialisierter juristischer Tätigkeiten scheint der Beruf eines Menschen (zu einer bestimmten Zeit, an einem bestimmten Ort) für die juristische Netzwerkforschung als Eingrenzungskriterium besonders naheliegend. So lassen sich Netzwerkstrukturen in abgegrenzten juristischen Berufsgruppen untersuchen, wie es bisher insbesondere für Anwälte und Richter in den USA und in Frankreich unternommen worden ist.<sup>87</sup> Die Zuordnung eines Individuums zu einem Beruf kann mitunter schwierig sein – man denke etwa an Rechtsreferendare in der Anwaltsstation oder Privatdozenten an juristischen Fakultäten. Die praktischen Auswirkungen dieses Problems lassen sich aber eindämmen, indem man für die Zuordnungsentscheidung wiederum Kriterien bestimmt, an denen die Entscheidung einheitlich ausgerichtet werden kann.
- *Affiliation*. Als weiteres Eingrenzungskriterium kommt (vor allem bei berufsgruppenbezogenen Studien) eine bestimmte institutionelle Affiliation in Betracht. Beispiele sind die Zugehörigkeit zu einem bestimmten Gesetzgebungsgremium (z.B. dem Bundestag) oder die Mitgliedschaft in einer bestimmten Vereinigung (z.B. der Deutschen Staatsrechtslehrervereinigung). In Deutschland könnten sich so etwa Bundesrichter (affiliert mit Bundesgerichten), BGH-Anwälte (affiliert mit der Rechtsanwaltskammer beim Bundesgerichtshof) oder Lehrstuhlinhaber an juristischen Fakultäten (affiliert mit deren Hochschulen) als interessante Gegenstände juristischer Netzwerkforschung erweisen. Soll das Zusammenspiel mehrerer Akteure mit verschiedenen Affiliationen untersucht werden, so kann es sinnvoll sein, die einzelnen Affiliationspunkte neben den Individuen als Objekte in das definierte Netzwerk aufzunehmen – man erhält dann ein bipartites Netzwerk, bei dessen Untersuchung einige Besonderheiten zu beachten sind.<sup>88</sup> Unabhängig davon können bei der Bestimmung von Affiliationen ähnliche Schwierigkeiten auftreten wie bei der Berufsbestimmung – insbesondere dann, wenn die Affiliation nicht durch einen Rechtsakt initiiert wird. Die Anknüpfung an einen affiliierenden Rechtsakt (z.B. Beitritt, Ernennung) kann in bestimmten Untersuchungsszenarien allerdings wiederum unzweckmäßig sein – etwa dann, wenn es darum geht, informelle Einflussprozesse sichtbar zu machen.

---

<sup>87</sup> Dazu bereits S. 18–22.

<sup>88</sup> Ein Netzwerk ist bipartit, wenn sich seine Objekte so in zwei Mengen einteilen lassen, dass alle Beziehungen nur zwischen Objekten *verschiedener* Mengen verlaufen. Dazu bereits in Fn. 50 (S. 21).

- *Partizipation*. Schließlich kann auch die Teilnahme an einzelnen Ereignissen, etwa dem Deutschen Anwaltstag oder einer wissenschaftlichen Konferenz, ein Kriterium zur Definition einer aus Menschen zusammengesetzten Objektmenge sein. Dies ist unter anderem dann relevant, wenn der Einfluss bestimmter Vernetzungsmöglichkeiten auf die Beziehungen zwischen denjenigen bestimmt werden soll, die das durch eine Veranstaltung gebotene Forum nutzen. Auch hier kann es wieder sinnvoll sein, die Veranstaltungen als Objekte in das Netzwerk aufzunehmen, und im Einzelnen können sich Abgrenzungsschwierigkeiten hinsichtlich der Teilnahme ergeben, die ähnlich wie die Probleme bei der Berufszuordnung durch regelgeleitete Entscheidungsprozesse abgemildert werden können.

Die Liste möglicher Differenzierungskriterien lässt sich beliebig fortsetzen. So können etwa zur Evaluation von Quotenregelungen auch Alter und Geschlecht als Auswahlfaktoren von Interesse sein. Entscheidend ist, dass die Kriterien mit Blick auf ihre Eignung für die Beantwortung der gegebenen Untersuchungsfrage gewählt, transparent gemacht und konsequent angewendet werden.

#### b) *Kollektive*

Anstatt Netzwerke mit einzelnen Menschen als Objekten zu definieren, kann man auch Kollektive als Objekte darstellen. Ein Kollektiv kann eine Menge von Rechtssubjekten sein, an die das Recht punktuell anknüpft, die aber nicht selbst Rechtssubjekt ist (z.B. eine Familie) – es kann aber auch eine abstrakte Einheit mit institutionalisierten Willensbildungsprozessen sein, die nur aufgrund einer rechtlichen Verfassung existiert und selbst Rechtssubjekt ist (z.B. ein Verein). In der Literatur zur Netzwerkforschung im Recht spielen bisher vor allem Kapitalgesellschaften, Gerichte und Staaten eine Rolle. Für die juristische Netzwerkforschung können aber weit mehr Kollektive von Interesse sein. Beispielfhaft seien genannt:

- Ehepaare, Haushalte oder Familien;
- Personen- und Kapitalgesellschaften, Konzerne oder Kartelle;
- Stiftungen und Vereine, Körperschaften und Anstalten des Öffentlichen Rechts;
- Gerichte, Behörden und ihre organisatorischen Einheiten;
- Staaten, supranationale und internationale Organisationen.

Häufig wird man die Menge der betrachteten Kollektive weiter einschränken wollen, was analog zum Vorgehen bei Individuen durch die Orientierung an ausgewählten Variablen erfolgen kann. Bei Kollektiven, die nicht durch Rechtsakt(e) konstituiert sind, kann die Existenz einzelner Objekte problematisch sein. Die Hauptschwierigkeit bei der Analyse von Netzwerken aus Kollektiven ist aber die Erkennung und Zuordnung von Beziehungen.<sup>89</sup>

<sup>89</sup> Dazu noch S. 57–59.

## 2. Rechtstexte

Auch Rechtstexte können als Objekte in Netzwerken repräsentiert werden. Hier lassen sich für die Zwecke der juristischen Netzwerkforschung fünf Textgattungen unterscheiden: Gerichtsentscheidungen (a)), Gesetzestexte (b)), Verwaltungsentscheidungen (c)), Verträge (d)) und juristische Literatur (e)).

### a) Gerichtsentscheidungen

Wie im Literaturüberblick deutlich geworden ist, hat die Darstellung von einzelnen Gerichtsentscheidungen als Objekte insbesondere im *Common Law* eine gewisse Tradition.<sup>90</sup> Die Eingrenzung auf bestimmte Gerichtsentscheidungen kann hier unter anderem anhand der folgenden Kriterien erreicht werden:

- *Autor.* Ohne die Eingrenzung auf bestimmte Autoren lassen sich Netzwerkstudien zu Gerichtsentscheidungen kaum durchführen. Wer aber ist der Autor einer Gerichtsentscheidung? Die überwiegende Anzahl der Studien zur Netzwerkforschung im Recht antwortet schlicht: „das Gericht“,<sup>91</sup> wobei nicht immer von vornherein klar ist, was genau mit „Gericht“ gemeint ist.<sup>92</sup> Je nach Untersuchungsinteresse kann man allerdings auch „der entscheidende Senat“, „die entscheidende Kammer“, „die entscheidenden Richter“ oder sogar „der Richter, der die Entscheidung verfasst hat“,<sup>93</sup> als Antwort geben.
- *Art.* Die Art der Gerichtsentscheidung ist ein weiteres, in der Literatur beliebtes Eingrenzungskriterium.<sup>94</sup> Ob etwa die Ausgrenzung von unbegründeten Beschlüssen aus der Grundgesamtheit aller Entscheidungen (eines bestimmten Autors) zweckmäßig ist, kann allerdings nur in Abhängigkeit von der konkreten Untersuchungsfrage entschieden werden. Soll beispielsweise das Zusammenspiel von Netzwerken aus Texten und sozialen Netzwerken untersucht werden, so könnte es durchaus von Interesse sein, ob eine bestimmte Konstellation von Richtern auffällig häufig unbegründet beschließt.
- *Gerichtsbearbeitung.* Gerade in Deutschland kann auch die Gerichtsbearbeitung (z.B. Arbeitsgerichtsbearbeitung, Sozialgerichtsbearbeitung oder allgemeine Verwaltungsgesichtsbearbeitung) ein hilfreiches Eingrenzungskriterium sein. Dies gilt insbesondere

<sup>90</sup> Dazu oben S. 13–14.

<sup>91</sup> Dazu bereits S. 13–14.

<sup>92</sup> Das zeigt etwa die Bemerkung von *Frankenreiter* betreffend die Untersuchung von *Derlén und Lindholm* zur Rechtsprechung des EuGH: „More precisely, the authors seem to use only judgments issued between 1954 and May 2011 by the highest branch of the CJEU, the Court of Justice.“ *Frankenreiter*, *German Law Journal* 18 (2017), S. 687 (688 Fn. 6).

<sup>93</sup> Präziser noch: „der Richter, der die Abfassung der Entscheidung verantwortet hat“.

<sup>94</sup> So liest man beispielsweise bei *Fowler u.a.*, *Political Analysis* 15 (2007), S. 324 (327 Fn. 7): „We excluded the 412 in-chambers’ opinions published in the *U.S. Reports* because they are not a decision on the merits, but rather have to do with such matters as extensions of time, injunctions, and stays.“

dann, wenn der Forscher sich für die Strukturen eines bestimmten Rechtsgebiets interessiert. Zwar könnte man auch versuchen, das Rechtsgebiet selbst zum Eingrenzungskriterium zu machen. Dieses Vorgehen fordert aber den Vorwurf mangelnder Objektivität geradezu heraus. Denn wie definiert man ein Rechtsgebiet?

- *Rechtsfrage*. Ähnlich problematisch ist die Eingrenzung auf Entscheidungen zu einer bestimmten Rechtsfrage oder – nach den Merkmalen in der *Supreme Court Database* – auf eine bestimmte *issue*. Ob eine Gerichtsentscheidung eine bestimmte Rechtsfrage betrifft, ist oft eine Wertungsentscheidung, deren konkrete Ausprägung die Ergebnisse sensibler Netzwerkmaße verzerren kann. Weniger subjektiv ist die Einschränkung auf Gerichtsentscheidungen, die bestimmte Normen betreffen. Auch hier stellt sich allerdings die Frage nach der Absicherung gegen Über- und Unterinklusivität in der Netzwerkdefinition:<sup>95</sup> Wann betrifft eine Entscheidung eine Norm?
- *Zeit*. Objektiv kann schließlich wiederum die Zeit als Eingrenzungskriterium herangezogen werden, wobei man in der Regel an das Entscheidungsdatum anknüpfen wird. Studien, deren Autoren keine eigene Datensammlung unternehmen, leiten ihren Betrachtungszeitraum allerdings in der Regel aus den Daten ab, die ihnen zur Verfügung stehen.<sup>96</sup>

#### b) Gesetzestexte

Aus der Perspektive des *Civil Law* scheint neben der Analyse von Gerichtsentscheidungen die Untersuchung von Gesetzestexten als Objekte in Netzwerken besonders interessant. Wie die Abbildungen 1.5 (S. 7) und 1.6 (S. 8) im ersten Kapitel zeigen, kann die juristische Netzwerkforschung die in Gesetzestexten enthaltenen Strukturen – etwa Hierarchien oder Verweise (Einschluss- oder Verweisungsbeziehungen) – sichtbar machen. Sie kann diese analysieren und die Strukturen unterschiedlicher Gesetzestexte oder unterschiedlicher Fassungen desselben Gesetzestexts miteinander vergleichen, um typische Muster zu erkennen und Veränderungen über die Zeit zu erfassen. Welche Darstellung am sinnvollsten ist, hängt von der Untersuchungsfrage und dort insbesondere davon ab, welche Beziehungen zwischen Gesetzestexten bzw. den in ihnen enthaltenen Normen analysiert werden sollen.

Die Eingrenzung der Objektmenge kann bei Gesetzestexten ähnlich wie bei Gerichtsentscheidungen zunächst nach Zeit und Autor erfolgen. Sofern man sich nur für die Gesetzgebung in einem bestimmten Bereich interessiert, steht man

<sup>95</sup> Eine Netzwerkdefinition ist *überinklusiv*, wenn sie Objekte oder Beziehungen erfasst, die der Definierende *nicht* erfassen wollte. Sie ist *unterinklusiv*, wenn sie Objekte oder Beziehungen *nicht* erfasst, die der Definierende erfassen wollte.

<sup>96</sup> Beispielsweise untersuchen *Cross, Smith und Tomarchio* das Zitienetzwerk des *U.S. Supreme Court* von 1885 bis 2005 – auf Basis eines Datensatzes, der von *LexisNexis* zur Verfügung gestellt wurde; *Cross, Smith und Tomarchio*, *Emory Law Journal* 57 (2008), S. 1227 (1237).

wieder vor einem bereits erwähnten Zuordnungsproblem: Wie identifiziert man ein Rechtsgebiet? Anders als bei Gerichtsentscheidungen lässt sich dieses Problem hier nicht durch Anknüpfung an eine etwaige Gerichtsbarkeit umgehen. Für Deutschland ist allerdings denkbar, auf die in Anspruch genommene Kompetenzgrundlage (Ziffern der Art. 73 I, 74 I GG) abzustellen. Will man Rechtsgebiete also als Kriterium zur Eingrenzung der betrachteten Textmenge nutzen, muss man im Vorfeld Kriterien festlegen, anhand derer sich einzelne Gesetzestexte einzelnen Rechtsgebieten zuordnen lassen – und damit *deduktiv* vorgehen. Die Netzwerkforschung stellt jedoch auch Methoden bereit, mit denen sich Rechtsgebiete *induktiv* – also aus den in (nicht nur Gesetzes-)Texten gefundenen Strukturen – bestimmen lassen,<sup>97</sup> sodass eine Vorfeldeingrenzung nach Rechtsgebiet unter Umständen unterbleiben kann.

### c) Verwaltungsentscheidungen

Als Textprodukte der Exekutive kann die juristische Netzwerkforschung auch Verwaltungsentscheidungen, allen voran Verwaltungsakte (bzw. deren funktionale Äquivalente in anderen Jurisdiktionen), in den Blick nehmen. Die Netzwerkforschung im Recht hat sich bisher auf Patente als Objekte in Patentnetzwerken konzentriert, wo sie sich auch im Praxiseinsatz befindet.<sup>98</sup> Patente können nach Zeit und zuständigem Patentamt verhältnismäßig leicht eingegrenzt werden. Will man die Untersuchung auf einzelne Technologiefelder beschränken, so kann man sich je nach Untersuchungszeitraum und betrachtetem Patentamt an der *International Patent Classification (IPC)*, der *Cooperative Patent Classification (CPC)* oder der *US Patent Classification (USPC)* orientieren. Da mit jedem Patent zahlreiche Kontextinformationen veröffentlicht werden, ist auch eine gezielte Einschränkung nach Anmeldern, Erfindern oder Herkunftsland der Verfahrensbeteiligten möglich. Für die Rechtswissenschaft sind Patentnetzwerke vor allem im Vergleich mit anderen juristischen Zitiernetzwerken (z.B. Rechtsprechungszitiernetzwerken) interessant.

Doch neben Patenten können noch weitere Arten von Verwaltungsentscheidungen zum Gegenstand juristischer Netzwerkforschung gemacht werden – etwa, um den Ablauf bestimmter Verwaltungsverfahren in der Praxis nachzuvollziehen und festzustellen, wo mögliche Reformen ansetzen könnten. Die betrachteten Entscheidungen lassen sich hier nach Autor (entscheidende Behörde) und nach Zeit ebenso eingrenzen wie nach Entscheidungsart (z.B. Baugenehmigung). Sofern entsprechende Daten vorliegen, kann man die betrachteten Entscheidungen auch nach Entscheidungsbetroffenen (z.B. bestimmte Bauträger oder Interessenverbände) auswählen. Je nach Untersuchungskontext kann es sich ferner lohnen, Dokumente in die Analyse einzubeziehen, die bei den Behörden zur Erfüllung

---

<sup>97</sup> Dazu unten S. 147–159.

<sup>98</sup> Zur einschlägigen Literatur schon S. 26–27. In der Praxis werden visuelle Darstellungen und quantitative Analysen von Patentnetzwerken unter anderem bei der Patentrecherche genutzt.

von Anzeigepflichten eingereicht werden (z.B. Versammlungsanzeigen). Deren Inhalte lassen sich oft geographisch verankern (z.B. kann sich eine Versammlungsanzeige auf einen bestimmten Ort beziehen) und hierüber gegebenenfalls weiter filtern. Als Objekte in räumlich eingebetteten Netzwerken machen Verwaltungsentscheidungen und Anzeigen so eine Kartographie der Grundrechtsausübung möglich.

#### d) Verträge

Die Untersuchung von Verträgen könnte einen weiteren Schwerpunkt der juristischen Netzwerkforschung bilden. Hier scheinen insbesondere zwei Arten von Verträgen interessant:

- *Völkerrechtliche Verträge*. Sowohl bilaterale Abkommen, etwa *Bilateral Investment Treaties* (BITs) und Doppelbesteuerungsabkommen (DBA), als auch multilaterale Abkommen lassen sich als Objekte in Netzwerken darstellen.<sup>99</sup> Dabei wird die Repräsentation der Verträge *allein* allerdings selten informativ sein. Vielmehr wird man zur Definition der Objektmenge die Menge der Verträge gemeinsam mit der Menge aller Staaten betrachten. Wie man dann weiter verfährt, hängt davon ab, ob Beziehungen zwischen Verträgen oder Beziehungen zwischen Staaten im Fokus der Analyse stehen sollen.
- *Privatrechtliche Verträge*. Die Untersuchung privatrechtlicher Verträge kann vor allem aus regulatorischer Perspektive interessant sein. Mit der Untersuchung von Vertragsnetzwerkstrukturen *im Recht* geht dabei die Frage nach der adäquaten Regulierung dieser Strukturen *durch Recht* einher, die in der Literatur zur Netzwerkforschung im Recht insbesondere im Kontext der *financial networks* diskutiert wird.<sup>100</sup>

Die Menge der untersuchten Verträge wird – unabhängig davon, ob es sich um völkerrechtliche oder privatrechtliche Verträge handelt – fast immer durch Anknüpfung an Beteiligte mit bestimmten Attributen (z.B. Banken mit einer bestimmten Bilanzsumme), die Nutzung bestimmter Plattformen (z.B. einer Börse) oder ein bestimmtes Thema (z.B. den Investitionsschutz) erfolgen müssen. Im Eingrenzungsprozess werden oft Wertungsentscheidungen erforderlich, die vom Forscher systematisch getroffen und transparent gemacht werden sollten.

---

<sup>99</sup> Siehe dazu die Nachweise in Fn. 15 (S. 16).

<sup>100</sup> Dazu bereits S. 22–28.

### e) Literatur

Für die Selbstreflexion der Rechtswissenschaft kann ferner die Repräsentation von Literaturbeiträgen als Objekte in einem Netzwerk interessant sein. Je nachdem, welche Beziehungen untersucht werden, können dabei nur die Beiträge selbst, die Beiträge und ihre Autoren oder aber auch nur die Autoren interessante Objekte sein. Mithilfe der Repräsentation von Literaturbeiträgen als Objekte kann man den Diskursstrukturen der Rechtswissenschaft nachspüren, Rezeptionsprozesse sichtbar machen, Forschungsfelder, Forschungslücken und Forschungstrends identifizieren oder – etwa durch Betrachtung eines Netzwerks aus Literaturbeiträgen und Gerichtsentscheidungen – den Einfluss der Rechtswissenschaft auf die Praxis untersuchen. Dabei kann man auf langjährige Forschungserfahrungen in der Bibliometrie (*bibliometrics*) und der Szientometrie (*scientometrics*) zurückgreifen.<sup>101</sup>

Ein- und Abgrenzungsprobleme ergeben sich hier vor allem im Umgang mit Sammelwerken (etwa Festschriften, Tagungsbänden oder Kommentaren), bei denen nicht von vornherein ausgemacht ist, ob diese in einem Netzwerk durch ein Objekt oder durch mehrere Objekte dargestellt werden sollten. Die Entscheidung richtet sich wieder nach der Untersuchungsfrage. Diese bestimmt auch, wie spätere Untersuchungsergebnisse zu interpretieren sind.

### 3. Rechtsbegriffe

Neben Rechtssubjekten und Rechtstexten lassen sich schließlich auch Rechtsbegriffe als Objekte in Netzwerken betrachten. Hier ist die Definition allerdings besonders problematisch: Was kennzeichnet einen Rechtsbegriff (*legal concept*)? Für eine objektiv ansetzende Objektdefinition kommen vor allem juristische Ausdrücke (*legal terms*) in Betracht, die in Rechtstexten enthalten sind. Das Definitionsergebnis kann dann sehr unterschiedlich ausfallen, je nachdem, von welchen Rechtstexten man ausgeht – etwa von Gesetzestexten, Gerichtsentscheidungen oder wissenschaftlichen Publikationen – und nach welcher Methode man die Termini auswählt, deren Beziehungen zueinander untersucht werden sollen:

- Die einfachste, wengleich subjektivste, Möglichkeit besteht darin, von der persönlichen Wahrnehmung des Untersuchenden auszugehen. Die Ergebnisse einer Analyse auf dieser Basis sind dann vor allem rechtswissenschaftssoziologisch interessant – im Vergleich zu anderen Analysen, die von anderen Forschern auf ebenso subjektiver Basis erstellt wurden.

---

<sup>101</sup> Visuell eindrucksvolle Einführungen in Bibliometrie und Szientometrie bieten *Börner*, Atlas of Science, 2010, und *Börner*, Atlas of Knowledge, 2015. Einen Überblick zur historischen Entwicklung von Bibliometrie, Szientometrie und der mit diesen verwandten Informatik (*informetrics*) geben *Hood und Wilson*, *Scientometrics* 52 (2001), S. 291–314, mit zahlreichen weiteren Nachweisen.

- Einen geringeren Grad an Subjektivität verspricht die Anknüpfung an die Wahrnehmung des rechtswissenschaftlichen Diskurses. Hier stellt sich allerdings das Problem, dass der Diskurs selbst keine Meinung hat, sondern meist mehrere verschiedene Begriffsbildungsvorschläge im wissenschaftlichen Austausch miteinander um die Gunst der Diskursteilnehmer ringen. Diese Schwierigkeit lässt sich nicht beseitigen – man kann aber vermeiden, dass Analyseergebnisse missverstanden werden, indem man die Prämissen, die man seiner Begriffsdefinition zugrunde legt, transparent macht und den Rezipienten der eigenen Arbeit so ermöglicht, sowohl die Wahl der Prämissen als auch die innere Schlüssigkeit dieser Prämissen kritisch zu prüfen.
- Eine weitere, möglicherweise noch weniger subjektive Vorgehensweise besteht schließlich darin, die Begriffe selbst aus der systematischen *statistischen* Analyse großer Mengen juristischer Texte zu gewinnen. Je nach Untersuchungsziel können die Texte, die der quantitativen Analyse unterzogen werden, auf einen Autor oder Autorentyp (z.B. Gesetzgeber) beschränkt oder aber verschiedener Urheberschaft sein. Zur Identifikation der Begriffe kann dann auf Techniken des *Natural Language Processing* (NLP) zurückgegriffen werden.<sup>102</sup>

Sofern man nicht die Beziehungen aller – wie auch immer konturierter – Rechtsbegriffe zueinander analysieren will, stellen sich zusätzlich zu den dargestellten Schwierigkeiten im Definitionsprozess die bereits mehrfach angesprochenen Fragen der Rechtsgebietsabgrenzung.

Die voranstehende Systematisierung möglicher Objekte in der juristischen Netzwerkforschung zeigt, dass die Definitionsmöglichkeiten fast unbegrenzt sind. Gleichzeitig tun sich bereits hier zahlreiche Probleme auf, wenn man die Facetten der potenziellen Objekte genau betrachtet. Die Schwierigkeiten vergrößern sich noch, wenn man neben den Objekten selbst auch die Beziehungen in den Blick nimmt, die zwischen ihnen bestehen können.

## II. Beziehungen

Sieht man davon ab, dass sich auch Beziehungen selbst als Objekte auffassen lassen,<sup>103</sup> so gehen die Möglichkeiten der Beziehungsdefinition notwendigerweise weit über die Möglichkeiten der Objektdefinition hinaus: Erstens können schon zwischen den Objekten einer Menge verschiedene Beziehungen bestehen. Zweitens können zwischen Objekten in verschiedenen Mengen jeweils gleiche Beziehungen bestehen. Drittens können auch zwischen Objekten aus mehreren ver-

<sup>102</sup> Eine leicht zugängliche Einführung bieten *Jurafsky und Martin*, *Speech and Language Processing*, 2009.

<sup>103</sup> Lässt man Beziehungen als Objekte zu, so kann man auch die Beziehungen zwischen Beziehungen betrachten – oder die Beziehungen zwischen Beziehungen zwischen Beziehungen, oder ... Dies führt zu einer unendlichen Rekursion. Allgemein zum Rekursionsproblem noch unten S. 59–61.

schiedenen Mengen und zwischen den verschiedenen Mengen selbst Beziehungen bestehen.

Zur Ordnung der Vielzahl möglicher Beziehungsdefinitionen werden nachfolgend drei Unterscheidungen eingeführt.<sup>104</sup> Diese Unterscheidungen orientieren sich maßgeblich an den Strukturen der betrachteten Beziehungen. Sie werden als *Beziehungsgrad* (1.), *Beziehungsprofil* (2.) und *Beziehungsform* (3.) bezeichnet.

### 1. Beziehungsgrad

Der Beziehungsgrad zeigt – wie der Grad einer Verwandtschaftsbeziehung – die Unmittelbarkeit bzw. Mittelbarkeit einer Beziehung an. Die Bestimmung des Grades einer Beziehung richtet sich nach folgender Leitfrage: Wie viele Beziehungen sind insgesamt zu betrachten, um festzustellen, ob zwischen zwei Objekten die gewünschte Beziehung besteht? Die unmittelbarste Beziehung ist jene ersten Grades, für deren Feststellung nur eine Beziehung betrachtet werden muss. Es folgt die Beziehung zweiten Grades, dann jene dritten Grades.

Prinzipiell ist die Zahl der Beziehungsgrade nach oben nur durch die Vorstellungskraft des Untersuchenden begrenzt. Je höher der Beziehungsgrad, desto höher ist jedoch auch das Abstraktionsniveau und desto geringer ist die intersubjektive Überprüfbarkeit sowohl der Beziehungsdefinition als auch der Ergebnisinterpretation. Die hiesige Darstellung beschränkt sich daher auf die ersten drei Beziehungsgrade, welche die bestehende Netzwerkforschung im Recht bereits abdecken. Die Struktur der Beziehungen dieser drei Grade ist in Abbildung 2.1 (S. 38) dargestellt.

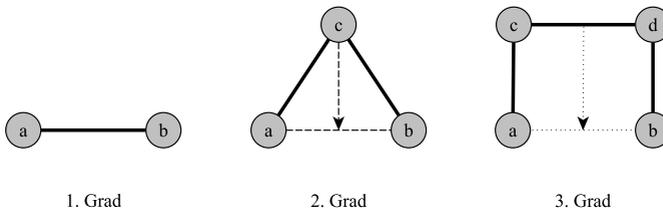


Abbildung 2.1: Strukturen der Beziehungen unterschiedlichen Grades

Anhand dieser Darstellung lassen sich für Beziehungen der ersten drei Grade einige Charakteristika herausarbeiten:

- 1. Grad. Beziehungen ersten Grades sind schon dann erkennbar, wenn man nur eine einzige Beziehung betrachtet. Konkrete Beispiele sind „a ist Mutter von b“ und „a zitiert b“.

<sup>104</sup> Ausführlich zu den Schwierigkeiten der Beziehungsdefinition für die allgemeine Netzwerkforschung *Zweig*, *Network Analysis Literacy*, 2016, S. 313–350.

- 2. *Grad*. Beziehungen zweiten Grades sind nur erkennbar, wenn man insgesamt zwei Beziehungen betrachtet. Eine Beziehung zweiten Grades zwischen zwei Objekten kann genau dann angenommen werden, wenn beide in der gleichen Beziehung zu *demselben* anderen Objekt stehen. Beziehungen zweiten Grades basieren auf genau *einer* Konjunktion (einem „und“). Konkrete Beispiele sind „*a* und *b* sind beide Eltern von *c*“ und „*a* und *b* werden beide von *c* zitiert“.
- 3. *Grad*. Beziehungen dritten Grades kann man nur finden, indem man insgesamt (mindestens) drei Beziehungen betrachtet. Zwischen zwei Objekten wird genau dann eine Beziehung angenommen, wenn diese jeweils in der gleichen Beziehung zu *zwei verschiedenen* anderen Objekten stehen und diese anderen Objekte wiederum in einer bestimmten Beziehung (in der Regel zweiten Grades) zueinander stehen. Beziehungen dritten Grades sind also Beziehungskonstruktionen auf Basis *zweier* Konjunktionen. Sie werden häufig eingesetzt, um abstrakte Ähnlichkeitskonzepte auszudrücken. Konkrete Beispiele hierfür sind Aussagen wie „*a* enthält Wort *c* und *b* enthält Wort *d*, und die Wörter *c* und *d* sind Synonyme, also haben *a* und *b* etwas gemeinsam“ oder „*a* hat in Test *X* einen Wert von *c*, *b* hat in Test *X* einen Wert von *d*, die Werte *c* und *d* liegen hinreichend nah beieinander, deswegen sind sich *a* und *b* ähnlich“. An diesen Formulierungen wird bereits deutlich, wie schwierig – und angreifbar – die Definition und Interpretation von Beziehungsstrukturen wird, wenn man sich auf höhere Abstraktionsniveaus begibt.<sup>105</sup>

## 2. Beziehungsprofil

Beziehungen lassen sich nicht nur nach ihrem Grad, sondern auch nach ihrem Profil unterscheiden. Als Beziehungsprofil wird hier die *inhaltliche* und *zeitliche* Struktur des Verhältnisses der Beziehungsbeteiligten bezeichnet, das in einer Beziehungsdefinition zum Ausdruck kommen soll. Nachfolgend sind die Profile einiger für die juristische Netzwerkforschung interessanter Beziehungen zusammengestellt, gruppiert nach ihrem jeweiligen Grad.

---

<sup>105</sup> Dennoch werden in der Literatur zum Teil Beziehungen dritten Grades untersucht, etwa bei *Siems*, *Arizona Journal of International & Comparative Law* 27 (2010), S. 747–784, und *Siems*, *Journal of Institutional Economics* 12 (2016), S. 579–602 (jeweils Beziehungen zwischen Staaten auf Basis der Differenzen ihrer Werte in einem eigens konstruierten Index).

## 1. Grad.

Bei den Beziehungen ersten Grades lohnt es sich, zwischen Produktions-, Transaktions-, Diffusions-, Reaktions- und Statusbeziehungen zu unterscheiden. Die Profile dieser Beziehungen sind in Abbildung 2.2 (S. 40) zusammengestellt, wobei  $t_0$  für die Ausgangssituation (noch keine Beziehung vorhanden) und  $t_1$  für die Beziehungssituation steht.

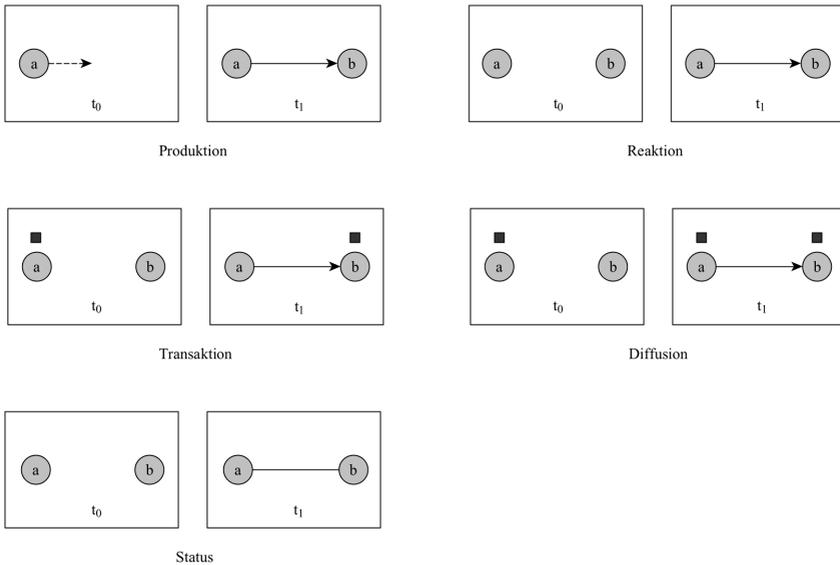


Abbildung 2.2: Profile typischer Beziehungen ersten Grades

Der Darstellung liegen folgende Erwägungen zugrunde:

- *Produktion.* Produktionsbeziehungen zeichnen sich dadurch aus, dass ein Objekt ein anderes erst erschafft. Im Ausgangszustand  $t_0$  existieren nur das produzierende Objekt und der Beginn eines Produktionsprozesses, der im Übergang von  $t_0$  zu  $t_1$  das produzierte Objekt hervorbringt. Produktionsprozesse lassen sich in Graphen durch Pfeile (gerichtete Linien) darstellen. Ein praktisches Beispiel für einen Produktionsprozess, das verdeutlicht, dass die zwischen  $t_0$  und  $t_1$  liegende Zeitspanne mitunter beträchtlich sein kann, ist die Erarbeitung einer Dissertation durch eine Doktorandin.
- *Transaktion.* Für Transaktionsbeziehungen ist charakteristisch, dass ein – wie auch immer geartetes – Gut im Übergang von  $t_0$  zu  $t_1$  von einem Objekt zu einem anderen wandert, wobei Wanderung bedeutet, dass es sich in  $t_1$  nicht mehr (auch) beim Ursprungsobjekt befindet.

Auch Transaktionsbeziehungen lassen sich durch Pfeile darstellen; laufen Pfeile in beide Richtungen, so kann man von einer *Interaktion* sprechen. Transaktionsbeziehungen, bei denen Eigentum übertragen werden soll, werden von den Parteien (oft unbewusst) durch Austauschverträge geregelt. Diese bilden (nach der deutschen Konzeption) den Rechtsgrund für den Eigentumserwerb; sie können ihrerseits als Statusbeziehungen modelliert werden.<sup>106</sup> Klassische Beispiele für eine Menge von Transaktionsbeziehungen sind die Wanderung eines Gerichtsfalles durch die Instanzen und der Bargeldumlauf.<sup>107</sup>

- *Diffusion*. Diffusionsbeziehungen zeichnen sich dadurch aus, dass die im Intervall von  $t_0$  nach  $t_1$  übertragene Substanz sich in  $t_1$  sowohl beim Ursprungs- als auch beim Zielobjekt befindet, was sie von den Transaktionsbeziehungen wesentlich unterscheidet.

Diffusionsbeziehungen werden in der Regel als Pfeile dargestellt, um die Diffusionsrichtung verfolgen zu können. Sie bewirken die Verbreitung von Krankheiten ebenso wie jene von Informationen und bauen in der Regel auf einer anderen Beziehung (z.B. einer Kommunikationsbeziehung) auf. Mit fortschreitender Digitalisierung findet ein immer größerer Teil der wirtschaftlichen Wertschöpfung über die Produktion von diffusionsfähigen Gütern statt. Hier steht das Recht vor einem Beziehungsproblem, wenn es versucht, Diffusionsbeziehungen zu verhindern und Transaktionsbeziehungen zu erzwingen.<sup>108</sup>

- *Reaktion*. Für Reaktionsbeziehungen (hier im Sinne von Reaktion *auf*, nicht Reaktion *mit*, verstanden) ist entscheidend, dass sie an den Charakteristika des Objekts, auf das reagiert wird, nichts ändern. Dadurch unterscheiden sie sich von Transaktionsbeziehungen. Diese graphisch schwer darstellbare Nuance lässt sich leichter erfassen, wenn man die Ausprägungen von Reaktionsbeziehungen betrachtet, die für die juristische Netzwerkforschung besonders relevant sind.

<sup>106</sup> Zu Statusbeziehungen noch unten S. 43.

<sup>107</sup> Die wesentlichen Schwierigkeiten im Dunstkreis der Bargeldabschaffung lassen sich also als Beziehungsprobleme verstehen: Dem Staat ist daran gelegen, sein Monopol auf die Produktionsbeziehung aufrecht zu erhalten und zu vermeiden, dass Geld zum Gegenstand von Diffusionsbeziehungen wird – Private sollen hinsichtlich des Geldes nur Transaktionsbeziehungen eingehen können. Während dies beim Bargeld durch physische Sicherheitsmerkmale weitgehend sichergestellt wird, sind bei digitalen Transaktionen und „digitalen Währungen“ (z.B. Bitcoin) zusätzliche Mechanismen erforderlich, um zu verhindern, dass Geld doppelt ausgegeben werden kann (*double spending problem*); dazu etwa *Barber u.a.*, *Bitter to Better – How to Make Bitcoin a Better Currency*, in: *Keromytis* (Hrsg.), *Financial Cryptography and Data Security*, 2012, S. 399 (402–403); *Alberts und Fry*, *Boston University Journal of Science and Technology Law* 21 (2015), S. 1 (3).

<sup>108</sup> Ein plastisches Beispiel bieten die Schwierigkeiten im Umgang mit *File Sharing*; dazu aus der empirisch grundierten juristischen Literatur *Strahilevitz*, *Virginia Law Review* 89 (2003), S. 505–595; *Feldman und Nadler*, *San Diego Law Review* 43 (2006), S. 577–618; *Liebowitz*, *Journal of Law and Economics* 49 (2006), S. 1–28.

- *Zitate und Verweise.* Zitate und Verweise sind für die zitierende bzw. verweisende Einheit – sei es nun ein Rechtstext oder ein Rechtssubjekt – zwei wesentliche Möglichkeiten, auf die in einer anderen Einheit enthaltenen oder von dieser propagierten Informationen zu reagieren. Dabei ändert sich für das zitierte Objekt bzw. das Objekt, auf das verwiesen wird, nichts über die Tatsache hinaus, nun auch Zitier- oder Verweisobjekt zu sein.

Die Funktionen von Zitaten und Verweisen sind im Einzelnen umstritten, wobei die Diskussion darunter leidet, dass schon die Vokabeln nicht einheitlich verwendet werden.<sup>109</sup> Sie hängen zudem davon ab, wer oder was durch wen oder was zitiert wird oder auf wen oder was durch wen oder was verwiesen wird. Dennoch scheint es sinnvoll, Zitate und Verweise voneinander abzugrenzen und als Oberbegriff „Referenzen“ zu verwenden. Man kann dann Zitate von Verweisen anhand des Verhaltens unterscheiden, das sie bei potenziellen Rezipienten hervorrufen sollen:

Während der Verweis den Rezipienten dazu anhält, sich beim Verweisziel weitere Informationen zu holen, zielt das Zitat darauf ab, dem Rezipienten ein solches Verhalten gerade zu ersparen, indem es auf einen (vielleicht auch nur vermeintlichen oder simulierten) gemeinsamen Erfahrungshorizont oder Wissensfundus Bezug nimmt. Das Zitat hat demnach eine (*text*)*schließende* Funktion, während dem Verweis eine (*text*)*öffnende* Funktion zukommt.

Die Netzwerkforschung im Recht differenziert in der Regel nicht zwischen Zitaten und Verweisen und die ihnen jeweils entsprechenden Beziehungen werden in Graphen durch Pfeile dargestellt. Diese Pfeile zeigen für gewöhnlich wie in Abbildung 2.2 (S. 40) vom Zitierenden zum Zitierten.<sup>110</sup>

Die Gleichsetzung von Zitaten und Verweisen mag bei der Analyse wissenschaftlicher Publikationen aus der Not gerechtfertigt sein, dass sich ohne manuelle Inspektion jeder Reaktionsbeziehung keine eindeutige Klassifikation vornehmen ließe (man denke nur an die äußerst heterogene Handhabung der Kürzel „s.“ und „vgl.“). Gerade bei der Untersuchung juristischer Textmengen, die nicht der Wissenschaft entspringen, kann der Unterschied im Informationsfluss aber systematisch erfasst werden: So lassen sich Bezug-

<sup>109</sup> Die Diskussion findet vor allem in den USA statt und konzentriert sich dort auf Zitate in Gerichtsentscheidungen (siehe dazu bereits die Nachweise in Fn. 8 [S. 14]). Im deutschen Sprachraum erfolgt nur selten eine explizite Auseinandersetzung mit den Funktionen von Zitaten, etwa bei *Nußberger*, JZ 61 (2006), S. 763–770, und bei *Hamann*, RW 2014, S. 501–534.

<sup>110</sup> In der Szientometrie werden die Pfeile zum Teil in entgegengesetzter Richtung gezeichnet, so etwa bei *Börner*, Atlas of Science, 2010, S. 121. Das liegt daran, dass Zitate und Verweise dort als Wege der Diffusion von Wissen oder Ideen begriffen werden. Dieser Deutung von Zitaten und Verweisen liegen allerdings einige Annahmen hinsichtlich des Verhaltens der Zitierenden bzw. Verweisenden zugrunde, deren Geltung nicht ohne Weiteres einsichtig ist – etwa, dass es bei Zitaten um Inhalte geht. Anlass zu Zweifeln bieten hier etwa *Simkin und Roychowdhury*, *Scientometrics* 62 (2005), S. 367–384. Das Thema „Referenzmotivationen“ soll hier nicht weiter vertieft werden; einen Einstieg in die einschlägige Literatur ermöglichen *Bornmann und Daniel*, *Journal of Documentation* 64 (2008), S. 45–80.

nahmen auf Gesetzesnormen in Gerichtsentscheidungen als Zitate auffassen, während sie innerhalb anderer Gesetzestexte – insoweit übereinstimmend mit der Sprachkonvention – Verweise darstellen. Dies ermöglicht der juristischen Netzwerkforschung, die Differenzen zwischen Zitaten und Verweisen in ihren Analysen zu berücksichtigen.

- *Evaluationen.* Neben der Zitat- oder Verweisungsbeziehung kommt die Reaktionsbeziehung in Form der Evaluationsbeziehung als Untersuchungsgegenstand der juristischen Netzwerkforschung in Betracht. Evaluation wird hier verstanden als Beurteilung eines Objekts durch ein anderes Objekt. Für das evaluierte Objekt ändert sich erst einmal nichts, mitunter werden ihm weder Tatsache noch Inhalt der Evaluation bekannt.

Evaluationen spielen bei der Analyse kleiner und mittelgroßer sozialer Netzwerke eine besondere Rolle. Hier werden die relationalen Daten, die der Netzwerkrepräsentation zugrunde liegen, nicht selten mit klassischen soziologischen Methoden erhoben.<sup>111</sup> Sie werden dann ebenfalls als Pfeile dargestellt, wobei die Richtung jeweils der Evaluationsrichtung entspricht (ein Pfeil von *a* nach *b* zeigt dann an, dass *b* von *a* in einer bestimmten, näher zu spezifizierenden Hinsicht beurteilt wurde).

Klassische Beispiele für Reaktionsbeziehungen in Form der Evaluation sind Situationen, in denen Bücher oder Experten empfohlen oder Angaben dazu gemacht werden, wie die eigene Beziehung zum Evaluierten einzustufen ist (z.B. Freundschaft oder Feindschaft).

Das Evaluationskonzept lässt sich aber auch in abstrakteren Kontexten einsetzen, etwa um das Verhältnis der Evaluierenden zu Propositionen – Wissen, Glauben, bewusste oder unbewusste Unkenntnis – darzustellen und zu untersuchen.<sup>112</sup> Dies kann für die juristische Netzwerkforschung nützlich sein, soweit es ihr darum geht, die Diskursentwicklung im Recht – sei es in Gesetzgebung, Rechtsprechung oder Rechtswissenschaft – zu verstehen.

- *Status.* Unter Statusbeziehungen lassen sich untechnisch alle Beziehungen zusammenfassen, deren Einrichtung zwischen den Beziehungsparteien unmittelbar einen Katalog von Rechten und Pflichten etabliert, der sich in wesentlichen Teilen *nicht* im Vokabular der Transaktionsbeziehungen beschreiben lässt, der aber nicht notwendig in Gesetzes- oder Vertragsform gegossen ist.

Je nachdem, ob durch die Statusbeziehung ein symmetrisches oder ein nicht symmetrisches Verhältnis etabliert wird, können Statusbeziehungen in Graphen als einfache Linien oder als Pfeile dargestellt werden. Beispiele für Statusbeziehungen im Recht sind die familienrechtlichen Status, von denen etwa die Mutterschaft nicht symmetrisch, die Lebenspartnerschaft hingegen symmetrisch ist. Ebenso fallen die nicht symmetrischen Beziehungen

<sup>111</sup> So schon das Vorgehen bei *Moreno*, *Who Shall Survive?*, 1934, S. 23–49.

<sup>112</sup> Hierzu entwickeln *Moldoveanu und Baum* eine ausdifferenzierte *epistemic description language*: *Moldoveanu und Baum*, *Epinets*, 2014, S. 23–33.

Arbeitgeber–Arbeitnehmer, Verein–Vereinsmitglied und Personengesellschaft–Gesellschafter nach dem hiesigen Verständnis unter die Statusbeziehungen.

An Abbildung 2.2 (S. 40) und den begleitenden Erläuterungen wird deutlich, dass sich für die meisten Beziehungen ersten Grades die Darstellung durch Pfeile anbietet. Da ein einziger (nicht animierter) Graph nicht zugleich  $t_0$  und  $t_1$  darstellen kann und für die Substanzen, die sich in Transaktion und Diffusion bewegen bzw. verbreiten, in der Standardrepräsentation eines Netzwerks kein Platz ist, birgt dies die Gefahr, dass sowohl die Rolle der Zeit als auch jene der unterschiedlichen Situationen, aus denen die als Pfeile dargestellten Beziehungen resultieren, bei der Analyse in Vergessenheit geraten.<sup>113</sup>

## 2. Grad.

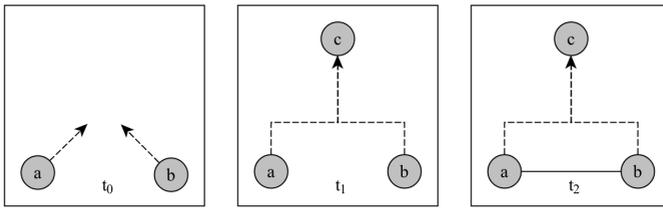
Bei den Beziehungen zweiten Grades lassen sich Co-Produktion, Co-Reaktion, Co-Status und Co-Okkurrenz unterscheiden, wobei im Rahmen der Co-Reaktion zwischen Co-Referenz und bibliographischer Kopplung differenziert werden kann. Damit der Definitionsprozess transparent wird, schließt ein Beziehungsprofil hier jeweils drei Zeitpunkte ein (Abb. 2.3, S. 45): einen Ausgangszeitpunkt  $t_0$ , einen Ereigniszeitpunkt  $t_1$ , der die Grundlage der Beziehungsdefinition bildet, und einen Analysezeitpunkt  $t_2$ , an dem die Beziehung festgestellt wird.

Der Darstellung liegen folgende Erwägungen zugrunde:

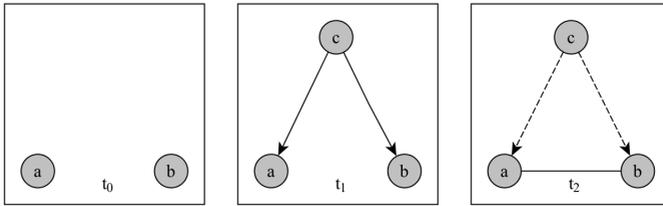
- *Co-Produktion.* Eine Co-Produktionsbeziehung zwischen  $a$  und  $b$  bringt zum Ausdruck, dass diese gemeinsam an einem Produktionsprozess (Beziehung ersten Grades) beteiligt waren oder sind. Sie ist vor allem in Form der Co-Autorenschaft von Rechtstexten interessant: Welche Richter waren an welchen Entscheidungen gemeinsam beteiligt? Welche Parlamentarier haben welchen Gesetzesentwürfen gemeinsam zugestimmt? Welche Rechtswissenschaftler haben welche Publikationen gemeinsam verfasst oder gemeinsam herausgegeben?

Wie die Co-Produktion zu interpretieren ist, hängt unter anderem davon ab, auf welche Rechtstexte sie sich bezieht. Bei Verfassungsrichtern etwa legt schon die Veröffentlichung eines einzelnen gemeinsamen Sondervotums eine übereinstimmende Positionierung hinsichtlich einer Einzelfrage nahe, der ein gewisses Gewicht beigelegt wird. Unter Parlamentariern böte sich eine solche Interpretation allenfalls beim gemeinsamen Verfassen oder Einbringen von Anträgen an, nicht aber bei identischen Stimmabgaben im Gesetzgebungsprozess. Dort wiederum können Abstimmungsmuster, die sich z.B. bei der Betrachtung des Abstimmungsverhaltens über die gesamte Legislaturperiode hinweg zeigen, Aufschluss sowohl über die Fraktionsdisziplin als auch über die (z.B. ideologische) Ähnlichkeit einzelner Parlamentarier geben.

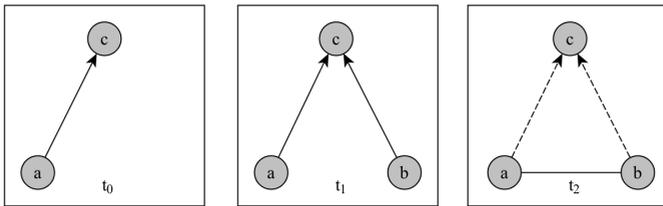
<sup>113</sup> Zur Rolle der Zeit in der juristischen Netzwerkforschung noch unten S. 187–193.



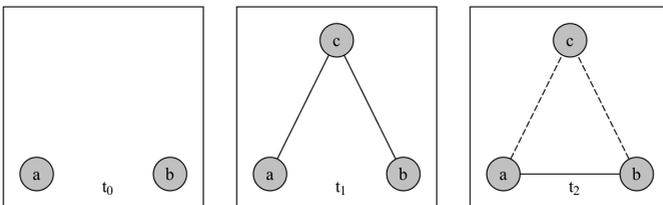
Co-Produktion



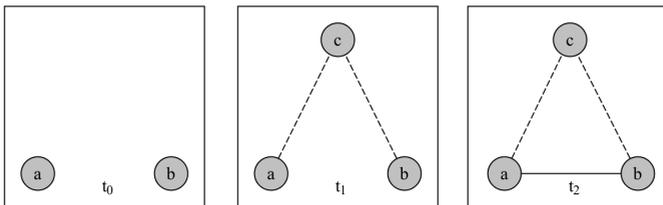
Co-Reaktion Var. 1: Co-Referenz



Co-Reaktion Var. 2: Bibliographische Kopplung



Co-Status



Co-Okkurrenz

Abbildung 2.3: Profile typischer Beziehungen zweiten Grades

Bei Rechtswissenschaftlern kann gemeinsame Autorenschaft ein Zeichen von Kollaboration, aber auch von Subordination sein (z.B. bei gemeinsamen Publikationen von Lehrstuhlinhabern und ihren Mitarbeitern); Ähnliches gilt für die gemeinsame Herausgeberschaft, die allerdings – etwa bei Festschriften – auch ein Indiz für eine gemeinsame Geschichte oder für die Zugehörigkeit zu derselben Schule sein kann.

- *Co-Reaktion*. Die Co-Reaktion bringt zum Ausdruck, dass *a* und *b* in der gleichen Reaktionsbeziehung (Beziehung ersten Grades) zu *c* stehen. Die juristische Netzwerkforschung ist hier vor allem an der Co-Reaktion unter Rechtstexten und an der Co-Reaktion unter Autoren von Rechtstexten interessiert. Für Texte und ihre Autoren werden in der Bibliometrie zwei Beziehungsarten unterschieden, die letztlich zwei Seiten derselben Medaille sind: Co-Referenz (*co-reference* bzw. *co-citation*) und bibliographische Kopplung (*bibliographic coupling*).<sup>114</sup>
- *Co-Referenz*. Zwischen *a* und *b* wird eine Co-Referenz-Beziehung angenommen, wenn es (mindestens) ein *c* gibt, das sowohl *a* als auch *b* referenziert, also *a* und *b* zitiert oder auf *a* und *b* verweist. Wenn beispielsweise ein Aufsatz sowohl „Das Kapital“ von *Karl Marx* als auch „The Wealth of Nations“ von *Adam Smith* zitiert, bedingt dies eine Co-Referenz-Beziehung sowohl zwischen den Werken als auch zwischen ihren Autoren.

Co-Referenz-Beziehungen ist inhärent, dass sie sich mit der Zeit unabhängig von den Beziehungsbeteiligten entwickeln: Immer dann, wenn ein Werk mit einem anderen erstmals gemeinsam referenziert wird, bildet sich eine neue Co-Referenz-Beziehung aus, und immer dann, wenn bereits co-referenzierte Werke wieder gemeinsam referenziert werden, gewinnt eine bestehende Co-Referenz-Beziehung zusätzlich an Gewicht.

Wie genau Co-Referenz-Beziehungen zu interpretieren sind, hängt stark vom Kontext ab; insbesondere können sie thematische Ähnlichkeit oder persönliche Beziehungen signalisieren. In der Rechtswissenschaft können Veränderungen in den Strukturen der Co-Referenz-Beziehungen von Texten und Autoren auf Interpretations- oder Bedeutungsverschiebungen hinweisen, die auch mit der Ausbildung neuer Rechtsbegriffe einhergehen können.

Beispielsweise hängt die Etablierung des Rechts auf ein menschenwürdiges Existenzminimum offensichtlich mit dem Aufbau einer Co-Referenz-Beziehung zwischen Art. 1 I GG (Menschenwürde) und

<sup>114</sup> Historisch geht die bibliographische Kopplung der Co-Referenz voraus: *Kessler*, *American Documentation* 14 (1963), S. 10–25 (bibliographische Kopplung); *Small*, *Journal of the American Society for Information Science* 24 (1973), S. 265–269 (Co-Zitation zwischen wissenschaftlichen Aufsätzen). Auf den möglichen Nutzen beider Konzepte in der juristischen Recherche macht bereits *Ogden*, *Law Library Journal* 85 (1993), S. 1 (47), aufmerksam.

Art. 20 I GG (Sozialstaatsprinzip) zusammen.<sup>115</sup> Man kann die Entwicklung dieser Beziehung über die Zeit (und damit die Rechtsprechung des BVerfG) untersuchen, um herauszufinden, ob ihre Verstärkung der Etablierung des Rechts auf ein menschenwürdiges Existenzminimum nachfolgte oder aber vorausging. Letzteres wiederum könnte bedeuten, dass die Schaffung des neuen Rechtsbegriffs den Endpunkt einer organischen Entwicklung darstellte, während Ersteres darauf hindeuten würde, dass das BVerfG hier einen abrupten Einschnitt in die verfassungsrechtliche Dogmatik vornahm.

- *Bibliographische Kopplung.* Die bibliographische Kopplung ist die Kehrseite der Co-Referenz-Beziehung: *a* und *b* sind bibliographisch gekoppelt, wenn es (mindestens) ein *c* gibt, das sie beide referenzieren. Das heißt, dass alle Aufsätze, die „The Wealth of Nations“ zitieren, erst einmal bibliographisch gekoppelt sind. Je mehr Referenzen die Aufsätze teilen, desto stärker ist dann ihre bibliographische Kopplung.

Im Gegensatz zu Co-Referenz-Beziehungen können sich bibliographische Kopplungsbeziehungen über die Zeit nur mithilfe mindestens eines der Beziehungsbeteiligten entwickeln: Jede Publikation legt ihre bibliographische Kopplung im Verhältnis zu allen älteren Publikationen selbst fest; auf ihre bibliographische Kopplung mit jüngeren Publikationen kann sie allerdings keinen Einfluss nehmen. Uneingeschränkt gilt dies freilich nur, soweit die zugrundeliegenden Zitiernetzwerke sich in ihrer internen Struktur lediglich durch Wachstum (und nicht etwa durch das Hinzufügen oder Entfernen von Beziehungen) verändern können. Das trifft auf Netzwerke aus Gerichtsentscheidungen oder rechtswissenschaftlichen Aufsätzen zu, nicht aber auf Netzwerke aus Gesetzestexten. Hier offenbart sich auch ein wesentlicher Unterschied zwischen dem als Analyseobjekt unter Netzwerkforschern beliebten *World Wide Web* und dem in der juristischen Netzwerkforschung prominentesten Netzwerk aus Entscheidungen des *U.S. Supreme Court*: Während Ersteres sich auch ohne Wachstum strukturell verändern kann (auf bestehenden Webseiten können Links hinzugefügt und entfernt werden), ist Letzteres nur aufgrund von Wachstum dynamisch (neue Entscheidungen werden veröffentlicht).<sup>116</sup>

Die Definition der bibliographischen Kopplung und die im Recht typische Natur ihres Veränderungsprozesses bedingen auch, dass sie vor allem im Maximalbereich (und hinsichtlich der Veränderungen in diesem) interessant ist: Bei rechtswissenschaftlichen Beiträgen oder Autoren, die verhältnismäßig stark miteinander gekoppelt sind, liegt eine gewisse thematische oder gedankliche Verwandtschaft (bzw. in Extremfällen womöglich eine Plagiatsvermutung) nahe, die nach ihrer Aufdeckung weiter untersucht werden kann.

<sup>115</sup> Erstmals als solches bezeichnet wurde das Recht auf ein menschenwürdiges Existenzminimum in BVerfGE 125, 175 (Hartz IV); danach aufgegriffen etwa in BVerfGE 132, 134 (Asylbewerberleistungsgesetz) und BVerfGE 137, 34 (existenzsichernder Regelbedarf).

<sup>116</sup> Näher zur Veränderung juristischer Netzwerke über die Zeit unten S. 187–193.

Bei Gerichtsentscheidungen kann die Analyse bibliographischer Kopplung auch Aufschluss über die Entwicklung von Rechtsprechungslinien geben. So lässt sich die Tatsache, dass eine höchstrichterliche Entscheidung die letzte höchstrichterliche Entscheidung zu einer bestimmten Vorschrift nicht nur zitiert, sondern mit dieser zudem stark bibliographisch gekoppelt ist, als Zeichen von Rechtsprechungskontinuität deuten. Ein geringerer Kopplungsgrad kann demgegenüber sowohl ein Zeichen für Rechtsprechungswandel sein als auch ein Signal dafür, dass sich die Rechtsfragen beider Entscheidungen lediglich teilweise decken.

- *Co-Status*. Eine Co-Status-Beziehung zwischen *a* und *b* zeigt an, dass diese in (mindestens) einer gleichen Statusbeziehung (Beziehung ersten Grades) zu *c* stehen. Wichtigstes Beispiel ist die gleichzeitige Zugehörigkeit zu Entscheidungsgremien oder Institutionen – etwa zu Vorstands- und Aufsichtsratsgremien, Spruchkörpern von Gerichten, parlamentarischen Ausschüssen, Parteien oder juristischen Fakultäten.

Co-Status-Beziehungen sind für die juristische Netzwerkforschung vor allem interessant, wenn sie in ihrem Zusammenspiel mit anderen Beziehungen ersten und zweiten Grades betrachtet werden. Oft stellt sich hier die Frage, inwieweit Co-Status-Beziehungen die Interessen Dritter beeinträchtigen und daher zu unterbinden sind. Klassische Beispiele hierfür sind Personalverflechtungen zwischen Unternehmen sowie zwischen Unternehmen und politischen Gremien.<sup>117</sup>

- *Co-Okkurrenz*. Eine Co-Okkurrenz-Beziehung zwischen *a* und *b* wird angenommen, wenn beide gemeinsam bei *c* auftreten. Die gestrichelten Linien schon in *t*<sub>1</sub> (Abb. 2.3, S. 45) bringen dabei zum Ausdruck, dass der Co-Okkurrenz die für den Co-Status charakteristische Anlage auf Dauer fehlt: Co-Status ist affiliationsbasiert, Co-Okkurrenz ereignisbasiert.

Was „gemeinsames Auftreten“ bedeutet, hängt davon ab, welche Objekte man betrachtet. Bei Rechtssubjekten ist insbesondere die Teilnahme an denselben Veranstaltungen relevant, etwa die Teilnahme zweier Führungskräfte an derselben Sitzung oder die Teilnahme zweier Professoren an derselben Konferenz. Bei Rechtstexten kommen vor allem zwei Arten von Co-Okkurrenz in Betracht: das Erscheinen in demselben Publikationsmedium als Phänomen der Co-Okkurrenz ganzer Texte und das Auftreten in demselben Text oder in einem bestimmten Ausschnitt desselben Texts als Phänomen der Co-Okkurrenz von Ausdrücken. Die Co-Okkurrenz von Ausdrücken kann zur bereits angesprochenen, um Objektivität bemühten Definition von Rechtsbegriffen herangezogen werden.<sup>118</sup> Das gemeinsame Auftreten von Rechtsbegriffen in Rechtstexten lässt sich dann wiederum als Phänomen der Co-Okkurrenz untersuchen.

<sup>117</sup> Zu Personalverflechtungen zwischen Unternehmen bereits oben S. 22–28.

<sup>118</sup> Dazu bereits S. 36–37.

Die Co-Okkurrenz ist eng mit den anderen, bereits erörterten Beziehungsprofilen verknüpft: Bei Rechtssubjekten hängt sie oft vom Co-Status ab, etwa dann, wenn die Mitgliedschaft in einem gewissen Gremium die Teilnahme an bestimmten Sitzungen dieses Gremiums erst ermöglicht. Gleichzeitig können bei mehreren Rechtssubjekten mit Co-Status gerade Co-Okkurrenz-Unterschiede Aufschluss über Koalitionen und Einflusstrukturen innerhalb von Gruppen geben.<sup>119</sup>

Bei Rechtstexten gehen Co-Produktion und Co-Okkurrenz von Texten häufig miteinander einher, ebenso Co-Okkurrenz von Texten und Co-Okkurrenz von Rechtsbegriffen sowie Co-Okkurrenz von Rechtsbegriffen und Co-Reaktion von Texten (sowohl in Form von Co-Zitation als auch in Form bibliographischer Kopplung). Die Korrelationsstärke der Beziehungen kann aber auch hier stark zwischen Objekten mit gleichen Beziehungsprofilen variieren – und die Unterschiede können auf inhaltliche und strukturelle Nuancen im Recht hindeuten. Für die juristische Netzwerkforschung kann es sich also lohnen, Co-Okkurrenz getrennt von den mit ihr im Einzelfall korrelierenden Beziehungsprofilen zu untersuchen.

Bei näherer Betrachtung von Abbildung 2.3 (S. 45) fällt auf, dass die Beziehungen zweiten Grades im Gegensatz zu jenen ersten Grades in Graphen allesamt durch einfache Linien dargestellt werden. Das liegt daran, dass sie Gleichheit mit Blick auf ein Vergleichskriterium zum Ausdruck bringen – und die Gleichheitsbeziehung ist *symmetrisch*.<sup>120</sup>

<sup>119</sup> Beispiel:  $a$ ,  $b$  und  $p$  sind alle Mitglieder des Aufsichtsrats von  $c$ , wobei  $a$  und  $b$  in der Vergangenheit an allen Gremiensitzungen teilnahmen,  $p$  sich aber nie blicken ließ. Dann haben  $a$ ,  $b$  und  $p$  mit Blick auf  $c$  jeweils identischen Co-Status, ihre Co-Okkurrenz in den Gremiensitzungen legt allerdings nahe, dass  $a$  und  $b$  im Aufsichtsrat anders positioniert sind als  $p$  (wie genau, hängt vom Kontext ab).

<sup>120</sup> Eine Beziehung ist im mathematischen Sinne eine *Relation*, d.h. eine Teilmenge des *Kartesischen Produkts* zweier Mengen. Die Menge aller Beziehungen zwischen Objekten in einem Netzwerk lässt sich als Relation  $R \subseteq V \times V$  auffassen, wobei  $V$  die Menge der Punkte des Graphen ist, durch den das Netzwerk repräsentiert wird.  $R$  ist *symmetrisch*, wenn aus der Tatsache, dass  $a$  mit  $b$  in Beziehung steht, für beliebig gewählte  $a$  und  $b$  folgt, dass  $b$  auch mit  $a$  in Beziehung steht (d.h.  $aRb \Rightarrow bRa$ ). Die Beziehung „ist verheiratet mit“ ist also beispielsweise symmetrisch. Demgegenüber ist die Beziehung „zitiert“ für rechtswissenschaftliche Aufsätze (im Ausgangspunkt) *antisymmetrisch* (d.h.  $\forall a, b \in V$  mit  $a \neq b$  gilt:  $aRb \Rightarrow \neg bRa$ ): Wenn Aufsatz  $a$  den Aufsatz  $b$  zitiert, kann (aufgrund der Publikationsreihenfolge „erst  $b$ , dann  $a$ “) nicht auch Aufsatz  $b$  den Aufsatz  $a$  zitieren. Abweichungen von der Antisymmetrie (zwei Aufsätze zitieren sich gegenseitig) deuten hier darauf hin, dass zwischen den Verfassern der Aufsätze enger Austausch (bis hin zur Personalunion) bestand.

### 3. Grad.

Beziehungen dritten Grades bringen in erster Linie Ähnlichkeit zum Ausdruck. Das Profil dieser Beziehungen wurde in Abbildung 2.1 (S. 38) bereits dargestellt. Ähnlichkeitsbeziehungen sind symmetrisch (genau wie die oben angesprochenen Gleichheitsbeziehungen, die eine Teilmenge der Ähnlichkeitsbeziehungen sind) und werden in Graphen als Linien repräsentiert. Auf die Schwierigkeiten, Beziehungen dritten Grades sauber zu definieren und zu interpretieren, wurde bereits hingewiesen.<sup>121</sup> Das hier typische Definitionsvorgehen sei noch einmal anhand eines Beispiels illustriert:<sup>122</sup>

Gegeben seien zwei Länder,  $a$  und  $b$ , sowie ein Index  $x$ , etwa der *Rule of Law Index*.<sup>123</sup> Angenommen,  $a$  erreiche auf dem Index den Wert  $c$  und  $b$  erreiche den Wert  $d$ . Die Differenz zwischen  $c$  und  $d$  liege unterhalb eines bestimmten, durch den Forscher festgelegten Schwellenwerts. Dann definiert der Forscher eine Beziehung zwischen  $a$  und  $b$ , die als einfache Linie dargestellt wird. Das Gewicht der Linie, die  $a$  und  $b$  verbindet, ist umso größer, je kleiner die Differenz zwischen  $c$  und  $d$  ist.<sup>124</sup>

Das ist unnötig kompliziert. Zur Untersuchung von Ähnlichkeitsbeziehungen höheren Grades stehen im Bereich des *Data Mining* Methoden zur Verfügung, die mehr Informationen mit weniger Vorurteilen besser verarbeiten können als die „manuelle“ Definition von Netzwerken aus Beziehungen dritten Grades.<sup>125</sup> Diese Methoden teilen bisweilen nicht nur das Erkenntnisinteresse, sondern auch den Namen mit ihren Verwandten aus der Netzwerkforschung. Sie basieren aber gerade nicht auf „Beziehungen“ bzw. „Linien“ im hiesigen Sinne, sondern auf Attributen von Objekten, welche durch die Methoden selbst in einem metrischen Raum verteilt und erst dadurch in Distanzbeziehungen relativ zueinander gebracht werden. Die juristische Netzwerkforschung sollte die Untersuchung von Beziehungen dritten und höheren Grades daher der (noch nicht wirklich existenten) *Legal Data Science* überlassen. Tabelle 2.1 (S. 51) fasst die eingeführten Differenzierungen zusammen.

<sup>121</sup> Dazu oben bei Abb. 2.1 (S. 38).

<sup>122</sup> Ähnlich das Vorgehen bei *Siems*, *Journal of Institutional Economics* 12 (2016), S. 579–602.

<sup>123</sup> Abrufbar unter [worldjusticeproject.org/rule-of-law-index](http://worldjusticeproject.org/rule-of-law-index).

<sup>124</sup> Zum Gewicht von Linien noch unten S. 51–55.

<sup>125</sup> Zur Messung von Ähnlichkeiten beim *Data Mining* etwa *Aggarwal*, *Data Mining*, 2015, S. 63–89, mit knapper Diskussion graphbasierter Methoden auf S. 85–87.

<i>Profil</i>	<i>Grad</i>	<i>Darstellung</i>
Produktion	1	Pfeile
Transaktion	1	Pfeile
Diffusion	1	Pfeile
Reaktion	1	Pfeile
Status	1	Pfeile oder Linien
Co-Produktion	2	Linien
Co-Reaktion	2	Linien
Co-Status	2	Linien
Co-Okkurrenz	2	Linien
Ähnlichkeit	3	Linien

Tabelle 2.1: Zusammenfassung der Beziehungsprofile

### 3. Beziehungsform

Neben ihrem Grad und ihrem Profil lassen sich Beziehungen auch nach ihrer Form unterscheiden. Unter der Form einer Beziehung werden hier alle Eigenschaften einer Beziehung zusammengefasst, für die es in der Graphentheorie besonders bezeichnete Gegenstücke gibt. Es geht also um Informationen über Beziehungen, welche die Übersetzung eines Netzwerks in einen Graphen unmittelbar beeinflussen. Eine Information, auf die dies zutrifft, wurde im vorangehenden Abschnitt bereits verwendet: die Richtung einer Beziehung. Daneben lässt sich eine Beziehung auch anhand ihrer Multiplizität, Intensität und Bewertung sowie nach der Anzahl ihrer Beteiligten charakterisieren. Im Einzelnen:

- *Richtung*. Eine Beziehung hat eine Richtung, wenn man sinnvollerweise davon sprechen kann, dass sie *von* einem Objekt *zu* einem anderen verläuft. Gerichtete Beziehungen werden in Graphen als *Pfeile* dargestellt. Wie bereits festgestellt, bietet sich für die meisten Beziehungen ersten Grades eine Repräsentation als *Pfeile* an, während für Beziehungen zweiten und dritten Grades Linien die natürliche Darstellungsform sind.<sup>126</sup> Der Grund ist (wie bereits erläutert), dass viele Beziehungen ersten Grades Aktionen bzw. Interaktionen repräsentieren, die *nicht symmetrisch* sind, während Beziehungen zweiten und dritten Grades, die Gleichheit oder Ähnlichkeit zum Ausdruck bringen, *symmetrisch* sind.

Bei nicht symmetrischen Beziehungen stellt sich die Frage, in welche Richtung die Pfeile zeigen sollten, durch die sie dargestellt werden. Zwar haben sich für die verschiedenen Beziehungsprofile in den meisten Fällen gewisse Standards herausgebildet, die auch der voranstehenden Darstellung zugrunde gelegt wurden. Letztlich hängt die Richtung aber von der eingenommenen Perspektive ab, die sich am theoretischen Unterbau und am Erkenntnisinteresse der Untersuchung im Einzelfall orientieren muss. Dies zeigt sich deutlich an den bereits angesprochenen Differenzen in der Darstellung von Zitaten: Wer an der

<sup>126</sup> Dazu oben bei Tab. 2.1 (S. 51).

Diffusion von Wissen interessiert ist und annimmt, dass der Diffusionsprozess sich in Zitaten wissenschaftlicher Publikationen untereinander niederschlägt, kann die Pfeile vom Zitierten auf das Zitierende richten. Demgegenüber sollte derjenige, der sich für Zitate vor allem als soziales Phänomen interessiert, die Pfeile vom Zitierenden auf das Zitierte zeigen lassen.

Eine Menge nicht symmetrischer Beziehungen lässt sich immer auch als Menge symmetrischer Beziehungen begreifen, die als einfache Linien dargestellt werden können. Dazu nimmt man eine symmetrische Beziehung immer dann an, wenn zwischen zwei Objekten eine nicht symmetrische Beziehung besteht.<sup>127</sup> Das kann die Analyse mitunter erleichtern, geht aber mit Informationsverlust einher: Man analysiert nur noch das Vorhandensein nicht symmetrischer Beziehungen *per se*, nicht aber die Ausrichtung der Beziehungen und die daraus möglicherweise ersichtlichen Ungleichgewichte.

- *Multiplizität*. Multiplizitäten sind für Beziehungen relevant, für die man sinnvollerweise davon sprechen kann, dass sie zwischen denselben Objekten *mehrfach* vorkommen. Dies trifft unter anderem auf Zitate zwischen Gerichtsentscheidungen zu: Eine Gerichtsentscheidung zitiert andere Gerichtsentscheidungen, diese aber unterschiedlich häufig. Um nicht alle Zitate gleich behandeln zu müssen, kann man diese Häufigkeiten erfassen und für jedes Zitat eine eigene Beziehung annehmen, die durch einen entsprechenden Pfeil (oder – im Falle einer ungerichteten Beziehung – durch eine entsprechende Linie) im Graphen repräsentiert wird. Die resultierenden Linien werden *Mehrfachkanten* oder auch *Multikanten* (*multiedges* bzw. *multi-edges*) genannt.<sup>128</sup>
- *Intensität*. Durch Intensitäten lassen sich Beziehungen erfassen, bei denen man sinnvollerweise davon sprechen kann, dass sie *mehr oder weniger stark* sind. Dies drängt sich in manchen Kontexten auf, etwa dann, wenn ein Netzwerk aus Individuen und Telefongesprächen betrachtet wird und die Gespräche unterschiedlich lange dauern, oder wenn ein Netzwerk aus Bahnhöfen und Schienen betrachtet wird, in dem die Gleisabschnitte zwischen den Bahnhöfen unterschiedlich lang sind. Beziehungsintensitäten werden im Graphen als *gewichtete Kanten* (*weighted edges*), also als Linien mit Gewichtsinformation dargestellt.<sup>129</sup>

Netzwerke, deren Beziehungen Multiplizitäten zugeordnet sind, lassen sich in Netzwerke mit Beziehungsintensitäten übersetzen (beispielsweise können Mehrfachzitate in Zitationsbeziehungen unterschiedlicher Intensität umgewan-

<sup>127</sup> Praktisch bedeutet das, dass man die Informationen über die Richtung der Beziehungen ignoriert.

<sup>128</sup> Brandes, Graphentheorie, in: Stegbauer und Häußling (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 345 (347); Newman, Networks, 2010, S. 110.

<sup>129</sup> Newman, Networks, 2010, S. 113. In der visuellen Repräsentation eines Graphen werden Gewichtsinformationen typischerweise in Linien unterschiedlicher Stärke übersetzt, siehe etwa Pfeffer, Visualisierung sozialer Netzwerke, in: Stegbauer (Hrsg.), Netzwerkanalyse und Netzwerktheorie, 2010, S. 227 (228).

delt werden). Das macht nicht nur die graphische Darstellung übersichtlicher, sondern kann auch aus technischen Gründen geboten sein, um mit dem korrespondierenden Graphen besser rechnen zu können. Demgegenüber ist das umgekehrte Vorgehen nicht zwangsläufig möglich, weil nicht alle Intensitäten (z.B. Zeit- und Distanzangaben) als Sammlungen einzelner Beziehungen aufgefasst werden können.

Es kann sich auch lohnen, Multiplizitäten und Intensitäten gleichzeitig zuzulassen: Will man beispielsweise Zahlungsströme als Transaktionsbeziehungen zwischen Personen oder Konten nachvollziehen, um auf kriminelle Aktivitäten (z.B. Geldwäsche oder Untreue) hindeutende Unregelmäßigkeiten zu erkennen, so sind dafür sowohl die Anzahl der Transaktionen als auch die transferierten Summen von Interesse – Erstere lassen sich in Form von Mehrfachkanten, Letztere durch Kantengewichte abbilden.

- *Bewertung.* Bewertungen dienen der Repräsentation von Beziehungen, bei denen man sinnvollerweise davon sprechen kann, dass sie (mehr oder weniger) *positiv oder negativ* sind. Beispielsweise können Menschen ihr Verhältnis zu anderen Menschen in einer Befragung als (mehr oder weniger) gut oder schlecht einschätzen; Rezensionen können sowohl positiv als auch negativ ausfallen; Banktransaktionen können sowohl eine positive als auch eine negative Veränderung des Kontostandes bewirken. Bewertungen werden in Graphen als Kanten mit (positivem oder negativem) Vorzeichen (*signed edges*) dargestellt, die sich von gewichteten Kanten nur dadurch unterscheiden, dass nun auch negative Gewichte zugelassen sind.<sup>130</sup>
- *Anzahl der Beteiligten.* In den bisherigen Ausführungen wurde davon ausgegangen, dass Beziehungen zwischen *genau zwei* Objekten bestehen. In der Realität existieren jedoch Beziehungen, die sich nur als Beziehung zwischen *mehr als zwei* Beteiligten unverfälscht repräsentieren lassen; außerdem sind auch Beziehungen eines Objekts zu sich selbst (Selbstbeziehungen, beispielsweise Selbstzitate) unter Umständen interessant. Selbstbeziehungen lassen sich ohne große Probleme erfassen. Ihre Entsprechungen in Graphen werden *Schleifen* (*loops*) genannt und visuell ebenso repräsentiert.<sup>131</sup>

Beziehungen mit mehr als zwei Beteiligten sind schwieriger darzustellen. Für das Recht sind diese Beziehungen unter anderem relevant, wenn es darum geht, die Wechselwirkungen zwischen Institutionen, von ihnen getroffenen Entscheidungen und mit ihnen affilierten Individuen zu erfassen. Bereits wenn mehrere Personen in demselben Gremium sitzen, besteht zwischen ihnen eine (über das Gremium vermittelte) multilaterale Beziehung, die sich nicht ohne Informationsverlust in bilaterale Beziehungen zwischen den einzelnen

<sup>130</sup> Wasserman und Faust, Social Network Analysis, 1994, S. 137.

<sup>131</sup> Harary, Graph Theory, 1969, S. 10; Brandes, Graphentheorie, in: Stegbauer und Häußling (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 345 (347).

Gremienmitgliedern übersetzen lässt.<sup>132</sup> Auch die Zusammenhänge zwischen Rechtsbegriffen, den sie verwendenden Gerichtsentscheidungen und den mit ihnen assoziierten Rechtsbereichen lassen sich nur als Beziehungen zwischen mehreren Beteiligten präzise beschreiben.

Für solche multilateralen Beziehungen – seien sie nun ersten, zweiten oder dritten Grades – sind *Hyperkanten (hyperedges)* in Graphen die naheliegende Darstellungsform; Beziehungen können dann nicht nur zwischen genau zwei Objekten, sondern zwischen beliebig vielen Objekten definiert werden.<sup>133</sup> Problematisch ist, dass Hyperkanten nicht mit den gängigen Speicherformen für Graphen kompatibel sind, auf denen auch die Instrumente der Netzwerkforschung aufbauen. Man behilft sich, indem man für jede multilaterale Beziehung in der Netzwerkdefinition ein Objekt einführt und die Beziehungsbeteiligten mit diesem Objekt in eine bilaterale Beziehung setzt, womit man ein bipartites Netzwerk (und daraus einen bipartiten Graphen) erhält (Abb. 2.4, S. 54).

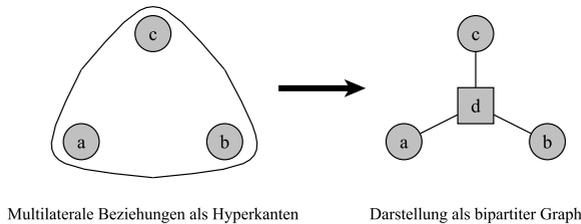


Abbildung 2.4: Übersetzung multilateraler Beziehungen in bilaterale Beziehungen

Die vorstehenden Ausführungen zeigen, dass sich Beziehungsformen in Graphen sehr differenziert darstellen lassen. Beziehungen, die nicht besonders ausgezeichnet sind, repräsentiert man durch einfache Linien, die implizit ein positives und konstantes Gewicht von 1 haben. In vielen Fällen wird man über die behandelten Charakteristika hinaus noch weitere Informationen zu den Beziehungen erfassen wollen, für die es keine unmittelbare graphentheoretische Übersetzung gibt. Das ist technisch ohne Weiteres möglich – welche Informationen betroffen sind, hängt aber stark vom Einzelfall ab. Ein wichtiges Beispiel für nützliche Zusatzinformationen ist der Zeitpunkt einer Beziehung. Dieser ist insbesondere für Netzwerke relevant, die sich über die Zeit verändern – etwa Zitiernetzwerke oder Netzwerke aus sozialen Interaktionen.

<sup>132</sup> Genau dies ist allerdings das Vorgehen bei der in Fn. 51 (S. 21) bereits angesprochenen Projektion (*one-mode projection*). Bei  $n$  Personen werden so aus  $n$  Beziehungen (jeweils zwischen Person und Gremium)  $\binom{n}{2} = \frac{n!}{(n-2)!2!}$  Beziehungen (jeweils zwischen Person und Person).

<sup>133</sup> Dazu *Bonacich, Holdren und Johnston*, *Social Networks* 26 (2004), S. 189 (194). Graphen mit Hyperkanten heißen Hypergraphen (*hypergraphs*); sie sind eine Generalisierung von Graphen. Grundlegend zu Hypergraphen *Berge, Graphes et Hypergraphes*, 1970, S. 373–474, auf den auch die Terminologie zurückgeht.

Abbildung 2.5 (S. 55) fasst die eingeführten Unterscheidungen zusammen. Mit dieser Systematik und den für die Objekte eingeführten Differenzierungen lassen sich nun die Probleme verstehen, die auftreten, wenn man Objekte und Beziehungen zum Untersuchungsgegenstand „Netzwerk“ zusammenfügt.<sup>134</sup>

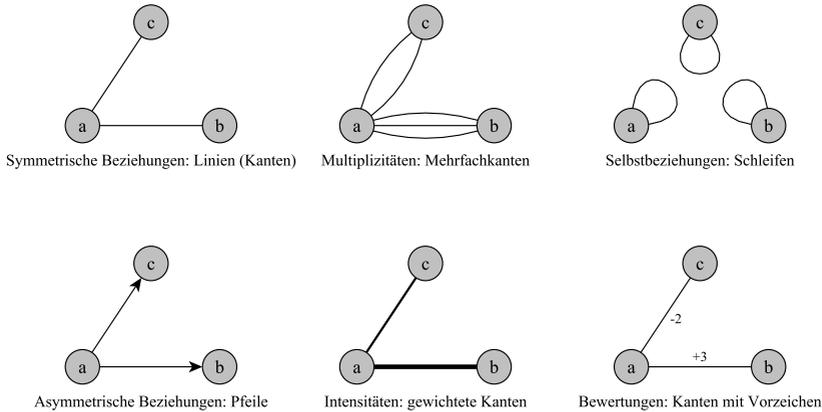


Abbildung 2.5: Beziehungsformen und ihre Darstellung in Graphen<sup>135</sup>

### III. Probleme

Bisher wurden Objekte und Beziehungen getrennt voneinander behandelt. Dabei ist deutlich geworden, dass der erste Schritt der juristischen Netzwerkforschung, die Netzwerkdefinition, schon dann zahlreiche Schwierigkeiten und Differenzierungsnotwendigkeiten mit sich bringt, wenn man Objekte und Beziehungen jeweils separat betrachtet. In der Praxis müssen Objekte und Beziehungen als gemeinsame Bestandteile eines Netzwerks jedoch regelmäßig gemeinsam gehandhabt werden; Objekt- und Beziehungsdefinition müssen aufeinander abgestimmt sein. Bei der Netzwerkdefinition sind dann grundsätzlich zwei Herangehensweisen denkbar, von denen eine als *objektgeleitet* und die andere als *beziehungsgeleitet*

<sup>134</sup> Zu den Differenzierungen bei der Objektdefinition oben S. 28–37.

<sup>135</sup> Graphen mit Pfeilen werden als gerichtete Graphen bzw. Digraphen (*directed graphs* bzw. *digraphs*), solche mit gewichteten Kanten als gewichtete Graphen (*weighted* bzw. *valued graphs*), solche mit Mehrfachkanten als Multigraphen (*multigraphs*) und solche, deren Kanten Vorzeichen haben, als signierte Graphen (*signed graphs*) bezeichnet; Graphen, die nur einfache Linien und keine Schleifen aufweisen, heißen schlichte Graphen (*simple graphs*). Zu Teilen der Terminologie auf Deutsch *Brandes*, Graphentheorie, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 345 (346–347). Eine der hiesigen ähnliche Darstellung findet sich bei *Barabási*, *Network Science*, 2016, S. 65. Wo im Folgenden ohne Zusatz von (dem oder den) Graphen gesprochen wird, wird aus dem Kontext deutlich, welche Graphtypen einbezogen sind.

tet gekennzeichnet werden kann. Die objektgeleitete Herangehensweise beginnt mit der Definition der Objektmenge und fragt dann nach der Beziehungsmenge, während die beziehungsgeleitete Herangehensweise mit der Definition der Beziehungsmenge beginnt und dann nach der Objektmenge fragt.

Ein Großteil der Netzwerkforschung ist primär objektgeleitet, was zumindest auch damit zusammenhängen wird, dass Objekte als Fokus einer Betrachtung für die meisten Menschen leichter zu erfassen sind als Beziehungen. Welche Herangehensweise man als Ausgangspunkt wählt, hängt aber letztlich vom Schwerpunkt des Erkenntnisinteresses ab: Möchte man vor allem die Strukturen eines bestimmten Bereichs, beispielsweise des Zivilrechts, verstehen, so wird man bei den Objekten beginnen, die sich diesem Bereich zuordnen lassen. Interessiert man sich hingegen für die Eigenheiten eines bestimmten Prozesses, etwa des Bargeldumlaufs, so wird man bei den Beziehungen beginnen, die den Prozess konstituieren.

Unabhängig davon, welche Herangehensweise gewählt wird, begegnet man im Definitionsprozess in der Regel drei Problemen: dem Selektionsproblem (1.), dem Inklusionsproblem (2.) und dem Rekursionsproblem (3.). Diese Probleme werden nachfolgend diskutiert und anhand des Untersuchungsinteresses „Strukturen des deutschen Gerichtssystems“ illustriert. Sie sind unterschiedliche Facetten des übergreifenden Netzwerkdefinitionsproblems, das diesen Abschnitt durchzieht, treten aber in der Regel in der Reihenfolge auf, in der sie nachfolgend behandelt werden. Bei ihrer näheren Betrachtung wird auch deutlich, dass eine dem Untersuchungsinteresse angemessene Netzwerkdefinition in der Regel eine Kombination von objekt- und beziehungsgeleiteter Herangehensweise erfordert.

### 1. Selektion

Das Selektionsproblem betrifft die Frage: Welche Art von Objekten und Beziehungen soll untersucht werden? Hier geht es darum, zunächst allgemein zwei Mengen potenzieller Netzwerkbestandteile zu bezeichnen, ohne dass es schon darauf ankäme, welche konkreten Elemente diese Mengen enthalten. Die Antwort auf die Selektionsfrage ist nicht immer so leicht, wie es auf den ersten Blick scheinen mag. Oft hat man zwar eine gewisse Vorstellung davon, welche Zusammenhänge analysiert werden sollen. Diese Vorstellung erweist sich bei näherem Hinsehen aber typischerweise als konkretisierungsbedürftig.

Interessiert sich der Forscher für die Strukturen des deutschen Gerichtssystems, so liegt in der soeben eingeführten Terminologie eine objektgeleitete Herangehensweise an die Netzwerkdefinition nahe. Dann stellt sich im ersten Schritt die Frage danach, welche Einheiten des deutschen Gerichtssystems zu Objekten des untersuchten Netzwerks gemacht werden sollen: Gerichtszweige, Gerichte, Spruchkörper, Richter oder Gerichtsentscheidungen? Oder aber Gerichte *und* Gerichtsentscheidungen – oder Gerichtsentscheidungen *und* Richter? Hier werden Unterscheidungen relevant, die in Abschnitt B.I. eingeführt wurden (S. 28–37). Letztlich geht es um die Wahl der *Auflösung* bzw. *Granularität*, mit welcher

der Forscher den Problembereich betrachten will. Die Festlegung der Auflösung schränkt die im zweiten Schritt erfolgende Wahl der Beziehungsart ein: Beziehungen zwischen Gerichten sind (jedenfalls zum Teil) gänzlich anderer Natur als jene zwischen Richtern oder jene zwischen Gerichten *und* Richtern – hier tauchen dann Unterscheidungen wieder auf, die in Abschnitt B.II. vorgestellt wurden (S. 37–55). Möglicherweise zeigt sich im zweiten Schritt, dass gerade die Beziehungsart, deren Untersuchung mit Blick auf das Erkenntnisinteresse besonders vielversprechend erscheint, nicht mit der gewählten Objektart zusammenpasst – dann muss der Forscher zum ersten Schritt zurückkehren, um die dortige Wahl entsprechend zu modifizieren.

Im Falle einer beziehungsgeleiteten Herangehensweise an die Netzwerkdefinition, die sich etwa anbietet, wenn man die Ausbreitung eines Rechtskonzepts in der Rechtsprechung als Diffusionsprozess untersuchen will, verläuft das Vorgehen genau umgekehrt: Hier stellt sich im ersten Schritt die Frage, welche Beziehungsart (mit welcher Art von Manifestation) untersucht werden soll – und die dortige Entscheidung bestimmt darüber, welche Objektmenge zur Analyse mit welcher Auflösung in Betracht kommt.

Kern des Selektionsproblems ist also die abstrakte Auswahl der Basiseinheiten des Netzwerks, ohne dass die Netzwerkelemente im Einzelnen bezeichnet werden. Die Identifikation der konkreten Netzwerkelemente fällt unter anderem in den Bereich des Inklusionsproblems (hierzu sogleich).

## 2. Inklusion

Durch die Abarbeitung des Selektionsproblems werden Objektmenge und Beziehungsmenge so konkretisiert, dass man die Frage stellen kann: Welche konkreten Elemente sollen als Teil dieser Mengen betrachtet werden? Die Erläuterungen in Abschnitt B.I. (S. 28–37) lassen bereits erahnen, dass die Beantwortung dieser Frage alles andere als trivial ist – der Teufel steckt im Detail. Welche Schwierigkeiten im Einzelnen auftreten, hängt davon ab, mit welcher Strategie man sich dem Inklusionsproblem annähert:

Man kann *deduktiv* vorgehen und für die abstrakten Ausdrücke, die man im Rahmen des Selektionsproblems als Mengenbezeichner festgelegt hat (z.B. „Gerichtsentscheidungen“), schrittweise Vertreter in der Realität suchen. Hierfür wird man die ausgewählten Ausdrücke gleich zu Beginn definieren und auslegen müssen. Zu Objekten des untersuchten Netzwerks werden dann alle beobachteten Einheiten, welche die Merkmale aufweisen, zu denen man im Wege der Deduktion gelangt ist. Umgekehrt kann man aber auch *induktiv* vorgehen und beobachtete Einheiten in der Realität schrittweise unter die gewählten Mengenbezeichner fassen. Dieses Vorgehen bietet sich an, wenn man bereits einen Datensatz hat, aus dem man Datenpunkte für die eigene Analyse auswählen kann. Es birgt aber die Gefahr, dass man mit Mengen arbeitet, die mit Blick auf die Untersuchungsfrage unvollständig sind, weil man auf eine genaue Definition und Auslegung der entscheidenden Begriffe gerade verzichtet und daher nicht die Frage stellt,

ob wirklich *alle* Einheiten erfasst werden, deren Betrachtung als Objekte für die Analysezwecke sinnvoll ist. Praktisch bietet es sich daher an, mehrfach zwischen induktiver und deduktiver Vorgehensweise zu wechseln und in diesem Wechselprozess Objekt- und Beziehungsmengen graduell zu konkretisieren.

Wieder am Beispiel des Untersuchungsinteresses „Strukturen des deutschen Gerichtssystems“: Angenommen, der Forscher hat sich im Rahmen des Selektionsproblems für Gerichtsentscheidungen und Zitate als Objekte und Beziehungen entschieden, sodass er damit beginnen kann, die einzelnen Elemente seines Netzwerks zu benennen. Er kann dabei nun etwa induktiv vorgehen und seine Objektmenge als die Menge der Gerichtsentscheidungen definieren, die er in einer bestimmten Datenbank findet. Das kann allerdings insbesondere in Deutschland zu einer höchst unvollständigen und noch dazu unsystematischen Definition der Objektmenge führen, da die Datenbanken oft viele, aber nicht alle Entscheidungen der höchsten Gerichte, einige Entscheidungen des EuGH und „wichtige“ Entscheidungen der Oberlandesgerichte enthalten, ohne dass die Auswahlkriterien vollständig nachvollziehbar sind.

Nun kann es gerade in den Anfängen der juristischen Netzwerkforschung erforderlich sein, mit solchen datensatzabhängigen Mengen, „definitionen“ zu arbeiten.<sup>136</sup> Wichtig ist, dass man sich die damit einhergehenden Limitationen der Netzwerkstudie immer wieder vor Augen führt und gegenüber deren potenziellen Rezipienten offenlegt, zu welchen *Exklusionen* die Handhabung des Inklusionsproblems geführt hat: Welche Gerichtsentscheidungen, deren Berücksichtigung ein außenstehender Beobachter stillschweigend annehmen könnte, werden gerade *nicht* in die Untersuchung einbezogen?

Wer im Umgang mit dem Inklusionsproblem nicht induktiv, sondern deduktiv verfährt, begegnet der ausschließenden Seite des Problems schon an früherer Stelle im Definitionsprozess. Im obigen Beispiel hat der Forscher als Objektmenge „Gerichtsentscheidungen“ definiert. Doch was ist überhaupt eine Gerichtsentscheidung? Soll der Forscher hierunter nur Urteile fassen oder auch Beschlüsse, insbesondere unbegründete Beschlüsse? Welche Gerichte oder Spruchkörper sollen als Autoren der Entscheidungen zur weiteren Eingrenzung herangezogen werden? Wenn es dem Forscher um die Analyse von Strukturen des deutschen Gerichtssystems geht, wird er wahrscheinlich die Entscheidungen von Gerichten ausschließen wollen, die nicht deutsch sind. Doch wie verhält es sich mit den Entscheidungen von Gerichten in Deutschland, die nicht staatlich sind? Und wann ist ein Gericht überhaupt „nicht deutsch“ oder „nicht staatlich“? Wieder werden

---

<sup>136</sup> Ein Beispiel für eine derartige Festlegung der Objektmenge findet sich bei Šadl und Olsen, *Leiden Journal of International Law* 30 (2017), S. 327 (336): „To examine how the CJEU uses effectiveness in its case law, we constructed a citation network of all cases that contain the word *effectiveness* (1,707 in total) from the network of all cases of the CJEU (approximately 10,000 judgments in total). We extracted all the information from Eurlex, which is a publicly available database of all legislation and case law of the EU.“ Hier wird „alle Entscheidungen“ definiert als „alle Entscheidungen, die über Eurlex verfügbar sind“.

die Unterscheidungen aus Abschnitt B.I. relevant (S. 28–37); nun allerdings nicht mehr im Rahmen abstrakter Definitionen, sondern bei deren Konkretisierung.

Das Zusammenspiel der Definitionen von Objekten und Beziehungen wird dann besonders problematisch, wenn man nicht (wie im voranstehenden Beispiel) Rechtstexte, sondern Rechtssubjekte und unter diesen wiederum nicht Individuen, sondern Kollektive als Objekte betrachten möchte: Sieht man von Statusbeziehungen ab, die nach ihrer Etablierung auch ohne die Handlungen Einzelner bestehen können, so sind alle Beziehungen von Kollektiven letztlich abhängig von den Aktionen einzelner Individuen. Damit stellt sich aber die Frage, auf Basis *welcher* Handlungen *welcher* Individuen die Beziehungsmenge zu konstruieren ist. Aus juristischer Perspektive liegt es nahe, diese Frage im Sinne juristischer Handlungszurechnung zu beantworten. Das wird bei Kollektiven ohne rechtliche Verfassung allerdings schwierig und kann auch bei rechtlich verfassten Kollektiven zu einer für das Untersuchungsziel ungeeigneten Definition der Beziehungsmenge führen – etwa dann, wenn rechtlich grenzwertige Aktivitäten von Unternehmen so organisiert sind, dass nach klassischen juristischen Maßstäben gerade *keine* Zurechnung erfolgen kann. Wie mit dieser Schwierigkeit bei der Inklusionsentscheidung umzugehen ist, muss im Einzelfall mit Blick auf das Ziel der geplanten Untersuchung entschieden werden – wichtig ist auch hier zunächst nur das Problembewusstsein.

### 3. Rekursion

Die bisherigen Ausführungen suggerieren, dass die Definition von Objektmenge und Beziehungsmenge unabhängig von den Elementen ist, welche die Definition erfüllen. Das *kann* so sein, muss es aber nicht: Vielmehr stellt sich als drittes, übergreifendes Problem der Netzwerkdefinition gerade die Frage, ob und wie die Netzwerkdefinition von den Elementen abhängen *soll*, die durch die Definition selbst bereits Teil des betrachteten Netzwerks geworden sind. Warum das Problem hier als Rekursionsproblem bezeichnet wird, klärt sich auf, wenn man das Beispiel zum Untersuchungsinteresse „Strukturen des deutschen Gerichtssystems“ noch einen Schritt weiter durchdenkt:

Angenommen, der Forscher hat sich im Rahmen der Selektion für die Untersuchung von Gerichtsentscheidungen und Zitaten als Objekte und Beziehungen entschieden. Außerdem hat er bei der Bearbeitung des Inklusionsproblems festgelegt, dass nur diejenigen Zitate zu Gerichtsentscheidungen untersucht werden sollen, welche in den Entscheidungen des BGH auftauchen, die in einem bestimmten Zeitraum getroffen wurden. Selbst wenn damit offenbar klar ist, worauf es dem Forscher ankommt, steht er doch vor der Frage, wie *genau* er die Grenzen seiner Netzwerkdefinition ziehen sollte: Will er Zitate nur dann als Beziehungen erfassen, wenn sie von einer BGH-Entscheidung zu einer anderen BGH-Entscheidung verlaufen, oder sollen auch Zitate von BGH-Entscheidungen zu Entscheidungen anderer Gerichte als Beziehungen erfasst werden? In letzterem Fall liegt es für den Forscher nahe, auch die zitierten, außerhalb der ursprünglichen Menge der BGH-Entscheidungen liegenden Entscheidungen als Objekte ins Netzwerk auf-

zunehmen und die ursprüngliche Definition damit *rekursiv* zu erweitern. Diesen Prozess kann er dann mit den hinzugekommenen Entscheidungen wiederholen – etwa, um Informationen über die allgemeine Relevanz der vom BGH zitierten Nicht-BGH-Entscheidungen in die Analyse einfließen zu lassen – und nach einer bestimmten Anzahl von Wiederholungen (*Iterationen*) abbrechen. Er kann den Prozess aber auch immer weiter fortführen, bis keine neuen Entscheidungen mehr hinzukommen, sodass das resultierende Netzwerk in sich geschlossen ist. Grundsätzlich betrifft das Rekursionsproblem also die Frage, nach wie vielen Iterationen der Netzwerkdefinitionsprozess abgeschlossen wird. Hier zeigt sich das Zusammenspiel von objektgeleiteter und beziehungsgeleiteter Herangehensweise besonders deutlich, auf das man sich oft einlassen muss, um zu einer auf die Untersuchungsfrage zugeschnittenen Netzwerkdefinition zu gelangen.

Das Rekursionsproblem wird auf die Spitze getrieben, wenn der Netzwerkdefinitionsprozess sich nicht (wie vorstehend geschildert) über Selektions- und Inklusionsproblem schrittweise vollzieht, sondern sich in einer bestimmten Anzahl rekursiver Operationen erschöpft. Dies ist insbesondere beim in der Sozialen Netzwerkanalyse beliebten *snowball sampling* der Fall, das die Beziehungen von Objekten einer Ursprungsmenge heranzieht, um weitere Objekte zum Netzwerk hinzuzufügen.<sup>137</sup> Ist die Anzahl der Elemente in der Ursprungsmenge gleich 1 (betrachtet man also etwa nur eine einzige Gerichtsentscheidung und alle Entscheidungen, die diese zitiert oder von denen diese zitiert wird), so nennt man das resultierende Netzwerk ein *ego-zentriertes Netzwerk* (*ego-centered network*), kurz *Egonetzwerk* (*ego network*).<sup>138</sup>

Die Betrachtung von Egonetzwerken bietet sich für die juristische Netzwerkforschung vor allem dann an, wenn Einfluss und Beeinflussung eines bestimmten Rechtssubjekts, Rechtstexts oder Rechtsbegriffs analysiert werden sollen. Sie ist allerdings definitionsbedingt nur in der Lage, unmittelbare Einflüsse und Beeinflussungen zu erfassen; sobald man sich also auch für mittelbare Effekte interessiert, muss man auf nicht ego-zentrierte Netzwerke zurückgreifen. Da man aus allgemeiner definierten Netzwerken stets die Egonetzwerke einzelner Elemente extrahieren kann, bietet es sich folglich an, zunächst (soweit praktisch möglich) dem hier skizzierten dreischrittigen Prozess zu folgen und Selektions-, Inklusions- und Rekursionsproblem in der oben genannten Reihenfolge abzuarbeiten.

Die vorstehenden Ausführungen zeigen, dass das Netzwerkdefinitionsproblem sowohl Probleme der Objekt- und Beziehungsdefinition als auch Probleme ihrer Interaktion umfasst. Vor diesem Hintergrund mag es verwundern, dass das Netzwerkdefinitionsproblem und die Konsequenzen der Entscheidungen, die im

<sup>137</sup> Grundlegend zum *snowball sampling* Goodman, *The Annals of Mathematical Statistics* 32 (1961), S. 148–170. Zu den Problemen von *snowball sampling* in der Netzwerkforschung siehe etwa Newman, *Social Networks* 25 (2003), S. 83–95.

<sup>138</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 42; Newman, *Networks*, 2010, S. 44. Die Anzahl der rekursiven Operationen, die zur Definition eines Egonetzwerks erforderlich sind, heißt dessen *Radius*.

Rahmen seiner Bewältigung getroffen werden müssen, in vielen Studien zu Netzwerken im Recht nur wenig bis gar keine Aufmerksamkeit erfahren.<sup>139</sup> Mögliche Erklärungen für diesen Befund können bei den Anreizstrukturen der wissenschaftlichen Publikationsinfrastruktur ebenso ansetzen wie bei den Zielen, die mit netzwerkanalytischen Studien verfolgt werden. Letztere sind Gegenstand des nächsten Abschnitts. Hier sei nur noch abschließend festgehalten:

„To represent an empirical phenomenon as a network is a theoretical act. It commits one to assumptions about what is interacting, the nature of that interaction, and the time scale on which that interaction takes place. Such assumptions are not ‘free,’ and indeed they can be wrong. [...] [T]he appropriate choice of representation is key to getting the correct result.“<sup>140</sup>

## C. Perspektiven

In der Literatur wird die Netzwerkforschung häufig *insgesamt* als Perspektive bezeichnet.<sup>141</sup> Bei näherem Hinsehen ist der Fokus der Arbeiten, die unter diesem Titel (oder einer seiner gängigen Varianten) firmieren, allerdings so unterschiedlich, dass sich die Netzwerkforschung genauer als *Bündel von Perspektiven* begreifen lässt. Diese Perspektiven können – je nach Untersuchungsfrage – auch im Rahmen der juristischen Netzwerkforschung eingenommen werden. Sie lassen sich nach ihren Untersuchungszielen (I.) und Untersuchungsebenen (II.) unterscheiden.

<sup>139</sup> Die Autoren schildern die Zusammenstellung ihres Datensatzes in der Regel nur äußerst knapp, ohne auf Definitionsschwierigkeiten einzugehen. Beispiele: *Cross, Smith und Tomarchio*, *Emory Law Journal* 57 (2008), S. 1227 (1237–1238); *Dibadj*, *San Diego Law Review* 46 (2009), S. 1 (8–9); *Derlén und Lindholm*, *European Law Journal* 20 (2014), S. 667 (672); *Derlén und Lindholm*, *German Law Journal* 18 (2017), S. 647 (656–657). Wie die Knappheit der Ausführungen zustande kommt, ist unklar. Vermutlich wirken freie Entscheidungen aufseiten der Autoren mit Vorgaben vonseiten des gewählten Publikationsmediums zusammen. Eine Ausnahme ist *Chandler*, *The Mathematica Journal* 10 (2007), S. 501 (502–508), der sich allerdings schon durch die Wahl des Publikationsmediums an ein technisch versiertes Publikum wendet. Verhältnismäßig ausführlich sind auch die Ausführungen bei *Lupu und Voeten*, *British Journal of Political Science* 42 (2012), S. 413 (424), deren inhaltliche Problematik die Wichtigkeit der Offenlegung selbst verdeutlicht: Die Autoren beginnen mit 7.319 Entscheidungen und 35.963 Zitaten des Europäischen Gerichtshofs für Menschenrechte und analysieren letztlich 2.222 Entscheidungen (30,4 %) und 16.863 Zitate (46,9 %).

<sup>140</sup> *Butts*, *Science* 325 (2009), S. 414 (416).

<sup>141</sup> So etwa bei *Smith*, *San Diego Law Review* 44 (2007), S. 309 (341); *Easley und Kleinberg*, *Networks, Crowds, and Markets*, 2010, S. 20; *Stegbauer und Häußling*, Einleitung in das Handbuch Netzwerkforschung, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 13 (13).

## I. Untersuchungsziele

Im Rahmen der Untersuchungsziele der Netzwerkforschung sind drei Schwerpunkte zu identifizieren: die Enthüllung von Mustern in Daten (*Exploration*, 1.), die Bestätigung von Hypothesen anhand vorhandener Daten (*Explikation*, 2.), und die Vorhersage von Daten auf Basis von Mustern in vorhandenen Daten (*Extrapolation*, 3.). Diese Ziele hängen zwar miteinander zusammen, gehen aber mit unterschiedlichen Forschungsansätzen einher, erfordern daher unterschiedliche Forschungsdesigns und sollten deshalb auch in der juristischen Netzwerkforschung getrennt voneinander verfolgt werden.

Im Folgenden werden die Untersuchungsziele der Netzwerkforschung vorgestellt. Dabei erfolgt zunächst eine allgemeine Charakterisierung (jeweils a)). Anschließend wird zuerst auf die Möglichkeiten der jeweiligen Zielsetzung in der juristischen Netzwerkforschung eingegangen (jeweils b)), bevor die Verantwortung der Rechtswissenschaft im Zusammenhang mit dem allgemeinen Netzwerkforschungsdiskurs skizziert wird, der das jeweilige Ziel auf dem Terrain des Rechts verfolgt (jeweils c)).

### 1. Exploration

#### a) Charakteristika

Die Forschung mit dem Ziel der Exploration beginnt bei den Daten und sucht in ihnen nach Mustern – sie ist datengeleitet (*data-driven*). Der zu untersuchende Datensatz wird in der Hoffnung zusammengestellt, Strukturen und Dynamiken in ihm zu entdecken, die sich so auch in Datensätzen aus gänzlich anderen Kontexten finden lassen. Man ist auf der Suche nach dem Universellen, sodass sich dieser Ansatz auch als *kontextfrei* bezeichnen lässt.<sup>142</sup> Kurz: Der explorativen Forschung kommt es auf das *Allgemeine* im Besonderen an. Zu dessen Identifikation arbeitet sie mit großen Datensätzen (in der Regel mehr als 1.000 Datenpunkte), die sie mit statistischen Methoden untersucht oder computergestützt modelliert.

Die Exploration ist in der Netzwerkforschung insbesondere die Domäne der statistischen Physiker. Diese untersuchen auch Netzwerke im Recht, vor allem Netzwerke aus Gerichtsentscheidungen und solche aus Patenten, auf universelle Strukturen – und werden regelmäßig fündig.<sup>143</sup> Dabei greifen sie auf Werkzeuge der deskriptiven und explorativen Statistik ebenso zurück wie auf solche der induk-

<sup>142</sup> So Zweig, Are Word-Adjacency Networks Networks?, in: Mehler u.a. (Hrsg.), Towards a Theoretical Framework for Analyzing Complex Linguistic Networks, 2016, S. 153 (155).

<sup>143</sup> Beispiele: Leicht u.a., The European Physical Journal B 59 (2007), S. 75–83; Karrer und Newman, Physical Review E 80 (2009), S. 046110-1–046110-12; Clough u.a., Journal of Complex Networks 3 (2015), S. 189–203.

tiven Statistik, wobei die Frage nach der statistischen Signifikanz identifizierter Strukturen in jüngerer Zeit immer häufiger gestellt wird.<sup>144</sup>

### b) Möglichkeiten

Die explorative juristische Netzwerkforschung ist zunächst für die Rechtsdogmatik von Interesse. Sie kann die Ergebnisse rechtsdogmatischer Analysen intersubjektiv überprüfbar machen, etwa indem sie Rechtsfiguren und Fallgruppen als Strukturen in Netzwerken aus Rechtstexten sichtbar macht und ihre Entwicklung über die Zeit verfolgt. Sie ist aber auch für Rechtstheorie, Rechtsgeschichte und Rechtsvergleichung geeignet. Hier kann sie helfen, Fragen wie die folgenden zu beantworten:

- Gibt es universelle Strukturen im Recht, die über Zeit und Raum konstant sind?
- Lässt sich die Vorstellung von Recht als System über die Theorie der komplexen adaptiven Systeme<sup>145</sup> mit der Vorstellung von Recht als Netzwerk vereinen?
- Gibt es quantifizierbare Unterschiede in Struktur und Dynamik von Rechtssystemen, die in der klassischen Rechtsvergleichung als *Common Law* klassifiziert werden, und jenen, die als *Civil Law* klassifiziert werden?

Wo ihre Erkundungen sich auf Sachverhalte richten, deren Kontrolle das Recht für sich in Anspruch nimmt, kann die explorative juristische Netzwerkforschung schließlich auch Rechtsdogmatik und Regulierungstheorie bei der Suche nach sachangemessenen Regelungen begleiten. So legt etwa die Feststellung, dass der Interbankenmarkt eine Struktur hat, die beim Ausfall besonders gut vernetzter Akteure insgesamt zusammenbricht, ein Regelungsregime nahe, das sich weniger an der Bilanzsumme denn an der Vernetzung der Akteure untereinander orientiert.<sup>146</sup>

Die Exploration von Rechtstatsachen durch die juristische Netzwerkforschung kann demnach die Entwicklung robuster Regelungsregimes und damit die Eindämmung systemischen Risikos unterstützen. Überall dort, wo sich zeigen lässt, dass die Funktion oder Funktionsfähigkeit eines bestimmten Sektors weniger von den Attributen individueller Akteure als von ihrer Vernetzung untereinander abhängt, können existierende oder zu schaffende Regelungen

<sup>144</sup> Beispiele: *Lancichinetti u.a.*, PLoS ONE 6 (2011), e18961-1–e18961-18; *Mirshahvalad u.a.*, PLoS ONE 7 (2012), e33721-1–e33721-7; *Wilson u.a.*, Annals of Applied Statistics 8 (2014), S. 1853–1891.

<sup>145</sup> Dazu insbesondere *Miller und Page*, Complex Adaptive Systems, 2007, sowie die weiteren Nachweise in Kapitel 1 Fn. 20 (S. 9).

<sup>146</sup> Die Struktur des Interbankenmarkts untersuchen beispielsweise *Pecora und Spelta*, Physica A 434 (2015), S. 1–12. Dort heißt es unter anderem: „In the light of these considerations, the Single Supervisory Mechanism implemented by the ECB is related to the financial stability of the biggest and most important Eurozone-based banks. We found that [...] results indicate that not all the financial institutions with high value of total asset are systemically important but only few of them.“ *Pecora und Spelta*, Physica A 434 (2015), S. 1 (11).

unter Berücksichtigung der Netzwerkstrukturen überdacht werden.<sup>147</sup> Die dabei absehbar auftretenden neuen Probleme – etwa, dass die „Systemrelevanz“ eines Akteurs mitunter aus einer Netzwerkposition resultiert, die er nicht vollständig selbst kontrollieren kann – sind wiederum von Interesse für die Rechtsdogmatik. Diese kann versuchen, die Erkenntnisse der Netzwerkforschung über die Figur des „Öffentlichen Interesses“ in ihr verwaltungsrechtliches Gebäude zu integrieren, oder sich die Frage stellen, ob die verfassungsrechtliche Schutzpflichtdogmatik in den hierarchischen Strukturen, die viele reale Netzwerke aufweisen, wirksamen Grundrechtsschutz gewährleisten kann.<sup>148</sup>

### c) Verantwortung

Umgekehrt kann – und sollte – sich die Rechtswissenschaft an der Erkundung ihres eigenen Terrains mit den Mitteln der Netzwerkforschung auch beteiligen, um die Weiterentwicklung der explorativen Netzwerkforschung zu unterstützen. Nicht-juristischen Studien zu Netzwerken im Recht liegen bisweilen zweifelhafteste Annahmen über Funktionsweise und Bedeutung von Recht zugrunde; viele Autoren schöpfen das Erkenntnispotenzial ihrer Daten mangels Interesse an den rechtlichen Einzelheiten nicht aus.<sup>149</sup>

<sup>147</sup> Besonders deutlich im Kontext der Finanzmärkte etwa *Battiston, Caldarelli und D’Errico*, The Financial System as a Nexus of Interconnected Networks, in: *Garas* (Hrsg.), *Interconnected Networks*, 2016, S. 195 (196): „The paradigm on which systemic importance is based must nowadays shift towards the inclusion of the ‚position‘ of an institution in the network.“

<sup>148</sup> Wo in juristischen Diskussionen von Netzwerken gesprochen wird, werden diese gern als Alternative zu hierarchischen Organisationen präsentiert, dazu zusammenfassend *Braun*, *Leitbilder im Recht*, 2015, S. 139–156. Dieser zumeist metaphorische Gebrauch des Netzwerkbegriffs ist irreführend, weil er den Blick auf die oft ungleiche Machtverteilung innerhalb von Netzwerken verstellt. Er kann sogar zu vollkommen fehlgeleiteten Maßnahmen führen, etwa bei der Planung und Umsetzung institutioneller Reformen.

<sup>149</sup> So generieren etwa die Arbeiten von *Tarissan und Nollez-Goldbach* zahlreiche Statistiken dazu, wie die Entscheidungen innerhalb einzelner Verfahren vor dem Internationalen Strafgerichtshof einander zitieren: *Tarissan und Nollez-Goldbach*, The Network of the International Criminal Court Decisions as a Complex System, in: *Sanayei, Zelinka und Rossler* (Hrsg.), *Emergence, Complexity and Computation*, 2013, S. 255–264; *Tarissan und Nollez-Goldbach*, Temporal Properties of Legal Decision Networks: A Case Study from the International Criminal Court, in: *Rotolo* (Hrsg.), *Legal Knowledge and Information Systems*, 2015, S. 111–120; *Tarissan und Nollez-Goldbach*, *Journal of Complex Networks* 4 (2016), S. 616–634. Die Autoren scheinen davon auszugehen, dass sich ihre Erkenntnisse und methodischen „Innovationen“ ohne Weiteres auf andere juristische Zitiernetzwerke übertragen lassen, obwohl die Arbeitsweise des Internationalen Strafgerichtshofs sich in vielerlei Hinsicht von der Arbeitsweise nationaler oder supranationaler Gerichte unterscheidet. Auch haben Zitate in einem Netzwerk aus Gerichtsentscheidungen, die einem einzigen Verfahren entstammen, ganz andere Motivationen als in einem Netzwerk aus Gerichtsentscheidungen, die überwiegend verschiedenen Verfahren entstammen, sodass kein einheitlicher Wichtigkeitsbegriff zugrunde gelegt werden kann. Die Diskussion dieser Einzelheiten könnte neue Perspektiven auf die Funktion verschiedener Gerichtstypen eröffnen, wird von den Autoren aber ausgespart.

Juristen können auf solche Mängel und ungenutzten Chancen im allgemeinen Diskurs zur Netzwerkforschung aufmerksam machen. Sie sollten sich zudem an der Erforschung des Zusammenspiels verschiedener Netzwerke beteiligen. Es ist nicht auszuschließen, dass das Recht auf Struktur und Dynamik nicht-juristischer Netzwerke auf eine spezifische Art und Weise einwirkt, die bei Studien ohne rechtswissenschaftliche Beteiligung im Dunkeln bleibt.

## 2. Explikation

### a) Charakteristika

Die explikative Forschung sucht nicht nur nach Mustern in Daten, sondern auch nach Gründen für diese Muster, weshalb sie jedenfalls in ihrem Kernstadium als hypothesengeleitet (*hypothesis-driven*) bezeichnet werden kann. Die Erzeugung der Hypothesen erfolgt in der Regel durch eine Beschäftigung mit dem Untersuchungskontext im Vorfeld der Analyse. Die Analyse selbst konzentriert sich dann auf die Interpretation der Befunde im Kontext. Man interessiert sich für das Kontextspezifische, sodass sich dieser Ansatz auch als *kontextuell* bezeichnen lässt.<sup>150</sup> Kurz: Der explikativen Forschung kommt es auf das *Besondere* im Besonderen an. Um dieses aufspüren zu können, arbeitet sie traditionell mit kleinen und mittelgroßen Datensätzen (in der Regel weniger als 1.000 Datenpunkte), die nicht selten speziell für die geplante Studie zusammengestellt werden.

Die explikative Netzwerkforschung hat ihre Wurzeln in der Soziologie und dort in der Sozialen Netzwerkanalyse (*social network analysis*, SNA). Deren Geburt wird klassischerweise mit den Arbeiten von *Jacob Levy Moreno* in Verbindung gebracht.<sup>151</sup> *Moreno* untersucht in seinem 1934 erschienenen Werk „Who Shall Survive?“ die Entwicklung und die Strukturen von Gruppen. Er stellt diese unter anderem mithilfe einer besonderen Art der Zeichnung dar, der er den Namen Soziogramm (*sociogram*) gibt (Abb. 2.6, S. 66).<sup>152</sup>

Wer im Buch von *Moreno* liest, wird dort zahlreiche Passagen finden, die sich eher der Exploration als der Explikation zuordnen lassen, und so baut auch die Explikation auf der Exploration auf. Entscheidend ist, dass *Moreno* sich auch für die materiellen Hintergründe der Strukturen interessiert, die er durch seine Erkundungen freilegt. Er bleibt nicht bei der Feststellung, dass ein Individuum eine auffällige Position in seinem Netzwerk hat, sondern setzt diese Position zu den Attributen des Individuums in Beziehung; er fragt nicht nur nach dem *Was*, sondern auch nach dem *Warum*. Ein solches Vorgehen ist für die explikative Netzwerkforschung charakteristisch. Dabei ist auffällig, dass gerade die explikative

<sup>150</sup> So *Zweig*, Are Word-Adjacency Networks Networks?, in: *Mehler u.a.* (Hrsg.), Towards a Theoretical Framework for Analyzing Complex Linguistic Networks, 2016, S. 153 (154).

<sup>151</sup> Siehe etwa *Wasserman und Faust*, Social Network Analysis, 1994, S. 10–13. Kritisch *Ziegler*, Zeitschrift für Klinische Psychologie 16 (1987), S. 339–352; differenzierend *Freeman*, The Development of Social Network Analysis, 2004, S. 7–8.

<sup>152</sup> *Moreno*, Who Shall Survive?, 1934, S. 26, 31–45.

Netzwerkforschung im Recht auch heute noch oft ohne induktive Statistik – also insbesondere ohne formale Hypothesentests – auskommt,<sup>153</sup> was ihr bisweilen den Vorwurf einbringt, rein deskriptiv zu sein.<sup>154</sup>

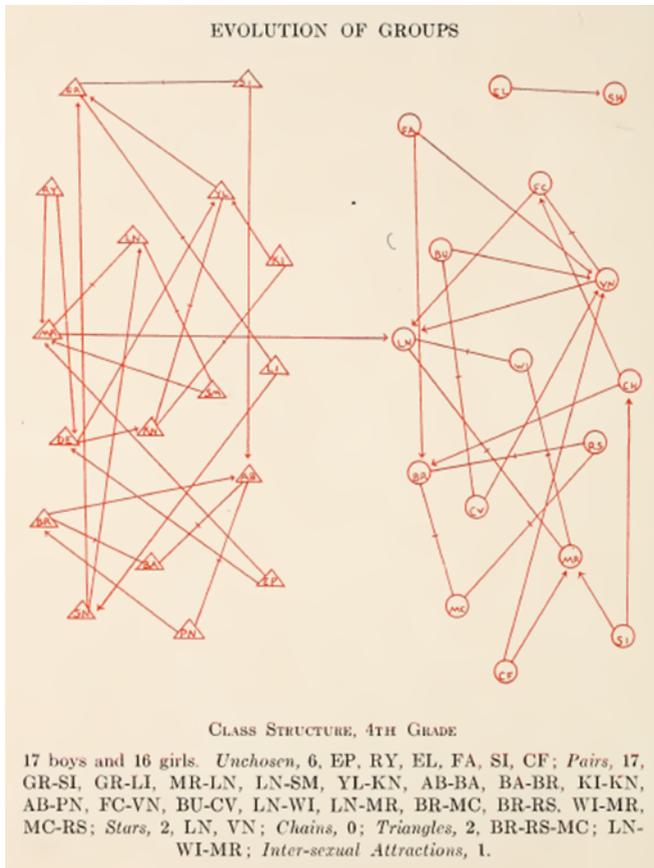


Abbildung 2.6: Morenos Soziogramm einer 4. Klasse<sup>155</sup>

<sup>153</sup> Dies kritisiert z.B. *Frankenreiter*, German Law Journal 18 (2017), S. 687 (691–692) an *Derlén und Lindholm*, German Law Journal 18 (2017), S. 647–686, die aus einem beschreibenden Vergleich eines Zitiernetzwerks des EuGH mit einem Zitiernetzwerk des *U.S. Supreme Court* folgern, die Rechtsprechung des EuGH folge der Logik eines *Case Law*-Systems.

<sup>154</sup> So etwa die Kritik bei *Whalen*, Michigan State Law Review 2016, S. 539 (556), und *Frankenreiter*, German Law Journal 18 (2017), S. 687 (692), der die Klage von *Whalen* wörtlich zitiert. Eine andere Frage ist, ob die Deskriptivität allein als Vorwurf taugt – vorwerfbar scheint allenfalls die Diskrepanz zwischen Anspruch und Realität zu sein, wenn die Darstellung suggeriert, man habe Zusammenhänge gefunden, diese Zusammenhänge dann aber nur auf anekdotische Evidenz gestützt werden.

Die explikative Netzwerkanalyse ist noch immer vor allem Domäne der Sozialwissenschaften.<sup>156</sup> In jüngerer Zeit geht allerdings auch dort der Trend zur Analyse größerer Datensätze, womit zwangsläufig eine gewisse Dekontextualisierung einhergeht. Umgekehrt lassen sich in den Naturwissenschaften Bemühungen feststellen, die in Daten gefundenen Muster im Kontext zu interpretieren. Diese Konvergenz der Perspektiven scheint mit der Verbreitung des Terminus „Netzwerkforschung“ zur Kennzeichnung des eigenen Forschungsgebiets einherzugehen.

### b) Möglichkeiten

Die explikative juristische Netzwerkforschung ist vor allem für die Rechtssoziologie und die Rechtsgeschichte interessant. Hier kann sie bei der Annäherung an die seit jeher drängende Frage helfen, welche Faktoren die Entwicklung und die Wirkung von Recht in welchen Situationen auf welche Art und Weise besonders stark beeinflussen. Die so gefundenen Anhaltspunkte können Arbeitsmaterial für die Rechtstheorie liefern. Diese kann sich dann – etwa im Wege der explorativen juristischen Netzwerkforschung – der Frage widmen, welche der gefundenen Spezifika orts- und zeitunabhängig vor allem in Rechtskontexten, nicht aber in Kontexten ohne Rechtsbezug gelten, welche Spezifika also *Rechtsspezifika* sind.

Für die Selbstbeobachtung der Rechtswissenschaft oder die Rechtswissenschaftstheorie kann die explikative Erforschung rechtswissenschaftlicher Netzwerke dabei helfen, einzelne Diskursphänomene (z.B. Rezeptionsprozesse) aufzudecken, zu verstehen und gegebenenfalls zu kritisieren.<sup>157</sup>

Wo sich die explikative juristische Netzwerkforschung mit der Entwicklung und Wirkung von Recht oder aber auch mit den Wechselwirkungen zwischen Recht und seinen Kontexten befasst, können ihre Ergebnisse schließlich auch die Rechtspolitik informieren. Zeigt sich etwa, dass eine Regelung ihr Ziel aufgrund unerwarteter Netzwerkdynamiken im geregelten Gebiet verfehlt, so kann die Rechtspolitik versuchen, durch Änderung der Rechtslage gegenzusteuern. Offenbart sich, dass bestimmte staatliche Institutionen Pathologien entwickeln,

<sup>155</sup> Quelle: *Moreno*, Who Shall Survive?, 1934, S. 38; ähnlich bereits erschienen in *Moreno*, Emotions Mapped by New Geography, The New York Times (3. April 1933), S. 17; zu den Hintergründen *Nolte*, The Philosophy, Theory and Methods of J. L. Moreno, 2014, S. 147. Ähnliche Darstellungen finden sich auch bei *Northway*, Sociometry 3 (1940), S. 144–150, und in *Northways* späteren Arbeiten, z.B. *Northway*, Sociometry 17 (1954), S. 272–281 (274, 276).

<sup>156</sup> Mit „Sozialwissenschaften“ sind hier unter anderem Soziologie, Politikwissenschaften und Ökonomie gemeint. Für Soziologie und Physik bringt *Stauffer* die unterschiedlichen Forschungsinteressen am Beispiel der Modellierung von Wahlergebnissen auf den Punkt: „Physiker finden die beobachtete Universalität der Wahlergebnisse interessant; Soziologen mögen eher fragen, warum ein bestimmter Kandidat eine bestimmte Wahl gewann.“ *Stauffer*, Small World, in: *Stegbauer und Häußling* (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 219 (224).

<sup>157</sup> In diese Richtung (wenngleich *nicht* explikativ) etwa *Katz u.a.*, Journal of Legal Education 61 (2011), S. 76–103, der die Ausbreitung von Ideen durch ein Netzwerk von *Law Schools* modelliert, die über die Einstellung von (an anderen *Law Schools* ausgebildeten) Professoren miteinander verbunden sind.

welche die Leistungsfähigkeit des Rechtsstaats beeinträchtigen, so ist dies ein Mandat zur Reform, dessen Umsetzung (möglicherweise politische Selbstrestriktion erforderlich) durch den von der juristischen Netzwerkforschung ausgehenden Transparenzdruck erleichtert werden könnte.

### c) Verantwortung

Es liegt in der Verantwortung der Rechtswissenschaft, auf zweifelhafte Prämissen hinsichtlich der Funktionsweise des Rechts, seiner Institutionen und Akteure hinzuweisen, die der nicht-juristischen Netzwerkforschung im Recht teilweise zugrunde liegen. Zwar ist nicht unwahrscheinlich, dass die rechtswissenschaftliche Perspektive bisweilen durch Selbstbetroffenheit und ein Übermaß an Detailwissen verzerrt ist. Dadurch kann sie aber gerade ein Gegengewicht bilden zur durch Nicht-Betroffenheit und ein Untermaß an Detailwissen verzerrten Perspektive, die bei Nicht-Rechtswissenschaftlern mitunter anzutreffen ist.

Die Einwirkung der Rechtswissenschaft auf den allgemeinen Netzwerkforschungsdiskurs ist notwendig, um zu verhindern, dass im Rahmen der explikativen Netzwerkforschung im Recht Scheinerkenntnisse produziert werden. Hier kann die explikative juristische Netzwerkforschung dazu beitragen, dass sich in der Literatur ein Bewusstsein für die Eigenheiten des Rechts und die damit einhergehenden Möglichkeiten und Grenzen der Erkenntnis durch Netzwerkforschung entwickelt.

## 3. Extrapolation

### a) Charakteristika

Die Forschung mit dem Ziel der Extrapolation baut auf Exploration und Explikation auf, um über diese hinauszugehen: Sie identifiziert Muster in vorhandenen Daten und versucht, anhand dieser Muster weitere Daten vorherzusagen (*prediction*).<sup>158</sup> Die gewünschte Vorhersage kann sich dabei auf Daten beziehen, die in der Zukunft liegen. Sie kann aber auch Daten betreffen, die schon in der Gegenwart existieren, dem Forscher aber unbekannt sind – etwa, weil es sich um ein kriminelles Netzwerk handelt, dessen Beteiligte im Verborgenen operieren oder ihre Spuren verwischen.<sup>159</sup> Kurz: Der extrapolierenden Forschung kommt es sowohl

<sup>158</sup> So nutzen beispielsweise *Fowler u.a.*, *Political Analysis* 15 (2007), S. 324 (336–343) die Maße, die sie für die Wichtigkeit einzelner Entscheidungen des *U.S. Supreme Court* entwickeln, um vorherzusagen, wie diese nach ihrer Veröffentlichung zitiert werden (dazu bereits S. 13–17).

<sup>159</sup> Beispiel: *Rhodes und Keefe*, *Journal of the Operational Research Society* 58 (2007), S. 1605–1611. Die akkurate Vorhersage fehlender Beziehungen in kriminellen Netzwerken ist oft Voraussetzung für deren effektive Destabilisierung (etwa durch gezielte Isolation von Schlüsselpersonen). Zu den Schwierigkeiten im Umgang mit kriminellen Netzwerken siehe auch *Carley, Lee und Krackhardt*, *Connections* 24 (2002), S. 79–92; *Carley*, *Computational & Mathematical Organization Theory* 12 (2006), S. 51–66; *Everton*, *Disrupting Dark Networks*, 2012, S. 365–383; *Duijn, Kashirin und Sloot*, *Scientific Reports* 4 (2014), S. 4238-1–4238-15.

auf das *Allgemeine* als auch auf das *Besondere* im Besonderen an – aber jeweils nur, soweit es für die angestrebte Vorhersage relevant ist.<sup>160</sup>

Die Extrapolation ist in der Netzwerkforschung vor allem das Gebiet der *Data Scientists* und der Informatiker, die sich in den Bereichen *Data Mining*, *Machine Learning* oder *Artificial Intelligence* spezialisiert haben.<sup>161</sup> Ihre Fähigkeit, Aussagen über zukünftige Entwicklungen zu treffen, macht sie besonders praxisrelevant und damit auch kommerziell attraktiv; gleichzeitig ist extrapolierende Netzwerkforschung ganz besonders „*easy to do badly*.“<sup>162</sup> Die extrapolierende Netzwerkforschung kommt nicht ohne induktive Statistik aus und wird in jüngerer Zeit – wie alle Forschung, die sich an Vorhersagen versucht – immer stärker von der bayesianischen Statistik beeinflusst.<sup>163</sup>

### b) Möglichkeiten

Die extrapolierende juristische Netzwerkforschung ist für die *Rechtspraxis* in zweierlei Hinsicht interessant. Erstens lässt sich die Extrapolation auf die Entwicklung des Rechts selbst beziehen. Aus dieser Perspektive kann sie einerseits Rechtsprechung und Verwaltung dabei unterstützen, Konsistenz in ihren Entscheidungspraktiken sicherzustellen. Andererseits kann sie Anwälten und ihren potenziellen Mandanten helfen, ihr Prozessrisiko besser einzuschätzen und ihre Argumentation vor Gericht an die erwartbaren Referenzpräferenzen (wer oder was soll zitiert werden?) der möglicherweise entscheidenden Richter anzupassen. Zweitens lässt sich die Extrapolation aber auch auf zukünftige Entwicklungen außerhalb des Rechts beziehen. Aus dieser Warte kann sie Private beispielsweise bei der Entscheidung unterstützen, mit wem sie in längerfristige Vertragsbeziehungen eintreten, und dem Staat insbesondere bei der Prävention – etwa von Straftaten oder Epidemien – behilflich sein.

<sup>160</sup> Allgemein zu den Voraussetzungen und Potenzialen der extrapolierenden Netzwerkforschung *Zweig*, *Central European Journal of Computer Science* 1 (2011), S. 137 (142–144, 149–150).

<sup>161</sup> Der Übergang zwischen den genannten Feldern ist fließend und die Terminologie wird auch in der Fachliteratur nicht einheitlich verwendet. Unterschiede bestehen vor allem mit Blick auf die Schwerpunktsetzung und das gängige Vokabular.

<sup>162</sup> *Larose und Larose*, *Data Mining and Predictive Analytics*, 2015, S. 6, dort bezogen auf *Data Mining* im Allgemeinen.

<sup>163</sup> Die *bayesianische Statistik* interpretiert Wahrscheinlichkeiten als Maß für die Glaubwürdigkeit von Aussagen, während die (im 20. Jahrhundert dominierende, weniger rechenintensive) *frequentistische Statistik* Wahrscheinlichkeiten als Ergebnisse unendlich oft wiederholbarer Zufallsexperimente auffasst. Beispiele für den Einsatz bayesianischer Statistik in der Netzwerkforschung: *Butts*, *Social Networks* 25 (2003), S. 103–140; *Koskinen und Snijders*, *Journal of Statistical Planning and Inference* 137 (2007), S. 3930–3938; *Guimerà und Sales-Pardo*, *PNAS* 106 (2009), S. 22073–22078; *Caimo und Friel*, *Social Networks* 33 (2011), S. 41–55; *Koskinen u.a.*, *Social Networks* 35 (2013), S. 514–527. Dass bayesianische Ansätze auch an anderen Stellen in der Rechtswissenschaft viel Potenzial aufweisen, zeigen unter anderem *Bent*, *Denver University Law Review* 91 (2014), S. 807–851, und *Bent*, *Buffalo Law Review* 63 (2015), S. 85–145.

Der *Rechtswissenschaft* kann die Extrapolation unter anderem helfen, das Verhältnis von Kontinuität und Wandel im Recht zu untersuchen. Geringere Kontinuität der Rechtsentwicklung kann in geringerer Vorhersagbarkeit der Rechtsprechung (oder auch Gesetzgebung) aus Merkmalen vergangener Entscheidungen (oder auch Gesetze) ihren Ausdruck finden. Etwaige Kontinuitätsunterschiede aufzuspüren und auf Unterschiede in Gremienbesetzungen, Entscheidungszeiträumen oder Jurisdiktionen zurückzuführen, verspricht für die Rechtssoziologie ebenso Erkenntnisgewinn wie für die Rechtsgeschichte und die Rechtsvergleichung. Für die Rechtstheorie ist dann wiederum interessant, welche Muster sich in den Kontinuitätsunterschieden und möglichen Begründungen erkennen lassen, die als Ausdruck charakteristischer Strukturen und Funktionen des Rechts gedeutet werden können.

### c) Verantwortung

Im Umgang mit der Extrapolation in der Netzwerkforschung kommt der Rechtswissenschaft eine besonders große Verantwortung zu. Wo sich die Extrapolation auf die Entwicklung des Rechts selbst bezieht, ergibt sich diese Verantwortung aus ihren *rechtsstaatlichen* Implikationen: Die Extrapolationsergebnisse lassen sich kommerzialisieren. Das hat erstens zur Folge, dass sich das Kräfteverhältnis zwischen etwaigen Konfliktparteien vor Gericht, sofern es überhaupt zum Prozess kommt, noch stärker als bislang zugunsten der zahlungskräftigeren Partei verschiebt. Zweitens besteht die Gefahr, dass potenzielle Prozessparteien aufgrund falscher Risikoprognosen durch irrende Algorithmen davon abgehalten werden, ihre Rechte (gerichtlich oder außergerichtlich) durchzusetzen. Drittens – und besonders problematisch – droht aber eine Erstarrung der Rechtsentwicklung: Wenn potenziellen Klägern auf Basis von Merkmalen vergangener Entscheidungen eine äußerst geringe Erfolgschance vorausgesagt wird, gelangen ihre Fälle möglicherweise sehr viel seltener zu Gericht. Richter erhalten dementsprechend seltener die Gelegenheit, Fehlentscheidungen zu korrigieren oder ihre Rechtsprechungslinien an gesamtgesellschaftliche Entwicklungen anzupassen. Dieses Erstarrungsrisiko verstärkt sich dadurch, dass viele Rechtsschutzversicherungen bei niedrigen Erfolgsprognosen, die noch dazu mit algorithmischer Objektivität daherkommen, eine Deckungszusage ablehnen werden.

Weder Rechtswissenschaft noch Gesetzgeber werden verhindern können, dass kommerzielle Anbieter auf Basis von (relationalen) Rechtsdaten über kurz oder lang versuchen werden, Rechtsextrapolation zu betreiben. Daher ist es umso wichtiger, dass die Rechtswissenschaft – allen voran die juristische Netzwerkforschung – diese Entwicklungen kritisch begleitet und sich in die Lage versetzt, die Methodik der Vorhersagen nachzuvollziehen und zu prüfen.

Wo die Extrapolation sich nicht auf die Rechtsentwicklung, sondern auf Entwicklungen außerhalb des Rechts bezieht, ergibt sich die Verantwortung der Rechtswissenschaft aus ihren *gesellschaftlichen* Implikationen: Schon jetzt ist es etwa möglich, das Kreditausfallrisiko einer Person oder ihr Risiko, straffällig

zu werden, anhand ihres Wohnortes oder ihrer Beziehungsdaten zu prognostizieren.<sup>164</sup> Das kann die Entwicklungs- und Gestaltungsmöglichkeiten Einzelner in Abhängigkeit von ihrer Netzwerkposition erheblich beeinträchtigen, obwohl sie über ihre Netzwerkposition nur eingeschränkt selbst entscheiden können.<sup>165</sup> Wie lässt sich dies mit dem Grundgedanken individueller Autonomie vereinbaren, der das (deutsche) Rechtssystem prägt? Will man eine Extrapolation auf Basis von Netzwerkdaten überhaupt zulassen – und wenn ja, wie kann sich der Einzelne gegen etwaige Extrapolationsfehler wehren?

Die Beantwortung dieser und ähnlicher Fragen muss unter Einmischung von Rechtsphilosophie und Rechtsdogmatik erfolgen, sofern das Recht seinen gesellschaftlichen Ordnungsanspruch wahren will. Die Rechtsphilosophie kann fragen, wie die immer stärkeren und immer stärker manipulierbaren Abhängigkeiten zwischen gesellschaftlichen Einheiten – seien es nun Individuen oder Kollektive – mit den Individualitäts- und Freiheitskonzepten des geltenden Rechts in Einklang gebracht werden *sollen*. Der Rechtsdogmatik kommt dann die Aufgabe zu, identifizierte Änderungsnotwendigkeiten mit Blick auf das geltende Recht in Rechtskonzepte *de lege ferenda* zu übersetzen, die Gesetzgebung und Rechtsprechung als Inspiration und Anleitung dienen können.

## II. Untersuchungsebenen

Unabhängig davon, wie ein Netzwerk zusammengesetzt ist und mit welchem Ziel es analysiert wird, lassen sich in der Netzwerkforschung drei Untersuchungsebenen unterscheiden: die Mikroebene (1.), die Mesoebene (2.) und die Makroebene (3.).<sup>166</sup> Diese Ebenen finden sich auch in der Struktur der nachfolgenden Kapitel wieder. Auf jeder der Untersuchungsebenen werden spezifische Fragen an das betrachtete Netzwerk herangetragen und spezifische Instrumente zu seiner Analyse genutzt. Während sich die Untersuchung auf der Mikroebene auf einzelne Objekte im Netzwerk konzentriert, nimmt sie auf der Mesoebene Gruppen von Objekten und auf der Makroebene das Netzwerk als Ganzes in den Fokus (Abb. 2.7, S. 72).<sup>167</sup> Um das Fundament für eine vertiefende Betrachtung im

<sup>164</sup> Zur Beurteilung der Kreditwürdigkeit Einzelner mithilfe von relationalen Daten etwa *Wei u.a.*, *Marketing Science* 35 (2016), S. 234–258.

<sup>165</sup> Beispiel: Ein Mieter hat typischerweise wenig Kontrolle darüber, wer in die benachbarten Wohnungen einzieht.

<sup>166</sup> In manchen Fällen kann eine weitere Ausdifferenzierung sinnvoll sein; dazu etwa *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 36–37. Die Differenzierung zwischen Mikro-, Meso- und Makroebene ist der Verfasserin zuerst begegnet bei *Börner*, *Atlas of Knowledge*, 2015, S. 4–11 (dort im Kontext der *Science of Science*). Sie liegt den meisten Arbeiten aus der Netzwerkforschung (manchmal explizit, oft implizit) zugrunde.

<sup>167</sup> Diese Differenzierung widerspricht nicht der Tatsache, dass Theoriebildung und Modellierung in der Netzwerkforschung ihrem Selbstverständnis nach auf der Mesoebene ansetzen, hervorgehoben unter anderem von *Stegbauer und Häußling*, Einleitung: Selbstverständnis der Netzwerkforschung, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 57 (59). Vielmehr werden auch Mikro- und Makrophänomene untersucht und vor dem Hinter-

Rahmen des dritten Kapitels zu legen, werden Mikro-, Meso- und Makroebene der Netzwerkforschung anhand ihrer typischen Untersuchungsfragen im juristischen Kontext nachfolgend jeweils knapp charakterisiert.

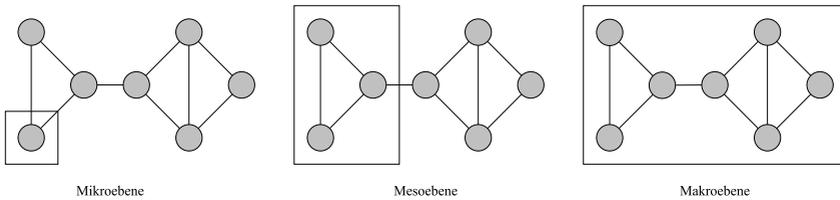


Abbildung 2.7: Untersuchungsebenen der Netzwerkforschung

### 1. Mikroebene

Auf der Mikroebene interessiert man sich für die *lokalen* Strukturen eines Netzwerks und für die Eigenschaften seiner einzelnen Objekte und Beziehungen. Die zentrale Fragestellung ist hier: Welche Elemente des betrachteten Netzwerks sind besonders wichtig? Diese Frage wird im Folgenden kurz als *Wichtigkeitsfrage* bezeichnet.

Die juristische Netzwerkforschung kann beispielsweise fragen (*mögliche juristische Erkenntnisperspektive(n) stets in Klammern*):

- Was sind die zentralen zivilrechtlichen Rechtsfiguren im deutschen Recht, wie haben sich diese über die Zeit entwickelt und wie sind sie mit den zentralen zivilrechtlichen Rechtsfiguren im französischen Recht zu vergleichen?  
(*Rechtsdogmatik, Rechtstheorie, Rechtsvergleichung*)
- Welche Personen haben besonderen Einfluss auf die Rechtsentwicklung im Sexualstrafrecht?  
(*Rechtssoziologie, Rechtsdogmatik*)
- Welche Entscheidungen haben die Rechtsprechung des Bundesverfassungsgerichts besonders geprägt?  
(*Rechtsdogmatik*)

Der kritische Leser wird an dieser Stelle mit Recht auf die Auslegungsbedürftigkeit des Wichtigkeitsbegriffs und seiner Artverwandten („zentral“, „besonderer Einfluss“, „besonders geprägt“) hinweisen, die auch ein Problem für die juristische Netzwerkforschung darstellt. Dieses Problem behandelt Kapitel 3 Abschnitt C.II. (S. 129–142).

---

grund einer mittleren Position – weder sehr wenige noch sehr viele Akteure, weder allwissend noch nichtswissend, weder vollkommen unabhängig noch vollkommen abhängig – gedeutet; dazu *Miller und Page*, *Complex Adaptive Systems*, 2007, S. 213–226.

## 2. Mesoebene

Auf der Mesoebene interessiert man sich für die *regionalen* Strukturen eines Netzwerks und die Eigenschaften seiner regionalen Struktureinheiten. Die zentrale Fragestellung ist hier: Welche Elemente des betrachteten Netzwerks gehören zusammen? Diese Frage wird im Folgenden kurz als *Gruppierungsfrage* bezeichnet.

Die juristische Netzwerkforschung kann beispielsweise fragen:

- Welche Charakteristika kennzeichnen den deutschen Verbraucherbegriff im Zivilrecht, wie hat sich dieser mit der Zeit entwickelt und wie verhält er sich zum europäischen Verbraucherbegriff?  
(*Rechtsdogmatik, Rechtsvergleichung*)
- Ist Wirtschaftsstrafrecht Wirtschaftsrecht oder Strafrecht? — Oder allgemeiner: Woran erkennt man überhaupt ein Rechtsgebiet?  
(*Rechtsdogmatik, Rechtstheorie*)
- Gibt es in der Staatsrechtswissenschaft Zitierkreise, die sich an politischen Präferenzen orientieren?  
(*Rechtssoziologie*)

Unklar ist, woran man die Zusammengehörigkeit von Netzwerkelementen erkennt. Mit diesem Problem befasst sich Kapitel 3 Abschnitt C.III. (S. 142–165).

## 3. Makroebene

Auf der Makroebene interessiert man sich für die *globalen* Strukturen eines Netzwerks. Die zentrale Fragestellung ist hier: Welche Eigenschaften sind für das betrachtete Netzwerk als Ganzes – insbesondere im Vergleich zu anderen Netzwerken – charakteristisch? Diese Frage wird im Folgenden kurz als *Charakterisierungsfrage* bezeichnet.

Die juristische Netzwerkforschung kann beispielsweise fragen:

- Gibt es rechtsraumspezifische, rechtsgebietsspezifische oder funktionspezifische Strukturen in juristischen Zitiernetzwerken?  
(*Rechtsvergleichung, Rechtsdogmatik, Rechtssoziologie*)
- Welche Netzwerke im Recht sind welchen Netzwerken außerhalb des Rechts auf welche Art und Weise besonders ähnlich?  
(*Rechtstheorie, Rechtssoziologie*)
- Wie beeinflusst das EU-Recht das deutsche Recht?  
(*Rechtstheorie, Rechtsdogmatik, Rechtsvergleichung*)

Hier steht man nun vor der Frage, was genau unter den Charakteristika eines Netzwerks zu verstehen ist. Dieser Frage widmet sich Kapitel 3 Abschnitt C.IV. (S. 165–186).

## Kapitel 3

# Methodik

In diesem Kapitel werden die methodischen Grundlagen der juristischen Netzwerkforschung eingeführt. Welche Werkzeuge stehen ihr zur Verfügung, um die im zweiten Kapitel geschilderten theoretischen Erwägungen in rechtswissenschaftliche Erkenntnis umzusetzen? Juristische Netzwerkforschung ist Arbeit mit Daten (A.). Die Methoden, die dabei zum Einsatz kommen, lassen sich nach ihrem Schwerpunkt in drei Kategorien einordnen: Visualisierung (B.), Quantifizierung (C.) und Modellierung (D.). Sie sind überwiegend auf Netzwerke mit *einer* Art von Objekten und *einer* Art von Beziehungen zugeschnitten, können aber auch für die Analyse von Netzwerken mit heterogener Zusammensetzung weiterentwickelt werden (E.).

### A. Arbeitsmaterial: Daten

Daten sind das Arbeitsmaterial der juristischen Netzwerkforschung – nicht nur als Rohstoff, sondern auch als Zwischenprodukt oder Endergebnis. Sie können aus der Realität stammen (*reale Daten*) oder speziell für Untersuchungszwecke erstellt worden sein (*fiktive Daten*). Die Arbeit mit Daten bringt einige Herausforderungen mit sich, die zunächst für reale juristische Daten skizziert werden (I.). Es folgt die Einführung zweier fiktiver Datensätze, anhand derer sich die Methoden der juristischen Netzwerkforschung illustrieren lassen, ohne den Unsicherheiten realer Daten ausgesetzt zu sein (II.). Schließlich werden einige Termini zur Beschreibung von Daten zusammengestellt, die im weiteren Verlauf dieser Arbeit immer wieder zum Einsatz kommen (III.).

#### I. *Reale Daten*

Juristische Netzwerkforschung zielt darauf ab, Recht als beobachtbares Phänomen besser zu verstehen. Dazu muss sie früher oder später mit realen Daten arbeiten. Das können Daten sein, die aus juristischer Sicht interessant sind, mit denen sich aber vornehmlich andere Wissenschaften beschäftigen, etwa Wirtschaftsdaten oder Sozialdaten (*juristische Daten im weiteren Sinne*). Es können allerdings ebenso Daten sein, die in erster Linie für Juristen relevant sind, z.B. Daten zu den Strukturen von Gesetzen, zur Entwicklung der Rechtsprechung oder zur

Ausgestaltung von Verträgen in der Kautelarpraxis (*juristische Daten im engeren Sinne*).<sup>1</sup>

Die Arbeit mit juristischen Daten bereitet einige grundsätzliche Probleme. Sie lassen sich einteilen in Probleme der Sammlung, der Analyse und der Kommunikation juristischer Daten.<sup>2</sup> Einige dieser Schwierigkeiten betreffen alle Disziplinen, die mit Daten arbeiten, andere sind rechtsspezifisch. Für die juristische Netzwerkforschung sind vor allem rechtsspezifische Herausforderungen relevant, die mit der Sammlung juristischer Daten im engeren Sinne verbunden sind. Diese Herausforderungen betreffen insbesondere die Beschaffung, Überprüfung und Aufbereitung juristischer Texte.<sup>3</sup> Hier geht es einerseits darum, Datenquellen zu finden, die sich als Grundlage einer juristischen Netzwerkstudie eignen (1.). Andererseits geht es darum, die Fehlerquellen juristischer Daten (in Abhängigkeit von der Datenquelle) zu identifizieren und den damit einhergehenden Unsicherheiten Rechnung zu tragen (2.).

### 1. Quellen juristischer Daten

Ein wesentliches Anwendungsfeld der juristischen Netzwerkforschung ist die Untersuchung von Strukturen in juristischen Texten.<sup>4</sup> Im Arbeitsalltag nutzen Juristen analoge und digitale Quellen, um sich aus solchen Texten zu informieren. Dabei greifen sie insbesondere auf Printmedien (Beispiel: die Entscheidungssammlung BVerfGE im Buchformat), kommerzielle juristische Datenbanken (Beispiel: *juris*) und die Internetauftritte juristischer Institutionen (Beispiel: die Websites<sup>5</sup> der Bundesgerichte) zurück. Für den wissenschaftlichen Einsatz weisen diese Quellen aber einige Defizite auf, welche die Entwicklung der juristischen Netzwerkforschung erschweren. Hier lassen sich *inhaltliche* (a)), *technische* (b)) und *rechtliche* (c)) Schwierigkeiten unterscheiden.

<sup>1</sup> Im deutschsprachigen juristischen Diskurs kommt der Ausdruck „juristische Daten“ außerhalb der Rechtsinformatik kaum vor. Was genau Daten zu „juristischen“ Daten macht, wird selten angesprochen. Eine Suche nach der Phrase „juristische Daten“ im Dezember 2017 ergab 55 Treffer in *juris*, 2 Treffer in *beck-online* und 129 Treffer in *jurion*. Die meisten dieser Treffer waren auf andere Phrasen zurückzuführen, die „juristische Daten“ als Präfix enthalten, etwa „juristische Datenbanken“, „juristische Datenverarbeitung“ oder „juristischer Datenbestand“.

<sup>2</sup> Dazu bereits *Coupette und Fleckner*, JZ 73 (2018), S. 379–389.

<sup>3</sup> Allgemein *Coupette und Fleckner*, JZ 73 (2018), S. 379–389.

<sup>4</sup> Dazu bereits S. 13–17 und S. 32–36.

<sup>5</sup> Zur Erinnerung: Eine Webseite (*web page*) ist ein einzelnes Dokument, das über das *World Wide Web* abrufbar und anhand seiner URL eindeutig identifizierbar ist (z.B. [www.bundesgerichtshof.de/DE/DasGericht/dasGericht\\_node.html](http://www.bundesgerichtshof.de/DE/DasGericht/dasGericht_node.html)). Eine Website (*website*) ist eine Menge von Webseiten, die unter einer Domain (z.B. [www.bundesgerichtshof.de](http://www.bundesgerichtshof.de)) abrufbar sind.

### a) Inhaltliche Schwierigkeiten

Inhaltliche Schwierigkeiten sind alle Probleme, welche die Definition des Untersuchungsgegenstands in ihrer praktischen Umsetzung (*Operationalisierung*) betreffen. Für die juristische Netzwerkforschung geht es hier um Herausforderungen bei der Operationalisierung der gewählten Netzwerkdefinition. Erste Unsicherheiten ergeben sich daraus, dass viele juristische Daten nicht von ihren ursprünglichen Produzenten, sondern von Dritten bereitgestellt werden. Wer diese Daten nutzt, erhält sie also nicht aus offizieller Quelle und muss hinsichtlich ihrer inhaltlichen Integrität auf den Drittanbieter vertrauen. Hinzu kommt, dass Drittanbieter juristischer Daten die Originaldaten meist in einer für ihre Nutzer aufbereiteten Form zur Verfügung stellen, sodass die Daten nicht unverändert sind. Eine weitere Herausforderung besteht schließlich darin, die inhaltliche Vollständigkeit der untersuchten Daten mit Blick auf die Definition des Untersuchungsgegenstands sicherzustellen – also zu gewährleisten, dass die Daten alles enthalten, was von der Definition erfasst wird.

Will man etwa „alle Zitate“ zwischen „allen Entscheidungen“ eines bestimmten Gerichts analysieren, so muss man den Ausdruck „alle“ zweimal konkretisieren. Bei „allen Entscheidungen“ besteht ein Problem darin, dass es in der Regel keine Quelle gibt, über die alle getroffenen oder auch nur alle veröffentlichten Entscheidungen zugänglich sind.<sup>6</sup> In ihrer Selbstbeschreibung geben viele Quellen sinngemäß an, „alle wichtigen“ Entscheidungen zu umfassen, wobei „wichtig“ wiederum unterschiedlich ausgelegt wird (z.B. „alle Entscheidungen, welche die entscheidenden Richter für wichtig halten“ oder „alle Entscheidungen, welche die Mitglieder einer besonderen Redaktion für wichtig halten“).<sup>7</sup> Hier bleibt dem Forscher oft nichts anderes übrig, als die Datenquelle in die Definition des Untersuchungsgegenstands einzubeziehen (z.B. „alle Entscheidungen, die in den

---

<sup>6</sup> Das betrifft bereits die Rechtsprechung der Bundesgerichte, insbesondere aber jene von Gerichten unterhalb der Bundesebene.

<sup>7</sup> Bemerkenswert ist, wie wenig sich auf den Websites der Gerichte über das Zustandekommen der landläufig als „amtlich“ bezeichneten Entscheidungssammlungen von BAG, BFH, BGH, BPatG, BSG, BVerfG und BVerwG in Erfahrung bringen lässt. Diese werden von den Mitgliedern der Gerichte (und im Falle des BGH auch von Mitgliedern der Bundesanwaltschaft) herausgegeben und in verschiedenen Verlagen publiziert. Deutlich ist die Beschreibung zur Sammlung BFHE auf der Website des BFH: „Die Senate beschließen bei jeder Entscheidung, ob diese zur amtlichen Veröffentlichung freigegeben wird. Das hängt davon ab, ob der Senat in seiner Entscheidung grundsätzliche oder in anderer Weise bedeutsame Aussagen getroffen hat. [...] Die zur amtlichen Veröffentlichung freigegebenen Entscheidungen werden nach entsprechender Neutralisierung [...] und nach Bildung eines Leitsatzes, der die Kernaussage der Entscheidung wiedergibt, in der von den Mitgliedern des Bundesfinanzhofs herausgegebenen – so genannten amtlichen – Sammlung der Entscheidungen des Bundesfinanzhofs – BFHE – veröffentlicht.“ Die Websites anderer Gerichte weisen zum Teil nicht einmal explizit auf die Existenz der „amtlichen“ Entscheidungssammlung(en) hin; Näheres ergibt sich allerdings zum Teil aus den Geschäftsordnungen der Gerichte, z.B. § 18 der Geschäftsordnung des BGH und § 31 der Geschäftsordnung des BVerfG.

ersten 100 Bänden BGHZ veröffentlicht wurden“ oder „alle Entscheidungen des BGH, die für die Jahre 2000 bis 2010 über *juris* verfügbar sind“) oder mit den Institutionen, welche die Daten generieren, direkt zu kooperieren (womit „alle Entscheidungen“ zu „alle zur Verfügung gestellten Entscheidungen“ wird).

Bei „allen Zitaten“ ist zunächst inhaltlich unklar, was überhaupt als Zitat zählt (jede – explizite oder implizite – Form der Bezugnahme auf eine andere Entscheidung oder nur Bezugnahmen, die einem bestimmten Muster folgen?). Außerdem stellt sich die Frage, ob mit „allen Zitaten“ tatsächlich „alle Vorkommen von Zitaten“ oder aber nur „alle zitierten Entscheidungen“ gemeint sind. Es liegt nahe, hier „alle Vorkommen von Zitaten“ zu antworten, da sich „alle zitierten Entscheidungen“ daraus ableiten lassen. Da Zitate häufig in Gruppen auftauchen, kann man sogar noch weiter gehen und für jede Entscheidung alle *Zitierblöcke* erfassen. Zitierblöcke sind Sequenzen von Zitaten, die in einem Dokument unmittelbar hintereinander stehen (in deutschen Gerichtsentscheidungen typischerweise in einem Klammerzusatz); aus ihnen lassen sich alle Vorkommen von Zitaten ableiten.<sup>8</sup> Der Unterschied zwischen „allen zitierten Entscheidungen“, „allen Vorkommen von Zitaten“ und „allen Zitierblöcken“ ist in Abbildung 3.1 (S. 77) dargestellt.

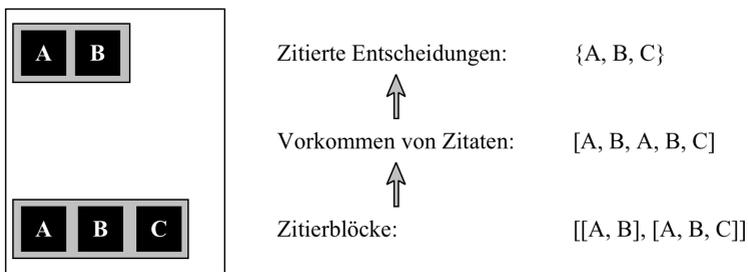


Abbildung 3.1: Möglichkeiten der Zitaterfassung<sup>9</sup>

<sup>8</sup> Liegen zuverlässige Informationen zu den Absätzen der untersuchten Gerichtsentscheidungen vor, so können die Zitate auch absatzbezogen erhoben werden; so geschehen etwa bei *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327–349. Da in einem Absatz mehrere Zitierblöcke vorkommen können, lässt sich das Ergebnis der einen Erhebung allerdings nicht aus dem Ergebnis der anderen ableiten.

<sup>9</sup> Der weiße Kasten auf der linken Seite stellt eine Entscheidung dar, die an zwei verschiedenen Stellen (graue Kästen) andere Entscheidungen zitiert, also zwei Zitierblöcke enthält. Im ersten Zitierblock werden die Entscheidungen A und B, im zweiten Zitierblock die Entscheidungen A, B und C zitiert (schwarze Kästen). In der abgebildeten Entscheidung gibt es also drei zitierte Entscheidungen und fünf Vorkommen von Zitaten. Die zitierten Entscheidungen sind strukturell eine Menge (in der Abbildung rechts gekennzeichnet durch geschweifte Klammern), denn es gibt keine Duplikate und die Reihenfolge der Zitate spielt keine Rolle. Die Vorkommen von Zitaten lassen sich unter anderem als Liste festhalten (indiziert durch eckige Klammern), da diese eine Reihenfolge hat und Duplikate zulässt. Zitierblöcke kann man sich als Liste von Listen vorstellen (angezeigt durch eckige Klammern, die ineinander verschachtelt sind).

Je detaillierter die Zitierdaten erfasst werden, desto vielfältiger sind die Möglichkeiten, die aus ihnen konstruierten *Rechtsprechungsziernetzwerke* (*judicial citation networks*) zu analysieren. Soll ermittelt werden, wie häufig einzelne Entscheidungen zitieren oder zitiert werden, so lassen sich zwei *Zählmodi* unterscheiden: Bei *binärer Zählung* von Zitaten ermittelt man für jede Entscheidung lediglich, *ob* diese eine andere Entscheidung zitiert bzw. von dieser zitiert wird. Bei *gewichteter Zählung* hält man zusätzlich fest, *wie häufig* eine Entscheidung eine andere zitiert bzw. von einer anderen Entscheidung zitiert wird. Die Entscheidung in Abbildung 3.1 (S. 77) zitiert nach binärer Zählung dreimal, nach gewichteter Zählung fünfmal; die Entscheidungen A und B werden nach binärer Zählung jeweils einmal, nach gewichteter Zählung jeweils zweimal zitiert (für C sind binäre und gewichtete Zählung identisch). Während es zur binären Zählung von Zitaten ausreicht, für jede Entscheidung die von ihr zitierten Entscheidungen zu erheben, ist für die gewichtete Zählung mindestens die Erhebung aller Vorkommen von Zitaten erforderlich. Ein Netzwerk, das die binäre Zählung ermöglicht, lässt sich als gerichteter Graph *ohne* Mehrfachkanten darstellen (jeder Pfeil von einem Punkt zu einem anderen Punkt kommt keinmal oder einmal vor). Demgegenüber muss ein Netzwerk, das die gewichtete Zählung ermöglicht, als gerichteter Graph *mit* Mehrfachkanten oder als gerichteter, gewichteter Graph *ohne* Mehrfachkanten dargestellt werden (Mehrfachzitate können entweder als Mehrfachkanten oder als Kantengewichte repräsentiert werden).

Es liegt nahe, dass sich mit einer Untersuchung auf Basis der binären Zählung von Zitaten andere Einblicke gewinnen lassen als mit der gewichteten Zählung.<sup>10</sup> Im Rahmen der Datensammlung sollte man daher nach Möglichkeit nicht nur die zitierten Entscheidungen, sondern zumindest alle Vorkommen von Zitaten erheben. In vielen Studien zu Rechtsprechungsziernetzwerken heißt „alle Zitate“ dennoch „alle zitierten Entscheidungen“.<sup>11</sup> Das hängt auch damit zusammen, dass die Informationen zu den zitierten Entscheidungen oft den Entscheidungsannotationen der juristischen Datenbanken entnommen werden können (wie zuverlässig die dortigen Angaben sind, hängt vom Datenbankanbieter ab). „Alle Vorkommen von Zitaten“ lassen sich hingegen typischerweise nur aus den Entscheidungstexten

<sup>10</sup> Der praktische Nutzen dieser Differenzierung wird im vierten Kapitel (S. 227–319) anhand von Zitierdaten zur deutschen Rechtsprechung untersucht.

<sup>11</sup> Beispiele: *Fowler und Jeon*, *Social Networks* 30 (2008), S. 16 (17–18); *Dibadj*, *San Diego Law Review* 46 (2009), S. 1 (9–10); *Lupu und Voeten*, *British Journal of Political Science* 42 (2012), S. 413 (424); *Schaper*, *The Structure and Organization of EU Law in the Field of Direct Taxes*, 2013, S. 144; *Derlén und Lindholm*, *European Law Journal* 20 (2014), S. 667 (672); *Tarissan, Panagis und Šadl*, *IEEE/ACM ASONAM* 2016, S. 661 (663); *Derlén und Lindholm*, *Journal of International Economic Law* 20 (2017), S. 257 (265). Ausnahmen (gewichtete Zählung): *Clark, Lauderdale und Katz*, *Political Analysis* 20 (2012), S. 329 (333–334); *Charlotin*, *Journal of International Economic Law* 20 (2017), S. 279 (286–287).

gewinnen (*Extraktion* von Zitaten) – mit Verfahren, die meist nur einen Teil „aller Zitate“ erfassen.<sup>12</sup>

### b) Technische Schwierigkeiten

Technische Schwierigkeiten sind alle Barrieren, die einer unmittelbaren automatisierten Weiterverarbeitung von Daten entgegenstehen und mit dem Bereitstellungsformat der Datenquellen verbunden sind. Unmittelbare automatisierte Weiterverarbeitung meint die strukturierte Speicherung und systematische Auswertung der Daten mithilfe von Computerprogrammen (*Code*) ohne (semi-)manuelle Vorbereitung.<sup>13</sup> Die Möglichkeit solcher Verarbeitungsmaßnahmen ist Voraussetzung dafür, dass größere Datenmengen mit vertretbarem Aufwand analysiert werden können.

Bei Printmedien beginnen die Probleme bereits mit der Digitalisierung. Denn für die automatisierte Weiterverarbeitung genügt es nicht, bedrucktes Papier als Bilddateien zu scannen. Damit der dort abgebildete Text maschinell ausgewertet werden kann, muss der Scan zusätzlich ein Texterkennungsverfahren (*optical character recognition*, OCR) durchlaufen.<sup>14</sup> Texterkennungsverfahren funktionieren heute zwar deutlich besser als noch vor einigen Jahren, haben aber nach wie vor Schwierigkeiten mit juristischen Texten in deutscher Sprache. Das liegt

<sup>12</sup> Ein konkretes Beispiel, in dem auch Zitierblöcke erhoben werden, findet sich im vierten Kapitel (S. 240–248).

<sup>13</sup> Auf die Ermöglichung unmittelbarer automatisierter Weiterverarbeitung zielt auch das Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen (Informationsweiterverwendungsgesetz – IWG). § 3 II IWG lautet: „Informationen sind in allen angefragten Formaten und Sprachen, in denen sie bei der öffentlichen Stelle vorliegen, zur Weiterverwendung zur Verfügung zu stellen; soweit möglich und wenn damit für die öffentliche Stelle kein unverhältnismäßiger Aufwand verbunden ist, sind sie vollständig oder in Auszügen *elektronisch* sowie *in einem offenen und maschinenlesbaren Format* zusammen mit den zugehörigen *Metadaten* zu übermitteln. Sowohl die Formate als auch die Metadaten sollten so weit wie möglich *anerkannten, offenen Standards entsprechen*.“ (Hervorhebungen hinzugefügt). Die zentralen Begriffe sind in § 2 Nr. 5–7 IWG definiert: „[Im Sinne des IWG] 5. ist maschinenlesbares Format ein Dateiformat, das so strukturiert ist, dass Softwareanwendungen bestimmte Daten, einschließlich einzelner Sachverhaltsdarstellungen und deren interner Struktur, leicht identifizieren, erkennen und extrahieren können, 6. ist offenes Format ein Dateiformat, das plattformunabhängig ist und der Öffentlichkeit ohne Einschränkungen, die der Weiterverwendung von Informationen hinderlich wären, zugänglich gemacht wird, 7. ist anerkannter, offener Standard ein schriftlich niedergelegter Standard, in dem die Anforderungen für die Sicherstellung der Interoperabilität der Software niedergelegt sind.“

<sup>14</sup> Geht es nicht nur um Texterkennung, sondern auch um die Erkennung anderer Inhalte, so spricht man von Dokumentenerkennung (*document image analysis*); zur Entwicklung beider Technologien im Überblick *Baird und Tombre*, *The Evolution of Document Image Analysis*, in: *Doermann und Tombre* (Hrsg.), *Handbook of Document Image Processing and Recognition*, 2014, S. 63–71, speziell zu Texterkennungsverfahren *Cao und Natarajan*, *Machine-Printed Character Recognition*, in: *Doermann und Tombre* (Hrsg.), *Handbook of Document Image Processing and Recognition*, 2014, S. 331–358, jeweils m.w.N.

daran, dass Texterkennung zu großen Teilen auf Mustererkennung basiert, die auf Vorwissen über erwartbare Muster angewiesen ist.<sup>15</sup> Jedem Zeichen entsprechen einige charakteristische Tintenstrukturen auf dem Papier, die von der verwendeten Schriftart abhängen, wobei insbesondere die Trennung einzelner Zeichen (*character segmentation*) schwierig sein kann.<sup>16</sup> Viele Texterkennungsverfahren nutzen außerdem Wörterbücher, um Fehler bei der Zeichenerkennung zu beheben.<sup>17</sup> Das kann dazu führen, dass richtig erkannte Folgen von Zeichen, die nicht in einem Wörterbuch stehen (und die auch nicht mittels Wortbildungsregeln aus bekannten Wörtern abgeleitet werden können), zu bekannten oder ableitbaren Wörtern „berichtigt“ (und damit verfälscht) werden.

An juristischen Texten ist problematisch, dass sie zahlreiche Muster beinhalten, die in alltagssprachlichen Texten kaum vorkommen, etwa mehrere Sonderzeichen mitten im Satz (z.B. „§§“), nach einem „h.“ ein „M.“ oder längere Ketten von Groß- und Kleinbuchstaben, die keine deutschen Wörter sind (z.B. „BlmSchG“). Die meisten Programme sind auf solche Muster (auch in ihrer Version für die deutsche Sprache) nicht eingestellt, da sie die Texterkennung nicht auf juristischen Texten „gelernt“ haben und das juristische Fachvokabular nicht kennen.<sup>18</sup> Dokumente mit großflächigen Schwärzungen oder Lücken im Text (mögliche Folgen von Anonymisierung), ältere Texte, die in nicht-standardisierten Schriftarten gedruckt oder in Fraktur gesetzt sind, sowie Drucke mit Tintenklecksen und fleckigem Papier bereiten noch größere Schwierigkeiten.<sup>19</sup> Die geschilderten Besonderheiten führen dazu, dass juristische Daten, die aus der Digitalisierung von Printmedien mithilfe von OCR als Texte hervorgehen, zahlreiche Fehler enthalten, die den weiteren Umgang mit diesen Texten (z.B. die Extraktion von Zitaten) erschweren. Dies ist insbesondere für Studien relevant, die historische Entwicklungen von

<sup>15</sup> Wie die Texterkennung im Einzelnen vonstatten geht, hängt von der Struktur und Arbeitsweise des verwendeten OCR-Systems ab. Zu den typischen Komponenten heutiger OCR-Systeme Cao und Natarajan, Machine-Printed Character Recognition, in: Doermann und Tombre (Hrsg.), Handbook of Document Image Processing and Recognition, 2014, S. 331 (336–338).

<sup>16</sup> Dazu Nobile und Suen, Text Segmentation for Document Recognition, in: Doermann und Tombre (Hrsg.), Handbook of Document Image Processing and Recognition, 2014, S. 257 (272–286).

<sup>17</sup> Baird und Tombre, The Evolution of Document Image Analysis, in: Doermann und Tombre (Hrsg.), Handbook of Document Image Processing and Recognition, 2014, S. 63 (66–67).

<sup>18</sup> Bisher gibt es noch kein frei zugängliches Korpus deutschsprachiger juristischer Texte, auf dem man ein Texterkennungsprogramm speziell zur Texterkennung für andere deutschsprachige juristische Texte „trainieren“ könnte. Zu den Bemühungen, ein juristisches Referenzkorpus zu entwickeln, und den bereits existierenden Ressourcen Vogel, Hamann und Gauer, Law & Social Inquiry 2017, DOI: 10.1111/lsi.12305.

<sup>19</sup> Zu den Problemen der Texterkennung bei historischen Dokumenten Berg-Kirkpatrick, Durrett und Klein, Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics 2013, S. 207–217. Besondere Herausforderungen bereitet auch die Aufbereitung von Dokumenten, in denen Fließtext und technische Zeichnungen gemischt vorkommen. Ein Beispiel hierfür sind Patente; dazu etwa Riedl u.a., International Journal on Document Analysis and Recognition 19 (2016), S. 155–172.

Gesetzgebung oder Rechtsprechung (z.B. die Entwicklung der ZPO seit ihrem erstmaligen Inkrafttreten oder die Rechtsprechung des Reichsgerichts) nachverfolgen wollen, da viele ältere Dokumente noch nicht oder nur ohne Texterkennung digitalisiert sind.

Doch auch die unmittelbare automatisierte Weiterverarbeitung juristischer Daten aus digitalen Quellen bereitet Schwierigkeiten. Programmierschnittstellen (*Application Programming Interfaces*, APIs) zu juristischen Datenbanken, die dem Forscher den systematischen Abruf großer Datenmengen ermöglichen, sind hier noch immer die Ausnahme.<sup>20</sup> Kommerzielle Datenbankanbieter stellen ihre Daten (z.B. die Texte von Gerichtsentscheidungen oder Kommentaren) oft so zur Verfügung, dass sie nur schwer automatisiert heruntergeladen werden können. Die bereitgestellten Formate sind ebenfalls nicht auf computergestützte Weiterverarbeitung ausgelegt: Daten, die als PDF- oder Word-Dokumente heruntergeladen werden, haben meist keine Struktur,<sup>21</sup> sodass schon einfache Fragen wie „Von welchem Gericht stammt die heruntergeladene Entscheidung?“ oder „Wer ist Bearbeiter des Kommentars?“ erst durch Informationsextraktion (*Text Mining*) beantwortet werden können (die persönliche Lektüre kommt als Auswertungsmodus bei großen Datenmengen nicht in Betracht).<sup>22</sup>

Insbesondere Gesetzestexte und Gerichtsentscheidungen werden zwar immer häufiger als semistrukturierte Dokumente, allen voran als XML-Dateien bereitgestellt.<sup>23</sup> Die XML-Versionen, die diesen Dateien zugrunde liegen, sind allerdings wiederum nicht einheitlich und die Dokumentenstrukturen (*Schemata*) sind teilweise so konstruiert, dass sie den an der automatisierten Auswertung interessierten

<sup>20</sup> Inspirierende Beispiele aus dem Ausland sind die APIs von *CourtListener* ([www.courtlistener.com/api/](http://www.courtlistener.com/api/)) und *PatentsView* ([www.patentsview.org/api/doc.html](http://www.patentsview.org/api/doc.html)).

<sup>21</sup> Struktur meint hier eine Organisation der Daten, die es ermöglicht, die Daten schnell nach verschiedenen Kriterien zu sortieren und abzufragen. Auch eine Gerichtsentscheidung, die für den Menschen deutlich sichtbar in Rubrum, Sachverhalt und Gründe unterteilt ist, hat in diesem Sinne keine Struktur. Allgemein unterscheidet man zwischen *strukturierten Daten* (*structured data*) und *unstrukturierten Daten* (*unstructured data*). Texte sind ein klassisches Beispiel für unstrukturierte Daten. Wenn Texte in Formaten bereitgestellt werden, die Strukturinformationen (z.B. als Markup oder Metadaten) mitführen, spricht man auch von *semistrukturierten Daten* (*semistructured data*) bzw. semistrukturierten Dokumenten; siehe etwa *Feldman und Sanger*, *The Text Mining Handbook*, 2007, S. 3–4.

<sup>22</sup> Eine Einführung in das *Text Mining* bieten *Feldman und Sanger*, *The Text Mining Handbook*, 2007.

<sup>23</sup> Beispiele sind die über die RSS-Feeds von BAG, BFH, BSG, BVerfG und BVerwG bereitgestellten Entscheidungen (der BGH stellt über seinen Feed lediglich PDF-Dokumente bereit, das BPatG bietet keinen Feed an), die Entscheidungen auf [www.rechtsprechung-im-internet.de](http://www.rechtsprechung-im-internet.de) und die Gesetzestexte auf [www.gesetze-im-internet.de](http://www.gesetze-im-internet.de). Zu den Potenzialen von XML im Recht beispielsweise *Vincent*, *S.M.U. Law Review* 53 (2000), S. 1395–1429; *Silverman*, *Washington Law Review* 2004, S. 175–221; *Governatori*, *International Journal of Cooperative Information Systems* 14 (2005), S. 181–216; *Cunningham*, *Washington University Law Review* 84 (2006), S. 313–373; *Athan u.a.*, *OASIS LegalRuleML*, in: *Vlek u.a.* (Hrsg.), *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, 2013, S. 3–12.

Forscher eher abschrecken.<sup>24</sup> Die seit geraumer Zeit bestehenden Bestrebungen, XML-Standards zur Repräsentation juristischer Texte zu schaffen, finden außerhalb der Rechtsinformatik (nicht nur in Deutschland) bisher wenig Beachtung.<sup>25</sup>

### c) Rechtliche Schwierigkeiten

Rechtliche Schwierigkeiten sind schließlich alle juristischen Gründe, die einer unmittelbaren automatisierten Weiterverarbeitung juristischer Daten entgegenstehen. Welche Faktoren das sind, hängt von der Datenquelle ebenso ab wie von der Gattung der betrachteten Texte. Sowohl bei analogen als auch bei digitalen Medien können Autoren- und Verlagsrechte problematisch sein, bei Gerichtsentscheidungen (vor ihrer Anonymisierung) außerdem die Rechte der darin genannten Personen. Werden die Daten über eine digitale Plattform bezogen, kommen die Rechte der Plattform hinzu.

In Deutschland genießen „Gesetze, Verordnungen, amtliche Erlasse und Bekanntmachungen sowie Entscheidungen und amtlich verfaßte Leitsätze zu Entscheidungen“ zwar nach § 5 I UrhG keinen urheberrechtlichen Schutz. Auf den Websites vieler Gerichte liest man dennoch, dass die dort abrufbaren Gerichtsentscheidungen nur zur nicht gewerblichen Nutzung kostenfrei bereitgestellt werden.<sup>26</sup> Bei der Nutzung aufbereiteter Entscheidungen stellt sich die Frage nach der Reichweite des § 5 I UrhG;<sup>27</sup> auch der Schutz von Datenbankwerken (§ 4 UrhG)

<sup>24</sup> Ein Beispiel für Abschreckung durch Schema ist die *Document Type Definition* der Gesetzestexte, die man über [www.gesetze-im-internet.de](http://www.gesetze-im-internet.de) erhält, verlinkt unter [www.gesetze-im-internet.de/hinweise.html](http://www.gesetze-im-internet.de/hinweise.html). Nach den persönlichen Erfahrungen der Verfasserin erstreckt sich die Abschreckungswirkung auf Rechtswissenschaftler, die mit Programmierkenntnissen ausgestattet sind, ebenso wie auf Informatiker, die an juristischen Projekten mitarbeiten.

<sup>25</sup> Zu diesen Bestrebungen siehe etwa *Vincent*, S.M.U. Law Review 53 (2000), S. 1395–1429; *Athan u.a.*, OASIS LegalRuleML, in: *Vlek u.a.* (Hrsg.), Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law, 2013, S. 3–12.

<sup>26</sup> So beispielsweise zu finden beim BGH, ähnlich bei BAG, BFH, BPatG, und BSG, weniger eindeutig bei BVerfG und BVerwG.

<sup>27</sup> Nach der Rechtsprechung des BGH können nichtamtlich verfasste Leitsätze gerichtlicher Entscheidungen als Bearbeitungen i.S.d. § 3 UrhG urheberrechtlich geschützt sein (BGH, Urteil vom 21. November 1991 – I ZR 190/89). Praktisch problematisch ist dabei nicht nur die den Urheberrechtsschutz begründende Schöpfungshöhe (dazu OLG Köln, Beschluss vom 28. August 2008 – 6 W 110/08), sondern auch die (teleologische) Auslegung und ggf. analoge Anwendung des § 5 I UrhG. Für die von der Dokumentationsstelle des BVerfG verfassten Orientierungssätze ist die Gemeinfreiheit i.S.d. § 5 I UrhG zuletzt mit ausführlicher Argumentation vom VGH Baden-Württemberg bejaht worden (VGH Baden-Württemberg, Urteil vom 7. Mai 2013 – 10 S 281/12), nachdem sie von der Vorinstanz noch verneint worden war (VG Karlsruhe, Urteil vom 3. November 2011 – 3 K 2289/09). Gegenstand des Verfahrens war ein Gleichbehandlungsanspruch nach § 3 I 1 Informationsweiterverwendungsgesetz (IWG) a.F. hinsichtlich der Überlassung von durch die Dokumentationsstelle aufbereiteten Entscheidungen des BVerfG, die bis dahin ausschließlich *juris* zur Verfügung gestellt worden waren. Der Rechtsstreit gelangte bis zum

und der Schutz des Datenbankherstellers (§§ 87a–87e UrhG) könnten problematisch sein.<sup>28</sup> Unklar ist beispielsweise, inwieweit die Sammlung von Daten aus einer juristischen Datenbank zur Durchführung einer juristischen Netzwerkstudie die Vervielfältigung eines nach Art oder Umfang wesentlichen Teils dieser Datenbank impliziert, die nach § 87b I 1 UrhG grundsätzlich dem Datenbankhersteller vorbehalten ist.<sup>29</sup> Dieses Recht schränkt § 87c I Nr. 2 UrhG zwar für den „eigenen wissenschaftlichen Gebrauch“ ein, „wenn und soweit die Vervielfältigung zu diesem Zweck geboten ist und der wissenschaftliche Gebrauch nicht zu gewerblichen Zwecken erfolgt“; nicht gestattet bleiben allerdings Verbreitung und öffentliche Wiedergabe, die nach § 87b I 1 UrhG ebenfalls ausschließlich dem Datenbankhersteller vorbehalten und von den Ausnahmeregelungen des § 87c I UrhG gerade nicht erfasst sind. Eine Vorfrage ist, ob juristische Datenbanken überhaupt den Schutz aus §§ 87a–87e UrhG beanspruchen können, soweit es um Dokumente geht, die nach § 5 I UrhG gemeinfrei sind – insbesondere dann, wenn diese nach dem Informationsweiterverwendungsgesetz (IWG) auch von den Stellen erlangt werden könnten, durch die sie dem Datenbankanbieter selbst zur Verfügung gestellt werden.<sup>30</sup> Und wie ist die Rechtslage, wenn für eine juristische Netzwerkstudie aus einer juristischen Datenbank gezielt diejenigen Dokumente herausgesucht werden, die in einer Entscheidungssammlung enthalten sind, welche von den Mitgliedern eines Gerichts herausgegeben wird?<sup>31</sup>

---

BVerwG, wo sich die Parteien anscheinend außergerichtlich einigten; siehe *Richter*, NVwZ 2016, S. 1143 (1143). Zuvor war ein Versuch der Klägerin gescheitert, sämtliche Richter des BVerwG aufgrund ihrer Mitgliedschaft im Verein der Bundesrichter bei dem Bundesverwaltungsgericht e.V. wegen Besorgnis der Befangenheit abzulehnen (BVerwG, Beschluss vom 29. Januar 2014 – 7 C 13.13).

<sup>28</sup> Die §§ 87a–87e UrhG dienen der Umsetzung der Datenbankrichtlinie (Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken, ABl L 77 vom 27. März 1996, S. 20–28), die bei der Auslegung und Anwendung der Normen zu berücksichtigen ist.

<sup>29</sup> Die Konkretisierung des Wesentlichkeitsbegriffs hat der Gesetzgeber der Rechtsprechung überlassen; dazu *Wiebe*, UrhG § 87b, in: *Spindler und Schuster* (Hrsg.), Recht der elektronischen Medien, 2015 (Rn. 16–26).

<sup>30</sup> Für die Anwendung des § 5 I UrhG auch im Rahmen der §§ 87a–87e UrhG etwa VGH Baden-Württemberg, Urteil vom 7. Mai 2013 – 10 S 281/12 (dazu bereits Fn. 27, S. 82), im Anschluss an einen Vorlagebeschluss des BGH (BGH, Beschluss vom 28. September 2006 – I ZR 261/03); zur europarechtlichen Grundlage der §§ 87a–87e UrhG bereits Fn. 28 (S. 83). Es ist allerdings nicht zur Entscheidung des EuGH über die Vereinbarkeit der dargestellten Rechtsauffassung mit der Datenbankrichtlinie gekommen, da die Revision vor dem BGH zurückgenommen wurde. Zur Diskussion in der deutschen Literatur *Wiebe*, UrhG § 5, in: *Spindler und Schuster* (Hrsg.), Recht der elektronischen Medien, 2015 (Rn. 3 m.w.N.).

<sup>31</sup> Gegen die Anwendung des § 5 I UrhG auf solche Entscheidungssammlungen ohne weitere Argumente etwa *Marquardt*, UrhG § 5, in: *Wandtke und Bullinger* (Hrsg.), Praxiskommentar zum Urheberrecht, 2014 (Rn. 13), da diese „trotz ihrer regelmäßigen Bezeichnung als ‚Amtliche Sammlung‘ private Sammlungen“ (Hervorhebungen weggelassen) seien; ebenso *Radke, Hilgert und Mardorf*, NVwZ 2008, S. 1070 (1072).

Die rechtlichen Rahmenbedingungen der automatisierten Weiterverarbeitung von Daten aus juristischen Datenbanken für rechtswissenschaftliche Zwecke werden in der (größtenteils über juristische Datenbanken abzurufenden) Literatur zu den einschlägigen Vorschriften – soweit ersichtlich – nicht thematisiert; hier besteht erheblicher Forschungsbedarf. Bis dieser Forschungsbedarf gedeckt ist, können Netzwerkforscher die ihren Studien zugrunde liegenden Daten kaum ohne Rechtsunsicherheit oder Rücksprache mit Dritten öffentlich zugänglich machen (vor allem, wenn diese Daten juristischen Datenbanken entstammen).<sup>32</sup> Dadurch wird anderen Forschern die Replikation der Ergebnisse juristischer Netzwerkstudien erschwert. Dies gilt insbesondere dann, wenn sich die Abdeckung der für eine Studie genutzten Datenbank – wie im juristischen Bereich üblich – mit der Zeit auch für Zeiträume verändert, die bereits in der Vergangenheit liegen, sodass auch die gleiche, mit zeitlicher Einschränkung versehene Suchanfrage zu verschiedenen Zeitpunkten unterschiedliche Ergebnisse liefern kann.<sup>33</sup>

Die geschilderten Herausforderungen sind nur die Spitze des Eisbergs. In Deutschland ist vor allem problematisch, dass es für den derzeitigen Stand der Technik kaum Präzedenzfälle oder *Best Practices* zur Sammlung juristischer Daten (insbesondere aus juristischen Datenbanken) gibt, an denen sich der Netzwerkforscher zur rechtlichen Absicherung orientieren könnte. Die juristische Netzwerkforschung (wie auch die empirische juristische Forschung im Allgemeinen) würde deutlich erleichtert, wenn sich die Wissenschaftler auf die technischen und inhaltlichen Herausforderungen der Arbeit mit juristischen Daten konzentrieren könnten, ohne rechtlichen Zweifeln ausgesetzt zu sein oder Erlaubnisse von Oligopolisten einholen zu müssen.

Auch das Verbot kommerzieller Nutzung ist selbst für die wissenschaftliche juristische Netzwerkforschung nur auf den ersten Blick bedeutungslos. Denn Forschung verursacht Kosten – und Forschung, deren Ergebnisse sich möglicherweise kommerzialisieren lassen (etwa als Bestandteil juristischer Suchmaschinen), kann eher fremdfinanziert werden. Die Gedanken der Forschungsförderung durch die Wirtschaft und der Forschungsk Kooperation mit der Wirtschaft sind der deutschen Rechtswissenschaft bisher weitgehend fremd, und die Gefahren für die wissenschaftliche Unabhängigkeit sind nicht zu unterschätzen. Der Blick in andere Länder zeigt jedoch, welcher (positive) Innovationsdruck für die etablierten Datenbankanbieter von jungen Unternehmen ausgehen kann, die juristische Daten systematisch auswerten.<sup>34</sup> In Deutschland gibt es derzeit auch für Wissenschaftler

<sup>32</sup> Dieses Problem wird für die Dokumentation des vierten Kapitels (S. 227–319) praktisch relevant.

<sup>33</sup> Da es in der Regel keine Möglichkeit gibt, Suchanfragen auf einen vergangenen Datenbankzustand zu beschränken, kann ein Forscher die genaue Datengrundlage einer Studie in solchen Fällen nur direkt von demjenigen erhalten, der sie durchgeführt hat – selbst dann, wenn das Vorgehen bei der Suche in der Studie vollständig offengelegt wird.

<sup>34</sup> Beispielsweise wurde das Startup *Ravel Law*, welches das Fallrecht der Bibliothek der *Harvard Law School* digitalisiert sowie quantitativ und visuell analysierbar gemacht hat, im Jahr 2017 von *LexisNexis* gekauft.

noch keine Möglichkeit, juristische Daten in größerem Umfang zu analysieren, ohne den hierfür erforderlichen Code selbst zu schreiben. Das liegt unter anderem daran, dass es rechtlich kaum möglich ist, die hierfür geeigneten Anwendungen zu entwickeln, ohne mit etablierten Datenbankanbietern zu kooperieren (die wiederum kaum Anreize haben, am *status quo* etwas zu ändern). Aus rechtsstaatlicher Perspektive wäre es allerdings jeder kommerziellen Lösung vorzuziehen, wenn die Produzenten juristischer Daten in Deutschland (also insbesondere Parlamente, Behörden und Gerichte) sich mit der Wissenschaft verbündeten, um diese Daten über das Internet vollständig öffentlich zugänglich zu machen und jedem Interessierten eine systematische Auswertung zu ermöglichen.

## 2. Fehlerquellen juristischer Daten

Die in inhaltlicher, technischer und rechtlicher Hinsicht problematische Quellenlage für juristische Daten, insbesondere für juristische Texte, lässt bereits erahnen, dass auch ihre Analyse mit vielen Schwierigkeiten verbunden ist. 100 % Prozent Genauigkeit sind bei computergestützten Verfahren, die mit begrenztem Wissen (d.h. *heuristisch*) und mit Näherungen (d.h. *approximativ*) arbeiten, kaum möglich.<sup>35</sup> Bei der Arbeit mit juristischen Texten in juristischen Netzwerkstudien (wie auch allgemein in empirischen juristischen Studien) können mindestens drei Typen von Fehlern zum Problem werden, die auch gemeinsam auftreten können: Abdeckungsfehler (a), Aufzeichnungsfehler (b) und Aufbereitungsfehler (c).

### a) Abdeckungsfehler

Ein Abdeckungsfehler liegt vor, wenn die gesammelten Daten nicht dasjenige Phänomen erfassen, das der Untersuchende zu erfassen glaubt. Hat ein Forscher beispielsweise zum Ziel, die Veränderung des Bürgerlichen Gesetzbuchs (BGB) über die Zeit nachzuzeichnen, so wird er dazu alle Änderungen am Normtext sammeln wollen (z.B. in Form der jeweils einschlägigen Änderungsgesetze oder in Form der jeweils geänderten Fassungen). Ein Abdeckungsfehler ist dann gegeben, wenn der Forscher seiner Studie einen Datensatz zugrunde legt, der nicht alle Änderungen enthält.

Die Diskrepanz zwischen tatsächlicher und angenommener Abdeckung kann verschiedene Gründe haben. Die Gründe können in der Sphäre der Datenquelle liegen, etwa wenn diese vorgibt, alle Änderungen eines Gesetzes nachzuweisen, die Nachweise aber unvollständig sind. Sie können aber auch in der Sphäre des Forschers liegen, etwa wenn der Forscher seine Daten aus Suchmaschinenergeb-

---

<sup>35</sup> 100 % Prozent Genauigkeit sind in der Regel auch nicht (bzw. erst recht nicht) durch händische Datenauswertung zu erreichen. Das Unbehagen, das manche Juristen mit computergestützten Verfahren verbinden, rührt daher, dass man oft mehr über die Defizite dieser Verfahren weiß, da man ihre Arbeitsweise anhand des Codes überprüfen kann und nicht auf die Sorgfalt einzelner Personen vertrauen muss.

nissen zusammenstellt, die auf Suchbegriffen basieren, welche das interessierende Phänomen zu weit oder zu eng erfassen.<sup>36</sup>

Je genauer ein Untersuchungsgegenstand durch Regeln beschrieben werden kann und je kleiner die Datenmenge, desto leichter lassen sich Abdeckungsfehler vermeiden; je komplizierter die regelhafte Beschreibung und je größer die Datenmenge, desto schwieriger lassen sie sich ausschließen. Die Daten, die juristischen Netzwerkstudien zugrunde liegen, sind zum Teil mit erheblichen Abdeckungsunsicherheiten behaftet, die nicht selten durch auslegungsbedürftige Ausdrücke (allen voran das Adjektiv „alle“) verdeckt werden – vielleicht auch, um Kritik aus den Reihen derjenigen vorzubeugen, die mit den typischen Schwierigkeiten der Datenarbeit wenig vertraut sind. In dieser Situation scheint es allerdings erst recht angebracht, die festgestellten Unsicherheiten zu betonen, um anderen Forschern zu ermöglichen, aus den erfahrenen Schwierigkeiten zu lernen.

### *b) Aufzeichnungsfehler*

Ein Aufzeichnungsfehler liegt vor, wenn der tatsächliche Inhalt eines Textes nach seinem Wortlaut nicht dem intendierten Inhalt entspricht. Aufzeichnungsfehler können als Tippfehler oder Denkfehler auftreten. Grund ist in der Regel menschliches Versagen aufseiten derjenigen, die juristische Texte produzieren (z.B. Gerichte), oder aufseiten derjenigen, die juristische Texte edieren (z.B. Mitglieder der Redaktion eines Datenbankanbieters).

Da der Forscher auf die Prozesse, die zu Aufzeichnungsfehlern führen, keinen Einfluss hat, kann er sie bei der Zusammenstellung seiner Daten nicht ausschließen, sondern nur verhindern, dass sie sich negativ auf die Qualität der späteren Analyse auswirken. Ein erster Schritt in diese Richtung besteht darin, mögliche analyserelevante Aufzeichnungsfehler mit Computerunterstützung zu identifizieren. In der juristischen Netzwerkforschung können unter anderem Fehler in Normzitierten, Rechtsprechungszierten oder auch in Kalenderdaten die Analyse verfälschen. Inwieweit sich solche Fehler erkennen lassen, hängt davon ab, wie sie sich konkret auswirken. Tippfehler können oft daran identifiziert werden, dass sie unsinnige Muster in den Daten produzieren, z.B. invalide Fundstellen oder inexistenten Kalenderdaten. Auch dass eine Entscheidung in BGHSt 31 eine Entscheidung in BGHSt 32 zitiert, ist unwahrscheinlich, insbesondere dann, wenn die zitierte Seite auch Anfangsseite einer Entscheidung in BGHSt 23 ist. Werden demgegenüber aufgrund von Denkfehlern existente, aber im konkreten Kontext unsinnige Normen zitiert, ist die Fehlererkennung deutlich schwieriger und gelingt unter Umständen erst im Rahmen des vergeblichen Versuchs, ein vermeintlich gefundenes Muster zu erklären.

---

<sup>36</sup> Ein Beispiel ist die Suche nach juristischen Aufsätzen mit quantitativen Elementen in einer Datenbank anhand des Kriteriums, dass der Wortstamm „empir“ im Aufsatztitel auftaucht. Dieses Kriterium ist zu weit, da nicht alles, was empirisch ist, auch quantitativ ist, und gleichzeitig zu eng, da der Volltext nicht einbezogen wird.

Aufzeichnungsfehler, die bei der Vorbereitung einer Analyse identifiziert werden, sollten transparent gemacht werden; in der Regel wird ihre Zahl im Verhältnis zum Datenvolumen recht klein sein. Wie mit Aufzeichnungsfehlern umzugehen ist, hängt vom Einzelfall ab. Eine Möglichkeit besteht darin, die betroffenen Strukturen aus der Analyse auszunehmen, also etwa das Zitat einer invaliden Fundstelle zu ignorieren, und diese Entscheidung wiederum offenzulegen.<sup>37</sup> Alternativ kann man die Fehler bewusst in die Analyse einschließen und eine Schätzung für die Ungenauigkeit angeben, die dadurch in die Ergebnisse der Untersuchung eingeführt wird – die Kernaussagen einer Untersuchung sollten gegenüber dieser Ungenauigkeit robust sein. Eine „Korrektur“ einzelner Fehler wird jedenfalls nur in seltenen Ausnahmefällen vorzunehmen sein.<sup>38</sup>

### c) Aufbereitungsfehler

Ein Aufbereitungsfehler liegt vor, wenn die Weiterverarbeitung der gesammelten Daten zu Verfälschungen in den resultierenden Daten geführt hat, mit denen anschließend weitergearbeitet wird. Aufbereitungsfehler sind der Sphäre des Forschers zuzuordnen. Welche Arten von Aufbereitungsfehlern bei der Vorbereitung einer juristischen Netzwerkstudie auftreten können, hängt von der verwendeten Datenquelle ab.

Die Digitalisierung von Printmedien kann zu Texterkennungsfehlern führen (Beispiel: die Zahl 1 wird als Buchstabe l erkannt). Erfasst man zum Aufbau eines Zitiernetzwerks Zitate in Texten, die digitalisiert vorliegen, so können Extraktionsfehler auftreten: Ein tatsächliches Zitat kann nicht oder falsch erfasst werden und ein anderer, ähnlich wie ein Zitat strukturierter Textbaustein kann fälschlich als Zitat erfasst werden. Die Nichterfassung von Zitaten wird insbesondere dann zum Problem, wenn die Texte, aus denen sie extrahiert werden sollen, zuvor digitalisiert wurden und Texterkennungsfehler zur Verfälschung typischer Zitiermuster geführt haben.

Das Risiko von Extraktionsfehlern lässt sich verringern, indem man die Programme, die zur Extraktion genutzt werden, anhand von Beispieltexten, deren Zitate bekannt sind, ausgiebig testet und zusätzlich eine Stichprobe der Extraktionsergebnisse manuell kontrolliert. Insbesondere dann, wenn die Zitate keinem geschlossenen Kanon fester Regeln folgen und man zur Extraktion ausschließlich regelbasierte Verfahren nutzt (also angibt, welche Struktur ein Zitat haben muss, um als solches erkannt zu werden), wird man allerdings kaum alle Zitate fehlerfrei erfassen können. Damit andere Forscher die angestellten Erwägungen nachvollziehen und auf den gesammelten Aufbereitungserfahrungen aufbauen können, sollte man den daraus hervorgehenden Unsicherheiten realer juristischer

<sup>37</sup> Dies ist das Vorgehen der Verfasserin für invalide Fundstellen im Rahmen der Fallstudie zur Rechtsprechung des BVerfG, dazu unten S. 240–248.

<sup>38</sup> Ein solcher Ausnahmefall kann etwa vorliegen, wenn man Daten aufbereitet, um sie der Öffentlichkeit zur Recherche zugänglich zu machen. In diesem Fall sollte aus den bereitgestellten Daten allerdings erkennbar sein, dass und wie sie korrigiert wurden.

Daten wiederum mit Transparenz begegnen (und auf die Entwicklung eines maschinenfreundlichen Standards zur Darstellung von Zitaten in juristischen Texten hinwirken).

## II. Fiktive Daten

Um die Methoden der juristischen Netzwerkforschung zu verstehen, ist es hilfreich, die im zweiten Kapitel skizzierten theoretischen Schwierigkeiten der Netzwerkdefinition im Blick zu behalten,<sup>39</sup> allerdings die vorstehend geschilderten Probleme der Sammlung realer Daten vorübergehend auszublenden. Das gelingt, indem man von fiktiven Daten (*mock data*) ausgeht, für die man annimmt, dass sie vollständig und fehlerfrei ein fiktives juristisches Netzwerk beschreiben, das man zuvor selbst definiert hat.

Da nicht alle Methoden, die für die juristische Netzwerkforschung relevant sind, auf jedes juristische Netzwerk passen, werden zur Veranschaulichung im Folgenden zwei Datensätze aus einem fiktiven Staat namens *Flatland* betrachtet:<sup>40</sup> ein Datensatz zu Zitaten der flatländischen Rechtsprechung (*Rechtsprechungsdaten*, 1.) und ein Datensatz zu Personalverflechtungen in Flatlands hundert größten Unternehmen (*Unternehmensdaten*, 2.).<sup>41</sup>

### 1. Datensatz 1: Rechtsprechungsdaten

Flatland hat ein quadratisches Staatsgebiet, das anhand seiner Diagonalen in vier Regionen eingeteilt ist: Norden (N), Osten (E), Süden (S) und Westen (W).<sup>42</sup> Das Gerichtssystem kennt nur eine Gerichtsbarkeit, die in drei Ebenen gegliedert ist. Auf der höchsten Ebene (*High Court*, HC) gibt es nur ein Gericht, das exakt im Zentrum von Flatland liegt (Region: H). Ihm unmittelbar untergeordnet sind vier Gerichte auf der mittleren Ebene (*Middle Courts*, MC), die jeweils für eine Region zuständig und innerhalb dieser zentral positioniert sind. In jeder Region gibt es auf der unteren Ebene (*Low Courts*, LC) drei weitere Gerichte (A, B und C), denen das für ihre Region zuständige Gericht auf der mittleren Ebene

<sup>39</sup> Zu diesen Schwierigkeiten oben S. 28–61.

<sup>40</sup> Die nachfolgende Schilderung ist inspiriert durch *Abbott*, Flatland, 1884 (veröffentlicht unter dem Pseudonym *A Square*). Es wurde bewusst ein fiktiver Staat gewählt, um keine ungerechtfertigten Annahmen über die Realität treffen zu müssen und das Szenario für jeden Leser gleich zugänglich zu machen. Und wer weiß: Vielleicht könnte die Rechtsvergleichung nicht nur von *Model Laws*, sondern auch von *Model States* profitieren?

<sup>41</sup> Die Daten wurden speziell für diese Arbeit generiert. Sie sind über den Online-Appendix abrufbar und können gemeinsam mit den entsprechend gekennzeichneten Code-Fragmenten genutzt werden, um die eingeführten Konzepte direkt anhand der Beispiele nachzuvollziehen. Im Online-Appendix findet sich auch der zur Erstellung der Daten verwendete Code, dessen Konsultation vor der Lektüre des Abschnitts zur Modellierung (S. 193–223) allerdings nicht zu empfehlen ist.

<sup>42</sup> Leser des Buches von *Abbott*, die sich Flatland bisher ohne Grenzen vorgestellt hatten, mögen der Verfasserin die Verweltlichung verzeihen.

unmittelbar übergeordnet ist. Die Gerichte auf der unteren Ebene sind nahe den Ecken ihrer Regionen positioniert und insgesamt gleichmäßig über Flatland verteilt. Abbildung 3.2 (S. 89) zeigt die geographische und hierarchische Struktur von Flatlands Gerichtslandschaft. Jedes Gericht hat zwei (nicht eingezeichnete) Kammern, die getrennt voneinander Entscheidungen treffen. Insgesamt gibt es also 34 verschiedene Spruchkörper.

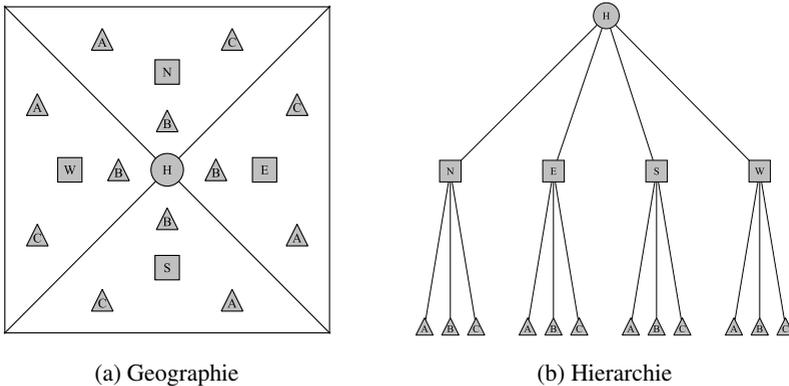


Abbildung 3.2: Gerichtslandschaft von Flatland<sup>43</sup>

Jeder Fall in Flatland endet mit genau einer Entscheidung, die vom Typ X oder vom Typ Y sein kann. Alle Entscheidungen werden begründet, veröffentlicht und nach ihrem Schwerpunkt eindeutig inhaltlich klassifiziert (Inhaltsziffern 0–9). Die Fälle – und damit auch die Entscheidungen – haben ein eindeutiges Aktenzeichen mit folgendem Format:

„Gerichtsregion-Gerichtsebene-Gerichtsbuchstabe:

Kammer-Eingangsjahr-Eingangsnummer“

Dabei ist der Gerichtsbuchstabe für die Gerichte auf höchster und mittlerer Ebene „0“ (Beispiel: H-HC-0:2-2004-173). Die Gerichte pflegen die Praxis, in ihren Entscheidungsbegründungen andere, bereits veröffentlichte Entscheidungen mit ihren Aktenzeichen zu zitieren, und das durchaus auch mehrfach. Wie sie die Entscheidungen für ihre Zitate auswählen, ob dabei also beispielsweise der Entscheidungstyp oder die Inhaltsziffer eine Rolle spielen, ist allerdings nicht bekannt.

<sup>43</sup> Die Formen sind angelehnt an Flatlands hierarchische Gesellschaftsstruktur; dazu *Abbott*, Flatland, 1884, S. 8–11 sowie unten S. 91.

Zur Illustration der Methoden der juristischen Netzwerkforschung soll davon ausgegangen werden, dass für den zehnjährigen Zeitraum von 2000 bis einschließlich 2009 fehlerfreie Zitierdaten für alle in Flatland veröffentlichten Gerichtsentscheidungen vorliegen. Zu jeder veröffentlichten Entscheidung wurden Aktenzeichen, Entscheidungsdatum, Entscheidungstyp, die aus dem Aktenzeichen ersichtlichen Informationen zum Spruchkörper und die zugeordnete Inhaltsziffer erhoben. Aus jeder Entscheidung wurden außerdem alle Zitate vollständig extrahiert und als Paare des Formats „Aktenzeichen der zitierenden Entscheidung, Aktenzeichen der zitierten Entscheidung“ in einer Liste erfasst – Mehrfachzitate äußern sich dadurch, dass mehrere gleiche Paare in der Liste auftauchen. Tabelle 3.1 (S. 90) zeigt Auszüge aus den gesammelten Daten. Insgesamt handelt es sich um 5.034 Entscheidungen mit 65.952 Zitaten.

<i>id</i>	<i>date</i>	<i>type</i>	<i>chamber</i>	<i>court</i>	<i>region</i>
S-LC-B:1-1999-195	2000-01-01	X	1	S-LC-B	S
W-LC-B:1-1999-96	2000-01-01	Y	1	W-LC-B	W
H-HC-0:1-1999-84	2000-01-03	X	1	H-HC-0	H
S-MC-0:1-1999-87	2000-01-04	Y	1	S-MC-0	S
S-LC-A:2-1999-93	2000-01-04	X	2	S-LC-A	S

<i>id</i>	<i>level</i>	<i>letter</i>	<i>year</i>	<i>month</i>	<i>day</i>	<i>content</i>
S-LC-B:1-1999-195	LC	B	2000	1	1	0
W-LC-B:1-1999-96	LC	B	2000	1	1	4
H-HC-0:1-1999-84	HC	0	2000	1	3	5
S-MC-0:1-1999-87	MC	0	2000	1	4	1
S-LC-A:2-1999-93	LC	A	2000	1	4	7

(a) Gerichtsentscheidungen (Punkte)<sup>44</sup>

<i>citing</i>	<i>cited</i>
S-LC-A:2-1999-93	H-HC-0:1-1999-84
W-LC-B:2-1999-41	H-HC-0:1-1999-84
H-HC-0:2-1999-111	H-HC-0:1-1999-84
W-MC-0:1-1999-67	H-HC-0:1-1999-84
W-MC-0:1-1999-67	S-LC-A:2-1999-93

(b) Rechtsprechungszitate (Pfeile)

Tabelle 3.1: Ausschnitt aus den für Flatland gesammelten Zitierdaten

<sup>44</sup> Die Spaltennamen der Gerichtsentscheidungen sind hier für Darstellungszwecke abgekürzt, aber in Tabelle 3.4 (S. 96) in voller Länge aufgeführt. Bei den Daten, die zu den Gerichtsentscheidungen gesammelt wurden, fällt auf, dass sowohl die Informationen zum Spruchkörper als auch die Informationen zum Entscheidungsdatum mehrfach auftauchen. Aus der Perspektive der Datenmodellierung sind solche *Redundanzen* (*redundancies*) grundsätzlich unerwünscht, da sie

Zitate haben eine natürliche Richtung (vom Zitierenden zum Zitierten) und einzelne Entscheidungen können andere Entscheidungen mehrfach zitieren. Es bietet sich daher an, die Daten als gerichteten Multigraphen mathematisch zu repräsentieren.<sup>45</sup> Dieser Graph wird nachfolgend als *Zitiergraph* bezeichnet; er hat 5.034 Punkte (einen Punkt für jede Entscheidung) und 65.952 Pfeile (einen Pfeil für jedes Zitat). Das ihm zugrundeliegende Netzwerk (mit Entscheidungen als Objekten und Zitaten als Beziehungen) wird *Zitiernetzwerk* genannt.

## 2. Datensatz 2: Unternehmensdaten

Flatlands Staatsvolk besteht aus zweidimensionalen geometrischen Figuren, die auch wirtschaftlich tätig sind.<sup>46</sup> Die Unternehmenslandschaft in Flatland ist streng reguliert. Jedes Unternehmen wird einer von fünf großen Industrien zugeordnet (Industrieziffern 0–4) und muss seine Tätigkeiten auf diesen Bereich beschränken. Unternehmen, die am Kapitalmarkt Geld aufnehmen wollen, müssen an der Staatsbörse (*FlatEx*) notiert sein. Das Management wird von einem Kontrollgremium (*Board*) überwacht, das sich aus zwei bis sechs Führungsfiguren zusammensetzt und dessen Größe per Dekret nach Faktoren bestimmt wird, die nicht öffentlich sind. Als Führungsfiguren kommen nur Figuren mit regulärer Form (alle Innenwinkel gleichgroß und alle Seiten gleichlang) in Betracht, die mindestens vier Ecken aufweisen;<sup>47</sup> Figuren, die so viele Ecken haben, dass man sie äußerlich von Kreisen nicht unterscheiden kann, werden als Kreise bezeichnet. Die Tätigkeit als Führungsfigur in einem der hundert größten Unternehmen des Landes setzt zudem eine staatliche Lizenz voraus. Die Anzahl der Lizenzen ist zur Qualitätssicherung jederzeit auf exakt eintausend Figuren beschränkt. Führungsfiguren, die lizenziert sind, können mehreren Kontrollgremien gleichzeitig angehören; nach welchen Kriterien die Unternehmen aus dem Kandidatenpool die Führungsfiguren für ihre Kontrollgremien auswählen, ist nicht bekannt.

Zu Illustrationszwecken soll angenommen werden, dass zum 31.12.2015 (Stichtag) umfassende Daten zu den 100 größten Unternehmen und den 1.000 lizenzierten Führungsfiguren in Flatland gesammelt wurden. Für jedes Unternehmen wurden Industrie, Börsennotierung und Größe des Kontrollgremiums festgehalten; zu jeder Führungsfigur wurde die Form registriert. Die Daten wurden anonymi-

---

nicht nur Speicherplatz verschwenden, sondern auch das Risiko erhöhen, dass Inkonsistenzen im Datenbestand auftreten (etwa, wenn Werte geändert und dabei einzelne redundante Spalten vergessen werden). Redundanzen können aber auch gezielt eingesetzt werden, um die Arbeit mit den Daten zu erleichtern. So verhält es sich hier: Man kann nun beispielsweise direkt auf das Entscheidungsjahr zugreifen (indem man die entsprechende Spalte abfragt), anstatt das Jahr jedes Mal aus dem Entscheidungsdatum zu extrahieren, und so für einzelne Jahre schnell Statistiken generieren oder Entscheidungen selektieren.

<sup>45</sup> Zu den verschiedenen Beziehungsformen und den Möglichkeiten ihrer Repräsentation als Linien in Graphen bereits S. 37–55.

<sup>46</sup> Das Staatsvolk wird ausführlich beschrieben bei *Abbott*, Flatland, 1884, S. 8–11.

<sup>47</sup> Die Gründe hierfür liegen in der Gesellschaftsordnung von Flatland; dazu *Abbott*, Flatland, 1884, S. 8.

siert, indem jedem Unternehmen ein eindeutiges Kürzel aus einem Großbuchstaben und einer Zahl zugeordnet wurde; jede Führungsfigur bekam eine laufende Nummer. Außerdem wurden alle Affiliationen, die aus den Mitgliedschaften von Führungsfiguren in den Kontrollgremien einzelner Unternehmen hervorgehen, in eine Liste aufgenommen (Mitgliedschaftsbeziehungen). Tabelle 3.2 (S. 92) zeigt Ausschnitte aus den erhobenen Daten.

<i>company_id</i>	<i>company_industry</i>	<i>company_listed</i>	<i>company_boardsize</i>
A0	4	0	5
B1	2	0	3
C2	0	1	6
D3	4	0	3
E4	4	0	2

(a) Unternehmen (Punkte)

<i>figure_id</i>	<i>figure_shape</i>	<i>company_id</i>	<i>figure_id</i>
0	Circle	A0	168
1	Square	A0	756
2	Circle	A0	841
3	Octagon	A0	463
4	Circle	A0	192

(b) Führungsfiguren (Punkte)

(c) Affiliationen (Linien)

Tabelle 3.2: Ausschnitt aus den für Flatland gesammelten Unternehmensdaten

Die Mitgliedschaft im Kontrollgremium hat als Beziehung zwischen einem Unternehmen und einer Führungsfigur keine natürliche Richtung. Affiliationen bestehen außerdem nur zwischen Objekten aus der Menge der Unternehmen und Objekten aus der Menge der Figures. Die Daten können mathematisch also als ungerichteter, bipartiter Graph (ohne Mehrfachkanten) repräsentiert werden, der nachfolgend *Verflechtungsgraph* genannt wird.<sup>48</sup> Der Verflechtungsgraph besteht aus 1.100 Punkten (ein Punkt für jedes der 100 Unternehmen und jede der 1.000 Führungsfiguren) und 408 Linien (eine Linie für jede Mitgliedschaft einer Führungsfigur in einem Kontrollgremium). Das ihm zugrundeliegende Netzwerk (zusammengesetzt aus den Unternehmen und Führungsfiguren als Objekte und den Gremienmitgliedschaften als Beziehungen) wird als *Verflechtungsnetzwerk* bezeichnet.

Aus dem Verflechtungsgraphen lassen sich mithilfe von Projektionen zwei weitere Graphen ableiten, in denen nur Unternehmen bzw. nur Führungsfiguren

<sup>48</sup> Ein *bipartiter Graph* ist ein Graph  $G = (V, E)$ , dessen Punktmenge sich so in zwei Teilmengen aufteilen lässt, dass sämtliche Linien von einer Menge in die andere Menge verlaufen, d.h. für den gilt:

$\exists V_1, V_2 \subseteq V$  mit  $V_1 \cup V_2 = V$  und  $V_1 \cap V_2 = \emptyset$ , sodass  $\forall \{u, w\} \in E$  gilt:  $u \in V_1 \Leftrightarrow w \in V_2$ .

als Punkte enthalten sind.<sup>49</sup> Sie werden *Unternehmensgraph* und *Personalgraph* genannt; die ihnen korrespondierenden Netzwerke werden als *Unternehmensnetzwerk* und *Personalnetzwerk* bezeichnet. Der Unternehmensgraph hat 100 Punkte (einen für jedes Unternehmen) und 557 Linien (einen für jeden Fall, in dem eine Führungsfigur in den Kontrollgremien zweier Unternehmen sitzt); der Personalgraph hat 1.000 Punkte (einen für jede Führungsfigur) und 724 Linien (eine für jeden Fall, in dem zwei Führungsfiguren gemeinsam in einem Kontrollgremium sitzen) – vorausgesetzt, man übersetzt mehrere gemeinsame Nachbarn zweier Punkte bei der jeweiligen Projektion in mehrere Linien zwischen diesen Punkten. Lässt man keine Mehrfachkanten zu – erfasst man also nur, ob zwei Unternehmen mindestens eine Führungsfigur teilen oder zwei Führungsfiguren gemeinsam in mindestens einem Kontrollgremium sitzen – oder übersetzt man mehrere gemeinsame Nachbarn in gewichtete Kanten, so hat der Unternehmensgraph 537 Linien (ggf. mit Gewichten), der Personalgraph 705.<sup>50</sup>

### III. Beschreibung von Daten

Ohne besondere Erwähnung sind auf den im vorigen Abschnitt beschriebenen, fiktiven Daten bereits einfache mathematische Operationen durchgeführt worden: Die Anzahl der Punkte in Zitiergraphen bzw. Verflechtungsgraphen entspricht der Anzahl der Entscheidungen bzw. Unternehmen und Führungsfiguren, die bei der Datensammlung für Zitiernetzwerk bzw. Verflechtungsnetzwerk identifiziert wurden; die Anzahl der Pfeile bzw. Linien entspricht der Anzahl der registrierten Zitate bzw. Gremienmitgliedschaften. Stellt man sich die gesammelten Daten als Tabellen vor, so stellt jede Zeile einer Tabelle eine *Beobachtung* (*observation*) dar, also eine Entscheidung, ein Zitat, ein Unternehmen, eine Führungsfigur oder eine Mitgliedschaftsbeziehung.<sup>51</sup> Die Anzahl der Punkte, Pfeile oder Linien der betrachteten Graphen lässt sich dann ermitteln, indem man die Zeilen der einschlägigen Tabelle zählt.

<sup>49</sup> Bei der *Projektion* eines bipartiten Graphen auf eine seiner zwei *partes* ( $V_1$  und  $V_2$ , sodass die Bedingung aus Fn. 48 [S. 92] erfüllt ist) wird zwischen zwei Punkten immer dann eine Linie angenommen, wenn sie im bipartiten Graphen mit demselben dritten Punkt in Verbindung stehen.

<sup>50</sup> Aus einem *bipartiten* Graphen *ohne* Mehrfachkanten werden hier durch die Projektionen also zwei in der Regel *nicht* bipartite Graphen *mit* oder *ohne* Mehrfachkanten und – falls man sich gegen Mehrfachkanten entscheidet – *mit* oder *ohne* gewichtete Kanten.

<sup>51</sup> Abstrakt sind die einzelnen Beobachtungen die Objekte (z.B. Gerichtsentscheidungen), an denen die interessierenden Größen (z.B. Spruchkörper) im Rahmen einer statistischen Untersuchung tatsächlich erfasst werden. Die Beobachtungen sind in der Regel nur ein Teil aller *Untersuchungseinheiten* (auch genannt *statistische Einheiten*), die für eine bestimmte Fragestellung relevant sind. Die Menge dieser Untersuchungseinheiten wird *Grundgesamtheit* (*population*) genannt; *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 1; *Fahrmeir u.a.*, Statistik, 2016, S. 13–14; *Moore, McCabe und Craig*, Introduction to the Practice of Statistics, 2017, S. 189.

Welche weiteren Rechenoperationen auf den Daten (mathematisch) zulässig sind, hängt davon ab, welche Arten von Informationen in den einzelnen Feldern der Tabelle enthalten sind. Jede Spalte der Tabelle steht dabei für ein abstraktes *Merkmal*, dessen konkreter Wert, genannt *Ausprägung*, für jede Beobachtung separat erfasst wird.<sup>52</sup> Jedes Merkmal hat eine *Skala*, auf der seine Ausprägungen gemessen werden. Die Skala eines Merkmals bestimmt, durch welche mathematischen Operationen sich seine Ausprägungen für mehrere Beobachtungen miteinander verknüpfen lassen und welche zusammenfassenden Kennzahlen zur Beschreibung der Gesamtheit aller Beobachtungen ermittelt werden können. Dabei wird klassischerweise zwischen Nominal-, Ordinal-, Intervall- und Verhältnisskala unterschieden (Tab. 3.3, S. 94).<sup>53</sup> Intervall- und Verhältnisskala werden oft zur *Kardinalskala* zusammengefasst; man spricht dann von *nominalen*, *ordinalen* und *metrischen* (d.h. kardinalskalierten) Merkmalen.<sup>54</sup>

<i>Skala</i>	<i>Op.</i>	<i>Deskriptive Statistiken u.a.</i>	<i>Juristisches Beispiel</i>
Nominal	=	Anzahl Ausprägungen	Gerichtsbarkeiten
	≠	Modus <sup>55</sup>	(z.B. Arbeitsger., Sozialger.)
Ordinal	<	Median <sup>56</sup>	Gerichtsstufen
	>	Quantile <sup>57</sup>	(z.B. AG, LG, OLG, BGH)
Intervall	–	Arithmetisches Mittel <sup>58</sup>	Verfahrenstermine
	+	Standardabweichung <sup>59</sup>	(z.B. 27.01.2010, 07.07.2010)
Verhältnis	*		Verfahrensdauern
	/		(z.B. 2 Monate, 2 Jahre)

Tabelle 3.3: Skalen von Merkmalen

<sup>52</sup> Toutenburg und Heumann, Deskriptive Statistik, 2009, S. 2; Fahrmeir u.a., Statistik, 2016, S. 13–14. Für Fälle, in denen die Ausprägung nicht festgestellt werden kann, sieht man in der Regel einen symbolischen Wert vor, der das Nichtvorhandensein eines tatsächlichen Werts signalisiert; in Datenbanken heißt dieser Wert für gewöhnlich NULL.

<sup>53</sup> Grundlegend Stevens, Science 103 (1946), S. 677–680; aus der jüngeren Lehrbuchliteratur Toutenburg und Heumann, Deskriptive Statistik, 2009, S. 6; Fahrmeir u.a., Statistik, 2016, S. 15–16; die Tabelle ist angelehnt an Stevens, Science 103 (1946), S. 677 (678). Die Nominalskala ist die niedrigste, die Verhältnisskala die höchste Skala. Auf einer höheren Skala sind immer auch alle Basisoperationen und Statistiken zulässig, die auf niedrigeren Skalen zulässig sind; dazu Fahrmeir u.a., Statistik, 2016, S. 17. Die Terminologie in der Literatur ist nicht ganz einheitlich, beispielsweise werden nominale Merkmale auch *kategorial* genannt (Ware, Information Visualization, 2013, S. 27) oder kategoriale Merkmale als Unterfall nominaler Merkmale begriffen (Ward, Grinstein und Keim, Interactive Data Visualization, 2015, S. 52).

<sup>54</sup> Toutenburg und Heumann, Deskriptive Statistik, 2009, S. 6–7; Fahrmeir u.a., Statistik, 2016, S. 16. Zur Skalierung der für Rechtsprechung und Unternehmen in Flatland erhobenen Merkmale unten Tab. 3.4–3.5 (S. 96–97).

Mit der Unterscheidung von nominalen, ordinalen und metrischen Merkmalen verwandt ist die Differenzierung zwischen *qualitativen* Merkmalen und *quantitativen* Merkmalen. Hier geht es darum, ob die Ausprägung eines Merkmals eine Eigenschaft signalisiert (qualitative Merkmale) oder ein Ausmaß angibt (quantitative Merkmale).<sup>60</sup> Daher werden nominale Merkmale als qualitativ und metrische Merkmale als quantitativ eingeordnet, während bei ordinalen Merkmalen die Einordnung davon abhängt, ob inhaltlich die Anordnungs-komponente (quantitativ) oder die Kategorisierungskomponente (qualitativ) überwiegt. Letzteres wird insbesondere dann der Fall sein, wenn ein ordinales Merkmal nur wenige verschiedene Ausprägungen annehmen kann.<sup>61</sup>

Schließlich kann man Merkmale auch generell nach der Anzahl ihrer möglichen Ausprägungen unterscheiden: Merkmale, die nur endlich viele oder abzählbar unendlich viele verschiedene Ausprägungen annehmen können, werden als *diskret*

<sup>55</sup> Der *Modus* ist die am häufigsten vorkommende Ausprägung, er ist nicht immer eindeutig; *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 50; *Finkelstein und Levin*, Statistics for Lawyers, 2015, S. 4; *Fahrmeir u.a.*, Statistik, 2016, S. 53.

<sup>56</sup> Der *Median* ist ein Wert, für den gilt, dass 50 % der beobachteten Ausprägungen kleiner oder gleich groß sind und 50 % der beobachteten Ausprägungen größer oder gleich groß sind. Für eine ungerade Anzahl von Beobachtungen ist der Median der Wert, der in der Mitte der nach der Größe sortierten Beobachtungswerte liegt. Für eine gerade Anzahl von Beobachtungen ist der Median das arithmetische Mittel der zwei Werte, die in der Mitte der nach der Größe sortierten Beobachtungswerte liegen, sofern das betrachtete Merkmal metrisch skaliert ist; anderenfalls ist er nicht eindeutig bestimmt. *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 52–53; *Finkelstein und Levin*, Statistics for Lawyers, 2015, S. 4; *Fahrmeir u.a.*, Statistik, 2016, S. 53.

<sup>57</sup> Ein *p-Quantil* ist jeder Wert  $x_p$ , für den ein Anteil  $p$  aller beobachteten Ausprägungen kleiner oder gleich und ein Anteil  $1 - p$  der beobachteten Ausprägungen größer oder gleich ist; das 0,25-Quantil, 0,50-Quantil (Median) und das 0,75-Quantil werden auch als *Quartile* bezeichnet; *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 55–57; *Fahrmeir u.a.*, Statistik, 2016, S. 60.

<sup>58</sup> Das *arithmetische Mittel* ist die Summe aller Ausprägungen, geteilt durch die Anzahl der Ausprägungen; *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 59–60; *Fahrmeir u.a.*, Statistik, 2016, S. 50; *Moore, McCabe und Craig*, Introduction to the Practice of Statistics, 2017, S. 28. Es ist der Wert, der typischerweise gemeint ist, wenn man vom *Durchschnitt* spricht.

<sup>59</sup> Während Modus, Median, Quantile und arithmetisches Mittel zu den *Lagemaßen* gezählt werden, gehört die Standardabweichung zu den *Streuungsmaßen*; *Finkelstein und Levin*, Statistics for Lawyers, 2015, S. 1. Die Standardabweichung misst, wie stark die einzelnen Ausprägungen durchschnittlich vom arithmetischen Mittel abweichen – in der Einheit, in der auch die Ausprägungen gemessen werden. Die Standardabweichung berechnet sich als Wurzel aus der Varianz bzw. der Stichprobenvarianz; *Fahrmeir u.a.*, Statistik, 2016, S. 65 (dort auch zum Unterschied zwischen Varianz und Stichprobenvarianz). Wo nachfolgend nicht anders vermerkt, wurden die Angaben zur Standardabweichung stets aus der Stichprobenvarianz ermittelt.

<sup>60</sup> *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 4; *Fahrmeir u.a.*, Statistik, 2016, S. 17.

<sup>61</sup> *Fahrmeir u.a.*, Statistik, 2016, S. 17.

bezeichnet,<sup>62</sup> während Merkmale mit überabzählbar unendlich vielen verschiedenen Ausprägungen *stetig* genannt werden;<sup>63</sup> Merkmale, die man nur diskret, aber so fein messen kann, dass sie sich wie stetige Merkmale behandeln lassen, heißen *quasi-stetig*.<sup>64</sup>

Alle eingeführten Adjektive werden genutzt, um zwischen verschiedenen *Merkmalstypen* zu unterscheiden.<sup>65</sup> Dabei handelt es sich streng genommen um drei verschiedene Typisierungen: Die Differenzierung zwischen nominalen, ordinalen und metrischen Merkmalen zielt auf die *Skala*, jene zwischen qualitativen und quantitativen Merkmalen auf die *Bedeutung* und jene zwischen diskreten und stetigen Merkmalen auf die *Erfassung* eines Merkmals. Während die Typisierung nach der Skala sowohl bei der Visualisierung als auch bei der Quantifizierung eine Rolle spielt, wird die Typisierung nach der Erfassung primär im Rahmen der Quantifizierung und jene nach der Bedeutung vor allem im Rahmen der Visualisierung relevant. Tabelle 3.4 (S. 96) und Tabelle 3.5 (S. 97) fassen die Typen der Merkmale zusammen, die für die Objekte aus den Rechtsprechungsdaten und den Unternehmensdaten in Flatland erhoben wurden.

<i>Merkmal</i>	<i>Skala</i>	<i>Bedeutung</i>	<i>Erfassung</i>
case_id	nominal	qualitativ	diskret
dec_date <sup>66</sup>	metrisch	quantitativ	diskret
dec_type	nominal	qualitativ	diskret
chamber	nominal	qualitativ	diskret
court	nominal	qualitativ	diskret
court_region	nominal	qualitativ	diskret
court_level <sup>67</sup>	ordinal	quantitativ	diskret
court_letter	nominal	qualitativ	diskret
dec_content	nominal	qualitativ	diskret

Tabelle 3.4: Merkmalstypen zu den Rechtsprechungsdaten (ohne Redundanzen)

<sup>62</sup> Die Anzahl möglicher Ausprägungen eines Merkmals ist dann *abzählbar unendlich*, wenn sie nicht endlich ist, man aber die einzelnen Ausprägungen mit den natürlichen Zahlen indexiert hintereinander schreiben könnte: „1. Erste Ausprägung, 2. Zweite Ausprägung, ...“.

<sup>63</sup> Die Anzahl möglicher Ausprägungen eines Merkmals ist dann *überabzählbar unendlich*, wenn sie nicht endlich ist und die natürlichen Zahlen nicht ausreichen, um alle möglichen Ausprägungen zu indexieren. Das kommt vor allem dann vor, wenn ein Merkmal jeden beliebigen Wert in einem Intervall der reellen Zahlen annehmen kann.

<sup>64</sup> *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 4–5; *Fahrmeir u.a.*, Statistik, 2016, S. 15.

<sup>65</sup> *Toutenburg und Heumann*, Deskriptive Statistik, 2009, S. 4–7; *Fahrmeir u.a.*, Statistik, 2016, S. 14, 18.

<sup>66</sup> Hätte man statt des Entscheidungsdatums den Entscheidungszeitpunkt erhoben, so wäre das Merkmal hinsichtlich der Erfassung als quasi-stetig einzuordnen.

<sup>67</sup> Wengleich dieses Merkmal nur drei verschiedene Ausprägungen annehmen kann, wird es hier als quantitativ eingeordnet, da die Kategorisierungskomponente mit dem Merkmal

<i>Merkmal</i>	<i>Skala</i>	<i>Bedeutung</i>	<i>Erfassung</i>
company_id	nominal	qualitativ	diskret
company_industry	nominal	qualitativ	diskret
company_listed <sup>68</sup>	ordinal	qualitativ	diskret
company_boardsize	metrisch	quantitativ	diskret

(a) Unternehmen

<i>Merkmal</i>	<i>Skala</i>	<i>Bedeutung</i>	<i>Erfassung</i>
figure_id	nominal	qualitativ	diskret
figure_shape <sup>69</sup>	ordinal	quantitativ	diskret

(b) Führungsfiguren

Tabelle 3.5: Merkmalstypen zu den Unternehmensdaten

Aus den Ausprägungen der erhobenen Objektmerkmale lassen sich zahlreiche zusammenfassende Statistiken ableiten (etwa anhand von Tab. 3.3, S. 94). So kann man beispielsweise die Informationen zur Größe der Boards in Flatlands 100 größten Unternehmen und zu den Formen von Flatlands 1.000 Führungsfiguren wie in Tabelle 3.6 (S. 97) dargestellt komprimieren.

<i>Statistik</i>	<i>Wert</i>	<i>Statistik</i>	<i>Wert</i>
Minimum	2	Minimum	Quadrat
Maximum	6	Maximum	Kreis
Modus	5	Modus	Kreis
0,25-Quantil	3	0,25-Quantil	Hexagon
0,50-Quantil (Median)	4	0,50-Quantil (Median)	Oktagon
0,75-Quantil	5	0,75-Quantil	Kreis
Arithmetisches Mittel	4,08	Arithmetisches Mittel	n/a
Standardabweichung	1,39	Standardabweichung	n/a

(a) Größe der Boards

(b) Form der Führungsfiguren

Tabelle 3.6: Zusammenfassende Statistiken zu zwei Merkmalen der Unternehmensdaten

In ähnlicher Weise können auch aus den Beziehungsdaten zusammenfassende Statistiken generiert werden. Etwa kann man für jedes Aktenzeichen in den Recht-

*court* vollständig erfasst wird und die Anordnungskomponente damit der einzige Grund ist, das Merkmal *court\_Level* überhaupt separat zu erheben.

<sup>68</sup> Eine Besonderheit ist, dass dieses Merkmal nur zwei Werte annehmen kann: 0 (für *nicht gelistet*) oder 1 (für *gelistet*). Solche Variablen nennt man *binär* oder auch *dichotom*; Fahrmeir u.a., Statistik, 2016, S. 49.

<sup>69</sup> Die Einordnung dieses Merkmals als ordinal, quantitativ und diskret rührt daher, dass die Gesellschaft Flatlands nach der Anzahl der Ecken hierarchisch geordnet ist und die Anzahl der Ecken (jedenfalls theoretisch) eine beliebige natürliche Zahl sein kann.

sprechungsdaten zählen, wie oft es in der Spalte *citing* bzw. in der Spalte *cited* vorkommt. Die resultierenden Zählzahlen lassen sich als zusätzliche Merkmale der einzelnen Entscheidungen auffassen, die metrisch skaliert (Skala), quantitativ (Bedeutung) und diskret (Erfassung) sind. So ergeben sich die zusammenfassenden Statistiken in Tabelle 3.7 (S. 98).

<i>Statistik</i>	<i>Wert</i>	<i>Statistik</i>	<i>Wert</i>
Minimum	0	Minimum	0
Maximum	71	Maximum	650
Modus	3	Modus	0
0,25-Quantil	4	0,25-Quantil	0
0,50-Quantil (Median)	10	0,50-Quantil (Median)	0
0,75-Quantil	19	0,75-Quantil	3
Arithmetisches Mittel	13,10	Arithmetisches Mittel	13,10
Standardabweichung	10,82	Standardabweichung	41,37

(a) Ausgehende Zitate (*citing*)

(b) Eingehende Zitate (*cited*)

Tabelle 3.7: Zusammenfassende Statistiken zur Anzahl der aus- und eingehenden Zitate

Einige der in Tabelle 3.7 (b) (S. 98) festgehaltenen Zahlen könnten auch Teil einer klassischen Zitationsanalyse sein.<sup>70</sup> Eine Gesamtschau dieser Zahlen lässt aber bereits vermuten, dass sie die zugrunde liegenden Daten wenig präzise beschreiben: Die Standardabweichung ist mehr als dreimal so hoch wie das arithmetische Mittel und dieses liegt wiederum weit jenseits von Modus und Median.<sup>71</sup> Würde man eine Entscheidung zufällig auswählen und für sie die Anzahl der eingehenden Zitate betrachten, könnte man kaum einen Wert in der Nähe des arithmetischen Mittels erwarten – ganz anders, als wenn man einen Menschen zufällig auswählen und dessen Körpergröße messen würde. Ähnliche Konstellationen kommen in Daten, mit denen die juristischen Netzwerkforschung arbeitet, relativ oft vor und führen dazu, dass neben den zusammenfassenden Statistiken zu den Ausprägungen eines Merkmals immer auch dessen *Häufigkeitsverteilung* (*frequency distribution*) betrachtet wird. Eine Häufigkeitsverteilung gibt entweder die Anzahl der Beobachtungen mit einer bestimmten Merkmalsausprägung (*absolute Häufigkeitsverteilung*) oder den Anteil der Vorkommnisse dieser Merkmalsausprägung an der Gesamtheit aller beobachteten Ausprägungen (*relative Häufigkeitsverteilung*) an.<sup>72</sup> Viele Häufigkeitsverteilungen, die in der juristische Netzwerkforschung eine Rolle spielen, unterscheiden sich deutlich von den um einen besonders wahrscheinlichen Wert herum symmetrisch abfallenden Häufigkeitsverteilungen, die in anderen Zusammenhängen häufig auftreten (z.B. bei der menschlichen Körpergröße). Das

<sup>70</sup> Zu Beispielen für solche Zitationsanalysen bereits Kapitel 2 Fn. 8 (S. 14).

<sup>71</sup> Zu Modus, Median, arithmetischem Mittel und Standardabweichung bereits oben Tab. 3.3 (S. 94) mit begleitenden Fußnoten.

<sup>72</sup> *Fahrmeir u.a.*, Statistik, 2016, S. 30.

bringt für ihre Analyse einige Besonderheiten mit sich, auf die im Rahmen von Visualisierung und Quantifizierung zurückzukommen sein wird.<sup>73</sup>

## B. Visualisierung

Visualisierung bezeichnet die Kommunikation von Informationen mittels graphischer Repräsentationen.<sup>74</sup> Damit kann sowohl der Prozess als auch das Medium der Kommunikation gemeint sein, je nachdem, ob der Ausdruck als abstraktes oder als konkretes Substantiv verwendet wird. Sei es als Balkendiagramm (*bar chart*) zur Darstellung von Wahlergebnissen, als Liniendiagramm (*line chart*) zur Kommunikation von Aktienkursentwicklungen oder als Karte (*map*) zur Orientierung in unbekannter Umgebung: Visualisierungen sind in der Lage, ihren Adressaten viele, möglicherweise komplexe Informationen mitsamt ihrer Zusammenhänge buchstäblich vor Augen zu führen. Dabei kommt ihnen zugute, dass sie mit dem Sehsinn den leistungsstärksten Sinn des Menschen ansprechen.<sup>75</sup>

Die Netzwerkforschung ist seit jeher eng mit der Visualisierung verbunden – man denke nur an *Eulers* Skizze aus dem ersten Kapitel (Abb. 1.2, S. 2) oder *Morenos* Soziogramm aus dem zweiten Kapitel (Abb. 2.6, S. 66). Im Recht sind visuelle Darstellungen, abgesehen von Klassikern aus dem Schuld- und Sachenrecht oder dem Staatsorganisationsrecht sowie den Skripten einiger Repetitoren, dagegen (jedenfalls heute) wenig verbreitet.<sup>76</sup> Vor diesem Hintergrund lohnt sich zunächst ein Blick auf die verschiedenen Funktionen, die Visualisierungen erfüllen

<sup>73</sup> Dazu unten S. 99–122 und S. 122–193.

<sup>74</sup> *Ward, Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 1, ausführlich zu den Grundlagen der Visualisierung auf S. 139–181 (Kapitel 4). Umfassende Leitlinien für die Produktion von Visualisierungen formuliert *Ware*, *Information Visualization*, 2013, zusammengestellt auf S. 445–457 (Appendix D). Zu den Klassikern der Visualisierungsliteratur gehören *Bertin*, *Sémiologie Graphique*, 2005 (Erstauflage 1967), *Tufte*, *The Visual Display of Quantitative Information*, 2001 (Erstauflage 1983), *Cleveland*, *The Elements of Graphing Data*, 1994 (Erstauflage 1985), und *Cleveland*, *Visualizing Data*, 1993.

<sup>75</sup> *Ware*, *Information Visualization*, 2013, S. 2.

<sup>76</sup> Man kann sich darüber streiten, ob Photographien oder Zeichnungen lebensweltlicher Situationen unter die oben angegebene Definition von Visualisierung fallen. Solche Bilder, wie sie insbesondere in der Rechtsikonographie zu finden sind, werden im Folgenden nicht näher behandelt, da sie für die juristische Netzwerkforschung keine Rolle spielen. Zum Verhältnis von Recht und Bildern allgemein *Boehme-Neßler*, *BilderRecht*, 2009, S. 109–162 m.w.N.; speziell zur historischen Entwicklung *Röhl*, *Bilder in gedruckten Rechtsbüchern*, in: *Lerch* (Hrsg.), *Die Sprache des Rechts*, 2005, S. 267–348; eine Auswahl von Bildern in juristischen Büchern aus der *Yale Law Library Collection* zeigen *Widener und Weiner*, *Law's Picture Books*, 2017. Eine auf die juristische Ausbildungsliteratur konzentrierte Bestandsaufnahme aus jüngerer Zeit findet sich bei *Bergmans*, *Visualisierungen in Rechtslehre und Rechtswissenschaft*, 2009, der die Definition des Begriffs „Visualisierung“ ebenfalls problematisiert und ihn letztlich sehr viel weiter auslegt als hier (S. 2–4, 17–24). Zu den Visualisierungen im hiesigen Sinne zählen nach der dortigen Terminologie vor allem Diagramme (auf S. 24 eingeordnet unter „Sonstige“) und Strukturbilder.

können (I.). Anschließend wird der Prozess der Erstellung von Visualisierungen aus Daten in seinen Grundzügen skizziert (II.), bevor das Augenmerk speziell auf die Visualisierung von Graphen gerichtet wird (III.). Die Erläuterungen sind darauf ausgelegt, einen Eindruck von der Rolle der Visualisierung in der juristischen Netzwerkforschung zu vermitteln und aufzuzeigen, wo die juristische Netzwerkforschung von einer stärkeren Einbeziehung der bereichsspezifischen Forschungsergebnisse profitieren könnte.

### I. Funktionen von Visualisierung

Visualisierungen können sowohl der Darstellung von Daten (*data visualization*) als auch der Darstellung von Konzepten (*concept visualization*) dienen. In beiden Fällen können sie verschiedene Ziele verfolgen, nach denen sie sich grob in zwei Klassen aufteilen lassen: *analytische* Visualisierungen und *illustrative* Visualisierungen.<sup>77</sup> Analytische Visualisierungen sollen dem Betrachter dabei helfen, neue Erkenntnisse zu gewinnen. Demgegenüber zielen illustrative Visualisierungen darauf ab, dem Betrachter Erkenntnisse zu vermitteln, die bereits an anderer Stelle gewonnen wurden.

Mit der Differenzierung zwischen analytischen und illustrativen Visualisierungen geht eine Unterscheidung nach primären Adressatenkreisen einher: Während analytische Visualisierungen sich vor allem an den Forscher selbst richten, sprechen illustrative Visualisierungen vorwiegend ein externes Publikum an. Damit stehen analytische Visualisierungen in der Regel am Anfang oder in der Mitte, illustrative Visualisierungen hingegen am Ende eines Forschungsprozesses – und die Informationsasymmetrie zwischen Autor und Adressat ist bei illustrativen Visualisierungen deutlich größer als bei analytischen Visualisierungen. Zur optimalen Ausfüllung der unterschiedlichen Visualisierungsfunktionen sind daher in der Regel auch verschiedene Visualisierungen erforderlich.<sup>78</sup>

Visualisierungen von Konzepten und Visualisierungen von Daten spielen in verschiedenen Phasen des Forschungsprozesses der juristischen Netzwerkforschung unterschiedliche Rollen. *Am Anfang* des Forschungsprozesses helfen analytische Visualisierungen von Konzepten dabei, den Untersuchungsgegenstand zu konkretisieren und in seinen Kontext zu setzen. So lässt sich etwa verdeutlichen, welche Beziehungen zwischen juristischen Texten und juristischen Akteuren be-

<sup>77</sup> Die Unterscheidung orientiert sich an jener bei *Tukey*, *Some Graphic and Semigraphic Displays*, in: *Bancroft* (Hrsg.), *Statistical Papers in Honor of George W. Snedecor*, 1972, S. 293 (293), die allerdings auf die Visualisierung von Daten zugeschnitten ist. Die dort beschriebene Tabellenersatzfunktion von Visualisierungen („Graphs from which numbers are to be read off“) wird hier als Unterfall der illustrativen Visualisierung von Daten begriffen. Eine feingliedrige Systematisierung einzelner Visualisierungsfunktionen findet sich bei *Munzner*, *Visualization Analysis & Design*, 2015, S. 42–65.

<sup>78</sup> Ebenso *Tukey*, *Some Graphic and Semigraphic Displays*, in: *Bancroft* (Hrsg.), *Statistical Papers in Honor of George W. Snedecor*, 1972, S. 293 (293), der zugleich feststellt, dass dieser Leitlinie nicht immer entsprochen wird.

stehen können, bevor mit der Analyse von Rechtsprechungsdaten begonnen wird (Abb. 3.3, S. 101). Die Unübersichtlichkeit des Resultats vermittelt einen Eindruck davon, welchen Ausschnitt der Realität beispielsweise eine Analyse von Rechtsprechungszitaten (nur) erfassen kann.

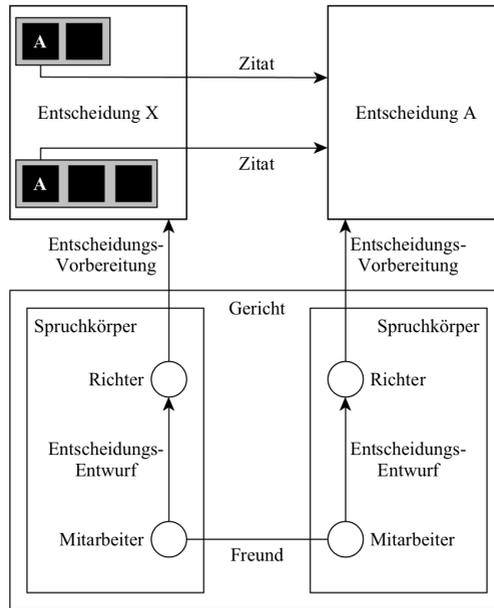


Abbildung 3.3: Rechtsprechungszitatre im Kontext<sup>79</sup>

*In der Mitte* des Forschungsprozesses werden analytische Visualisierungen von Daten eingesetzt, um den ausgewählten Untersuchungsgegenstand besser kennenzulernen und Spuren möglicher Zusammenhänge zwischen erhobenen Merkmalen zu entdecken, deren Weiterverfolgung sich lohnen könnte. Visualisierungen mit dieser Funktion werden im weiteren Verlauf der Arbeit immer wieder genutzt, um das Vorgehen der juristischen Netzwerkforschung zu veranschaulichen. *Am Ende* des Forschungsprozesses können schließlich illustrative Visualisierungen von Konzepten und illustrative Visualisierungen von Daten zum Einsatz kommen, um die Ergebnisse einer Untersuchung zu kommunizieren.<sup>80</sup>

<sup>79</sup> Die Abbildung zeigt (nur) einen Ausschnitt der möglichen Beziehungen. Hinzu kommen insbesondere Beziehungen zu anderen juristischen Texten (etwa Zitate von Normen bei den Entscheidungen und Autorenschaft von Literaturbeiträgen bei den Richtern) und Beziehungen zu anderen juristischen Akteuren (Beispiel: Richter und ihre Mitarbeiter haben Verwandte und akademische Lehrer).

<sup>80</sup> Jenseits der juristischen Netzwerkforschung im engeren Sinne könnte die illustrative Visualisierung von Konzepten auch rechtsdogmatische Diskussionen bereichern, soweit es in ihnen um

Auch die Literatur zur Netzwerkforschung im Recht enthält zahlreiche Beispiele für den Einsatz von Visualisierungen,<sup>81</sup> wobei die Mehrzahl der Darstellungen die Visualisierung von Daten betrifft.<sup>82</sup> Legt man die obige Differenzierung zugrunde, müssten diese Visualisierungen alle illustrativ sein, denn in den Publikationen geht es gerade darum, dem externen Publikum intern erzielte Erkenntnisse zu vermitteln. Mitunter scheinen allerdings auch analytische Visualisierungen unmittelbar den Weg in eine Veröffentlichung gefunden zu haben.<sup>83</sup> Außerdem sind selbst (offenbar) illustrative Visualisierungen in der Informationsvermittlung oft wenig effektiv – und mitunter tauchen auch Visualisierungen auf, die rein *dekorativ* sind.<sup>84</sup> Die Probleme solcher Visualisierungen lassen sich besser verstehen, wenn man neben den spezifischen Schwierigkeiten der Visualisierung von Graphen (dazu unten S. 109–122) auch den Prozess betrachtet, in dem Visualisierungen typischerweise entstehen (dazu sogleich).

---

Rechtsbeziehungen geht. Würden beispielsweise Autoren, die sich zur rechtlichen Behandlung von Anweisungs- oder Nebenbesitzfällen äußern, in Form von Graphen darlegen, wie sie unter Zugrundelegung verschiedener Geschehensabläufe die Rechtsbeziehungen zwischen den Beteiligten zu verschiedenen Zeitpunkten sehen, und sich dabei an eine einheitliche Notation halten, so ließen sich die einzelnen Positionen schneller miteinander vergleichen und formulierungsbedingte Missverständnisse könnten von sachlichen Differenzen leichter abgegrenzt werden. Unabhängig davon würde die Erstellung solcher Graphen die Möglichkeit eröffnen, etwaige Vagheiten und Inkohärenzen in der jeweiligen Position vor Veröffentlichung auszumerzen, sodass der illustrativen Visualisierung hier zusätzlich eine analytische Funktion zukäme.

<sup>81</sup> Zu dieser Literatur bereits S. 11–28.

<sup>82</sup> Beispiele für konzeptionelle Visualisierungen finden sich hingegen bei *Blumoff*, *Georgia State University Law Review* 24 (2008), S. 1003 (1003); *Siems*, *Arizona Journal of International & Comparative Law* 27 (2010), S. 747 (757); *Torrance und West*, *Virginia Journal of Law & Technology* 20 (2017), S. 466 (484).

<sup>83</sup> Beispiele: *Boulet, Mazzeza und Bourcier*, *Network Analysis of the French Environmental Code*, in: *Casanovas u.a.* (Hrsg.), *AI Approaches to the Complexity of Legal Systems*, 2010, S. 39 (45–48); *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457 (491–492); *Katz und Bommarito*, *Artificial Intelligence and Law* 22 (2014), S. 337 (353); *Li u.a.*, *Journal of Business and Technology Law* 10 (2015), S. 297 (336–339); *Derlén und Lindholm*, *German Law Journal* 18 (2017), S. 647 (667, 678); *Torrance und West*, *Virginia Journal of Law & Technology* 20 (2017), S. 466 (489).

<sup>84</sup> Beispiele: *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457 (459, 489–490); *Puig*, *The European Journal of International Law* 25 (2014), S. 387 (387); *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327 (338).

## II. Prozess der Visualisierung

Der Prozess, in dem aus Daten eine Visualisierung erzeugt wird, heißt Visualisierungsprozess (*visualization process*). Die Grundstruktur des Visualisierungsprozesses ist unabhängig davon, ob er mit oder ohne Computerunterstützung durchlaufen wird (Abb. 3.4, S. 103): Man transformiert die ursprünglichen Beobachtungen (*Rohdaten*) zunächst in *strukturierte Daten* (Schritt 1), um ihnen danach *visuelle Strukturen* zuzuordnen (Schritt 2) und diese schließlich in eine *Visualisierung* zu übersetzen (Schritt 3).<sup>85</sup>

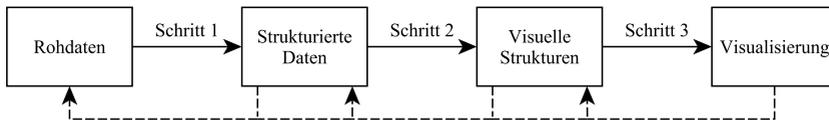


Abbildung 3.4: Grundstruktur des Visualisierungsprozesses<sup>86</sup>

Je nachdem, an welcher Stelle im Visualisierungsprozess Probleme auftreten und mit welchem Ziel visualisiert wird, können die einzelnen Schritte auch mehrfach durchlaufen werden, was in Abbildung 3.4 (S. 103) durch die gestrichelten Pfeile angedeutet ist. Außerdem entfällt der erste Schritt, sofern genau die Daten, die visualisiert werden sollen, bereits in strukturierter Form vorliegen. Das kann vorkommen, wenn man mit fiktiven Daten oder mit Daten arbeitet, die von anderen gesammelt und strukturiert wurden. Auch in diesen Fällen ist es aber unter Umständen erforderlich, zum Zwecke der Visualisierung weitere strukturierte Daten zu generieren – etwa dann, wenn man nicht jede Beobachtung einzeln darstellen, sondern zusammenfassende Statistiken präsentieren möchte.

Der praktisch oft problematische Übergang von Rohdaten zu strukturierten Daten (Schritt 1) wurde bereits in Abschnitt A. thematisiert.<sup>87</sup> Wie aus visuellen Strukturen eine Visualisierung wird (Schritt 3), hängt maßgeblich von der verwendeten Technik ab und kann im Online-Appendix anhand von Code-Beispielen nachvollzogen werden. Der Übergang von strukturierten Daten zu visuellen Strukturen (Schritt 2) erfolgt, indem man den Datenpunkten, die abgebildet werden

<sup>85</sup> Ward, Grinstein und Keim, *Interactive Data Visualization*, 2015, S. 139.

<sup>86</sup> Darstellung in Anlehnung an Ward, Grinstein und Keim, *Interactive Data Visualization*, 2015, S. 139 Figure 4.1, deren Abbildung allerdings auf interaktive, computergestützte Visualisierungen zugeschnitten ist. In der Visualisierungsliteratur finden sich unterschiedliche Darstellungen des Visualisierungsprozesses (auch genannt *visualization pipeline*), deren Schritte sich auf die hier gewählte, bewusst schlicht gehaltene Prozessdarstellung abbilden lassen; beispielsweise Ware, *Information Visualization*, 2013, S. 4 Figure 1.2; Telea, *Data Visualization*, 2015, S. 124 Figure 4.1 und S. 125 Figure 4.2.

<sup>87</sup> Dazu S. 85–88.

sollen,<sup>88</sup> graphische Symbole (*graphical symbols, marks*) zuordnet.<sup>89</sup> Graphische Symbole können grundsätzlich nulldimensional (Punkte), eindimensional (Linien), zweidimensional (Flächen) oder dreidimensional (Volumina) sein.<sup>90</sup> Darüber hinaus können sie sich unter anderem in ihrer Position (in zwei oder drei Dimensionen, mit oder ohne gemeinsame Skala), Form, Größe (in einer, zwei oder drei Dimensionen) und Farbe (Farbton [*hue*], Helligkeit [*lightness*] und Sättigung [*saturation*]) unterscheiden.<sup>91</sup> Diese (und weitere) Faktoren werden als *visuelle Variablen* (*visual variables, visual channels*) bezeichnet; sie können eingesetzt werden, um dem Betrachter die Ausprägungen der Merkmale zu kommunizieren, die zu den dargestellten Datenpunkten erfasst wurden.<sup>92</sup>

Die Zuordnung einer visuellen Variablen zu einem Merkmal wird auch als visuelle Kodierung (*visual encoding*) dieses Merkmals bezeichnet.<sup>93</sup> Eine visuelle Variable kann maximal ein Merkmal kodieren, ein Merkmal kann aber auch in

---

<sup>88</sup> Wo im Folgenden von Datenpunkten gesprochen wird, sind sowohl einzelne Beobachtungen als auch zusammenfassende Statistiken gemeint. In der Visualisierungsliteratur wird in der Regel schlicht von „den Daten“ (*the data*) gesprochen, was aufgrund der etwas anderen Bedeutung dieses Ausdrucks in der Statistikkultur für Texte, die auf Material aus beiden Disziplinen zurückgreifen, ein gewisses Verwirrungspotenzial birgt.

<sup>89</sup> *Munzner, Visualization Analysis & Design*, 2015, S. 95; *Ward, Grinstein und Keim, Interactive Data Visualization*, 2015, S. 147.

<sup>90</sup> Dreidimensionale Symbole sind wenig verbreitet, was auch damit zusammenhängt, dass dreidimensionale Visualisierungen ganz besondere Probleme mit sich bringen; dazu im Überblick *Munzner, Visualization Analysis & Design*, 2015, S. 117–130.

<sup>91</sup> Die Aufzählung ist nicht abschließend; ebenso denkbar sind beispielsweise Unterschiede in Ausrichtung und (bei digitalen Visualisierungen) Bewegung. Zu den verschiedenen Dimensionen von Farbe und ihren Einsatzmöglichkeiten in der Visualisierung *Brewer, Color Use Guidelines for Mapping and Visualization*, in: *MacEachren und Taylor* (Hrsg.), *Visualization in Modern Cartography*, 1994, S. 123–147. Im Anschluss an die dort verwendete Terminologie werden nachfolgend auch Weiß, Schwarz und Grautöne als Farben bezeichnet. Sie unterscheiden sich von anderen Farben dadurch, dass sie weder Farbton noch Sättigung haben; a.a.O. S. 126. Ob Helligkeit als Dimension von Farbe aufgefasst oder separat behandelt werden sollte, wird unterschiedlich beurteilt und hängt auch davon ab, welches Ziel die Einteilung verfolgt. In der Computergraphik ist die Repräsentation von Farbe mit Farbton, Sättigung und Helligkeit als Komponenten seit *Smith, SIGGRAPH 78 Conference Proceedings* 1978, S. 12–19, gängig; Beispiele aus der Einführungsliteratur zur Visualisierung sind *Telea, Data Visualization*, 2015, S. 84, und *Munzner, Visualization Analysis & Design*, 2015, S. 96. Für eine separate Behandlung der Helligkeit etwa *Ware, Information Visualization*, 2013, S. 118–119, und *Ward, Grinstein und Keim, Interactive Data Visualization*, 2015, S. 152.

<sup>92</sup> *Bertin, Sémiologie Graphique*, 2005, S. 42; *Munzner, Visualization Analysis & Design*, 2015, S. 96; *Ward, Grinstein und Keim, Interactive Data Visualization*, 2015, S. 147. Die Autoren nehmen unterschiedliche Differenzierungen vor; die hiesige Darstellung orientiert sich an *Munzner, Visualization Analysis & Design*, 2015, S. 94–115, lässt aber einige der dort angesprochenen Faktoren aus, um die Darstellung nicht zu überfrachten.

<sup>93</sup> *Munzner, Visualization Analysis & Design*, 2015, S. 95; *Ward, Grinstein und Keim, Interactive Data Visualization*, 2015, S. 147.

mehreren verschiedene Variablen kodiert werden.<sup>94</sup> Welche visuellen Variablen zur Kodierung infrage kommen, hängt insbesondere vom Typ des betrachteten Merkmals ab. Während sich Farbton und Form zur Kodierung qualitativer Merkmale eignen, können Position, Größe, Helligkeit und Sättigung Ausprägungen quantitativer Merkmale kommunizieren.<sup>95</sup> Bei der Vermittlung der kodierten Informationen sind die visuellen Variablen unterschiedlich effektiv, insbesondere unterscheiden sie sich darin, wie genau die dargestellten Werte vom Betrachter im Vergleich zu objektiven Messungen wahrgenommen werden (Abb. 3.5, S. 106).<sup>96</sup> Die Effektivität einer visuellen Variablen hängt zudem davon ab, welche weiteren Variablen zur Kodierung von Merkmalen der dargestellten Datenpunkte eingesetzt werden; beispielsweise verschmelzen Länge und Breite eines graphischen Symbols zu einer einzigen Wahrnehmung von dessen Fläche.<sup>97</sup>

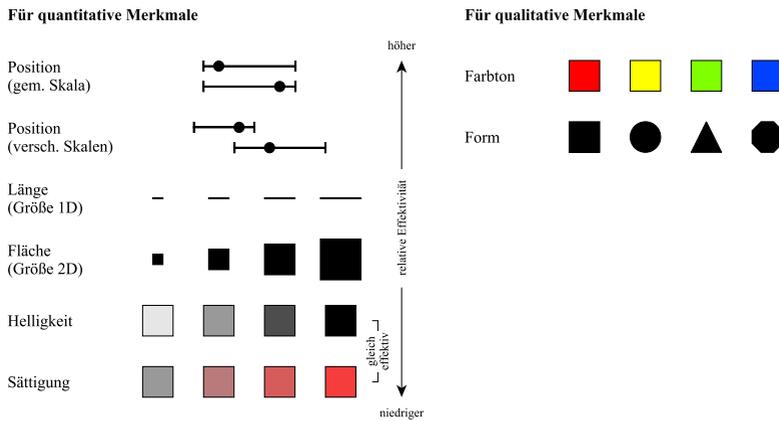
---

<sup>94</sup> Ware, *Information Visualization*, 2013, S. 159; Munzner, *Visualization Analysis & Design*, 2015, S. 97. Die redundante Kodierung eines Merkmals erleichtert es dem Betrachter, Unterschiede zwischen Datenpunkten mit verschiedenen Merkmalsausprägungen wahrzunehmen, hat aber den Nachteil, dass dann weniger visuelle Variablen zur Kodierung anderer Merkmale zur Verfügung stehen.

<sup>95</sup> Munzner, *Visualization Analysis & Design*, 2015, S. 103–106. Die Autorin unterscheidet zwischen kategorialen und geordneten Merkmalen, wobei erstere über den *identity channel* und letztere über den *magnitude channel* zu kodieren sind. Die dort kategorial genannten Merkmale heißen im hiesigen Sprachgebrauch qualitativ; unter den dort als geordnet bezeichneten Merkmalen werden hier ordinale und metrische Merkmale zusammengefasst. Sofern bei ordinalen Merkmalen allerdings der Identitätsaspekt überwiegt (dazu bereits oben S. 93–99) und dieser in der Darstellung hervorgehoben werden soll, scheint ihre Kodierung in visuellen Variablen angemessen, die auf qualitative Merkmale zugeschnitten sind. Die Einteilung der visuellen Variablen wurde daher bewusst an der Unterscheidung qualitativ/quantitativ orientiert.

<sup>96</sup> Die wahrgenommene Größe von Sinneseindrücken folgt einem Potenzgesetz, dessen Basis die physische Intensität des Stimulus ist: Größe des Sinneseindrucks = (Physische Intensität des Stimulus)<sup>Exponent</sup>. Nur, wenn der Exponent exakt 1 ist, wird ein Unterschied in der physischen Intensität des Stimulus unverzerrt wahrgenommen. Das gilt annähernd etwa für die Länge (eindimensionale Größe), nicht aber für die Fläche (zweidimensionale Größe). Bei letzterer ist der Exponent 0,7, was bedeutet, dass die Verdopplung der Fläche eines graphischen Symbols in einem Symbol resultiert, das weniger als doppelt so groß wahrgenommen wird. Speziell für Länge und Fläche Stevens und Guirao, *Journal of Experimental Psychology* 66 (1963), S. 177–186; allgemein Stevens, *Psychophysics*, 1975, S. 13–19; im Überblick Munzner, *Visualization Analysis & Design*, 2015, S. 103–106. Was aus den einschlägigen Messungen und experimentellen Resultaten für die Kodierung einzelner Merkmalsausprägungen in der Praxis folgt, ist allerdings nicht ganz klar.

<sup>97</sup> Ware, *Information Visualization*, 2013, S. 162–168; Munzner, *Visualization Analysis & Design*, 2015, S. 106–109.

Abbildung 3.5: Visuelle Variablen<sup>98</sup>

Wie die Übersetzung von strukturierten Daten in visuelle Strukturen vonstatten geht, sei am Beispiel von explorativen Visualisierungen zu den Rechtsprechungsdaten aus Flatland verdeutlicht. Diese sind bereits strukturiert: Sie haben die Form von Tabellen, in denen jede Zeile für eine Beobachtung und jede Spalte für ein Merkmal steht; alle Ausprägungen eines Merkmals wurden auf derselben Skala gemessen. Fehlende Werte, die einer Sonderbehandlung bedürftig sind, kommen nicht vor. Außerdem sind die Daten so gespeichert, dass sie sich problemlos mit Computerprogrammen weiterverarbeiten lassen.

In Abschnitt A. (S. 93–99) wurde bereits festgestellt, dass die maximale Anzahl der Zitate, die eine Entscheidung erhält, mit 650 sehr stark von allen Durchschnittswerten abweicht.<sup>99</sup> Man kann versuchen, sich das genauer anzusehen, indem man zählt, wie häufig jede Anzahl von Zitaten vorkommt, und jeder Anzahl ein Rechteck als zweidimensionales Symbol zuordnet. Kodiert man die Anzahl selbst als Position des Rechtecks (auf der x-Achse) und trägt die Häufigkeit ihres Vorkommens als Länge des Rechtecks (auf der y-Achse) ab, während die Breite des Rechtecks für alle Rechtecke konstant bleibt, so ergibt sich ein klassisches *Histogramm* (*histogram*), das – wie Abbildung 3.6 (a) (S. 107) zeigt – als Übersicht zu den Zitierhäufigkeiten in den Rechtsprechungsdaten nicht sehr informativ ist. Daran ändert sich auch dann nichts, wenn man (als Ergebnis von Erwägungen in Schritt 3 des Visualisierungsprozesses) die Länge der Rechtecke auf der y-Achse logarithmisch skaliert oder die kumulierte bzw. *kumulative Häufigkeitsverteilung* als Linie abträgt (Abb. 3.6 (b), S. 107).<sup>100</sup>

<sup>98</sup> Darstellung in Anlehnung an *Munzner*, *Visualization Analysis & Design*, 2015, S. 94, 102.

<sup>99</sup> Zur Erinnerung: Modus und Median liegen bei 0, das arithmetische Mittel liegt bei 13, 10.

<sup>100</sup> Andere Möglichkeiten zur Darstellung derselben Informationen werden im Rahmen der Quantifizierung auf der Makroebene vorgestellt (S. 165–186). Unabhängig davon, dass Histogramme nicht für alle Häufigkeitsverteilungen lesbar sind, haben sie die unschöne Eigenschaft,

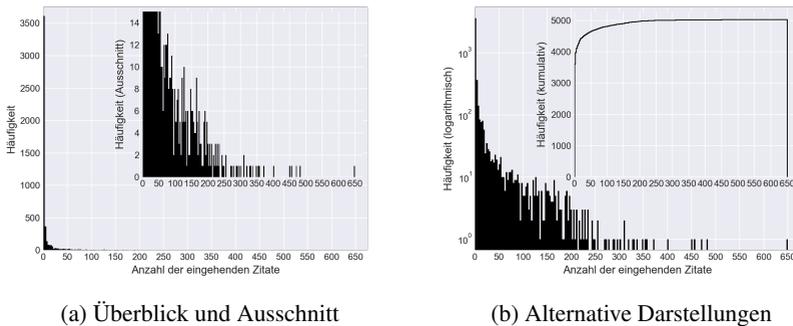


Abbildung 3.6: Häufigkeitsverteilung der eingehenden Zitate in Flatlands Rechtsprechung als Histogramme

Unbefriedigend an den Histogrammen erscheint auch, dass hier die Beobachtungen bereits aggregiert worden sind (die einzelnen Entscheidungen tauchen nur als Teil der aus ihnen ermittelten Häufigkeiten auf) und dass kein einziges der ursprünglich erhobenen Merkmale (etwa die Inhaltsziffer oder die Ebene des entscheidenden Gerichts) Eingang in die Visualisierung gefunden hat. Das lässt sich ändern, indem man jeder einzelnen Beobachtung zunächst einen Punkt als nulldimensionales Symbol zuordnet. Eine Möglichkeit, mehrere potenzielle Zusammenhänge gleichzeitig zu erkunden, besteht dann darin, das Datum einer Entscheidung als Position des Punkts auf der x-Achse und die Anzahl der erhaltenen Zitate als Position auf der y-Achse zu kodieren. Wenn man nun zusätzlich die inhaltlichen Kategorien der Entscheidungen auf verschiedene Farbtöne abbildet und die Position des entscheidenden Spruchkörpers in der Gerichtshierarchie über verschiedene Größen und Formen kodiert, so kann man gleichzeitig drei Fragen zur Zitierhäufigkeit stellen: (1) Hängt die Zitierhäufigkeit mit dem Alter einer Entscheidung zusammen (Temporalzusammenhang)?<sup>101</sup> (2) Hängt die Zitierhäufigkeit mit der Position des entscheidenden Spruchkörpers in der Gerichtshierarchie zusam-

eindimensionale Werte (die Häufigkeit des Vorkommens einer bestimmten Anzahl von Zitaten) zweidimensional zu kodieren. Damit haben sie – was an der verbalen Beschreibung der Kodierungserwägungen bereits deutlich wird – einen hohen Anteil von „redundanter Datentinte“ (*redundant data-ink*), die entfernt werden könnte, ohne die Aussage der Darstellung zu verändern; dazu *Tufte*, *The Visual Display of Quantitative Information*, 2001, S. 91–105, insbes. S. 96, 100–102.

<sup>101</sup> Grundsätzlich plausibel, denn je älter eine Entscheidung, desto mehr Entscheidungen nach ihr hätten die Möglichkeit gehabt, sie zu zitieren.

men (Hierarchiezusammenhang)?<sup>102</sup> (3) Hängt die Zitierhäufigkeit mit dem Inhalt einer Entscheidung zusammen (Inhaltszusammenhang)?<sup>103</sup>

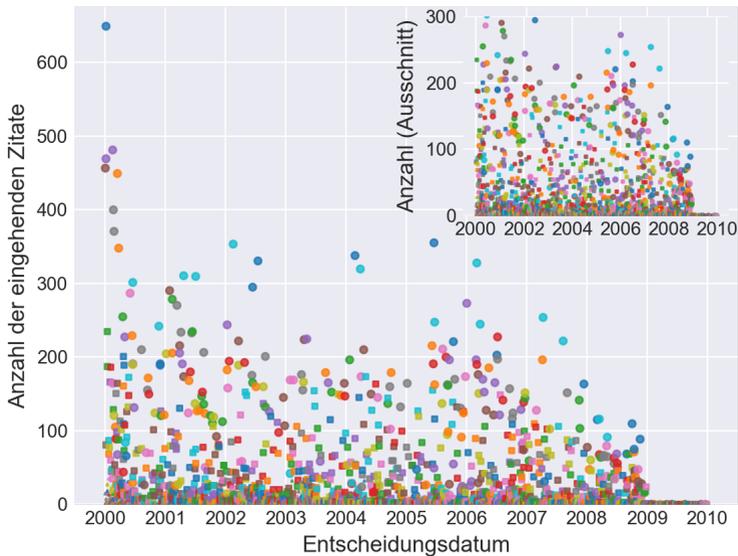


Abbildung 3.7: Anzahl der eingehenden Zitate in Abhängigkeit vom Entscheidungsalter

Das Resultat (Abb. 3.7, S. 108) ist ein Streudiagramm (*scatter plot*), das allerdings vor allem die Grenzen der explorativen Visualisierung verdeutlicht: Optisch drängen sich Temporalzusammenhang und Hierarchiezusammenhang auf. Doch bleibt der vermutete Temporalzusammenhang auch bestehen, wenn man alle anderen Merkmale konstant hält – oder wenn man nicht die Anzahl der erhaltenen Zitate, sondern die Anzahl der zitierenden Entscheidungen als Position auf der y-Achse kodiert? Und ist es Zufall, dass die höchsten Zitierwerte (offenbar) von einigen wenigen Sachmaterien erreicht werden? Zur Beantwortung solcher Fragen reichen visuelle Werkzeuge oft nicht aus, was auch damit zusammenhängt, dass die menschliche Wahrnehmung von Trends und Strukturen in den Daten, die einer Visualisierung zugrunde liegen, davon beeinflusst wird, in welchem Größenverhältnis Breite und Höhe der Darstellung zueinander stehen (*aspect*

<sup>102</sup> Grundsätzlich plausibel, denn in einem hierarchischen System haben die unteren Gerichte mehr Anreize, sich an der Rechtsprechung höherer Gerichte zu orientieren und dies auch nach außen deutlich zu machen, als umgekehrt.

<sup>103</sup> Plausibel etwa dann, wenn man annimmt, dass es auch in Flatland ein Prozessrecht gibt, sodass Entscheidungen, die dieses betreffen, eine höhere Wahrscheinlichkeit haben, für andere Entscheidungen relevant zu werden, als Entscheidungen zu einer beliebigen Sachmaterie.

*ratio*) und wie sich die Datenpunkte über die Darstellungsfläche verteilen.<sup>104</sup> Über diese Faktoren wird in Schritt 3 des Visualisierungsprozesses entschieden, der auch mehrfach durchlaufen werden kann (s. Abb. 3.4, S. 103). Damit bergen Visualisierungen stets die Gefahr, insbesondere diejenigen Zusammenhänge zu offenbaren, die ihr Urheber von vornherein erwartet hat.

### III. Visualisierung von Graphen

Die bisher angestellten Überlegungen gelten für die Visualisierung jeder Form von Daten in ähnlicher Weise. Das Besondere der juristischen Netzwerkforschung ist nun aber, dass sie mit *relationalen* Daten arbeitet: Daten, die neben Informationen zu einzelnen Objekten (z.B. zu Gerichtsentscheidungen) vor allem auch Informationen zu den Beziehungen zwischen diesen Objekten (z.B. zu Zitaten) enthalten und für die es sich gerade deshalb anbietet, sie mathematisch als Graphen zu repräsentieren. Wie Abbildung 3.6 (S. 107) und Abbildung 3.7 (S. 108) andeuten, können viele Informationen zu Graphen auch mithilfe klassischer Diagrammtypen visualisiert werden. Diese Diagrammtypen stoßen allerdings an ihre Grenzen, wenn es darum geht, Beziehungsinformationen ohne Aggregation darzustellen, also jeden Punkt und jede Linie so zu kodieren, dass die Information darüber erhalten bleibt, welcher Punkt mit welchen anderen Punkten auf welche Weise verbunden ist. Eine solche Darstellung erfordert, dass der gesamte Graph visualisiert wird. Doch wie stellt man Graphen graphisch dar? Eine Möglichkeit besteht darin, sie als Diagramme aus Punkten und Linien zu zeichnen (1.). Alternativ kann man Graphen aber auch mithilfe ihrer Matrizen visualisieren (2.).

#### 1. Graphdiagramme

Die Darstellung von Graphen als Diagramme aus Punkten und Linien (*node-link diagrams*)<sup>105</sup>, die hier als *Graphdiagramme* bezeichnet werden, ist in dieser Arbeit bereits mehrfach aufgetaucht. Die meisten bisher in dieser Form abgebildeten Graphen sind allerdings konzeptuell<sup>106</sup> oder klein<sup>107</sup>, sodass ihre Elemente per Hand positioniert werden konnten. Dieses Vorgehen kommt für Graphen mit hunderten oder tausenden Punkten und Linien, beispielsweise für den Zitiergraphen und den Verflechtungsgraphen aus Flatland oder auch für die Gesetzesbäume aus dem ersten Kapitel<sup>108</sup>, allerdings nicht in Betracht. In solchen Fällen muss die Position der einzelnen Elemente eines Graphen mit Computerunterstützung

<sup>104</sup> Dazu die Beispiele bei *Cleveland, Harris und McGill*, *The Bell System Technical Journal* 62 (1983), S. 1659 (1669 Fig. 7) (für die Ausbreitung von Punkten in Streudiagrammen) und *Heer und Agrawala*, *IEEE Transactions on Visualization and Computer Graphics* 12 (2006), S. 701–708 (für die *aspect ratio* in Liniendiagrammen).

<sup>105</sup> *Ware*, *Information Visualization*, 2013, S. 222; *Munzner*, *Visualization Analysis & Design*, 2015, S. 201; *Ward, Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 326.

<sup>106</sup> Beispiele: Abb. 2.1–2.4 (S. 38, 40, 45, 54).

<sup>107</sup> Beispiele: Abb. 1.3 (a) (S. 4), Abb. 1.4 (S. 6) und Abb. 2.6 (S. 66).

<sup>108</sup> Siehe Abb. 1.5 (S. 7) und Abb. 1.6 (S. 8).

berechnet und der Graph durch ein Computerprogramm automatisch gezeichnet werden. Hierzu sind dann explizite Regeln anzugeben, nach denen der Computer die horizontale und vertikale Position der einzelnen Elemente bestimmen kann.<sup>109</sup> Da die Positionen der Linien von den Positionen der Punkte abhängen, muss man die Punktpositionen zuerst berechnen. Zur Ermittlung der Linienpositionen ist dann vor allem festzulegen, ob die Linien auf dem kürzesten Weg (und damit gerade), über mehrere Ecken und geschwungen von einem Endpunkt zum anderen verlaufen sollen und wie potenzielle Überschneidungen mit Punkten oder anderen Linien zu behandeln sind.<sup>110</sup>

#### a) Merkmalsbasierte Positionierung

Bei Bäumen, deren Punkte sich sowohl hierarchisch als auch sequenziell ordnen lassen, bietet es sich an, diese Ordnungen in horizontale und vertikale Koordinaten zu übersetzen, wie es beispielsweise bei den Gesetzesbäumen im ersten Kapitel unternommen wurde.<sup>111</sup> Für ein Netzwerk, dessen Objekten geographische oder zeitliche Informationen zugeordnet sind, kann es sich lohnen, die Beziehungsstruktur unter Berücksichtigung dieser Informationen zu betrachten. Das ist für die Geographie mit Länge und Breite, die sich direkt in x- und y-Koordinaten übersetzen lassen, leichter als für die Zeit, die im Alltag eindimensional erfasst wird. Dieses Problem kann man bewältigen, indem man ausnutzt, dass sich Tage und Monate im Jahresverlauf wiederholen (Zyklizität der Zeiterfassung). So kann man die Zeit in zwei Teile zerlegen und diese als horizontale und vertikale Punktpositionen kodieren – beispielsweise als einen diskreten Jahresteil und einen diskreten Monat-und-Tag-Teil.<sup>112</sup> Genauso ist es aber möglich, temporale Informationen nur auf eine Positionskomponente (in der Regel die horizontale) abzubilden und die andere Komponente zur Kodierung eines weiteren interessanten Merkmals zu verwenden (Abb. 3.8, S. 111).<sup>113</sup>

<sup>109</sup> Das Ergebnis der Positionierung bezeichnet man als *Layout* des Graphen; die Regeln, aus denen es hervorgeht, als *Layout-Algorithmus*.

<sup>110</sup> Hiervon abzugrenzen ist die Frage, wie genau Punkte und Linien mit (teilweise) identischen Koordinaten sich in der endgültigen Visualisierung sichtbar überschneiden. Die Antwort darauf hängt vor allem davon ab, in welcher Reihenfolge die einzelnen Elemente des Graphen gezeichnet werden (eine Frage, die in Schritt 3 des Visualisierungsprozesses beantwortet wird).

<sup>111</sup> Siehe Abb. 1.5 (S. 7) und Abb. 1.6 (S. 8). Dort wurde die Hierarchie als vertikale und die Sequenz als horizontale Position kodiert. Die Umkehrung dieser Zuordnung ist ebenfalls möglich; welche Kodierung für eine Visualisierung am besten geeignet ist, hängt vom Einzelfall ab.

<sup>112</sup> Allgemein ist es schwierig, zyklische Daten in *einer* Visualisierung so darzustellen, dass alle Muster, die in ihnen auftreten könnten, gleichsam erkennbar sind. Dazu *Tukey*, *Some Graphic and Semigraphic Displays*, in: *Bancroft* (Hrsg.), *Statistical Papers in Honor of George W. Snedecor*, 1972, S. 293 (297–301); im juristischen Kontext *Coupette und Fleckner*, *JZ* 73 (2018), S. 379 (387–388).

<sup>113</sup> So wird etwa in den Visualisierungen der Rechtsprechung des *U.S. Supreme Court*, die man mit *CourtListener* erstellen kann, die vertikale Position dazu genutzt, die Variablen *decision*

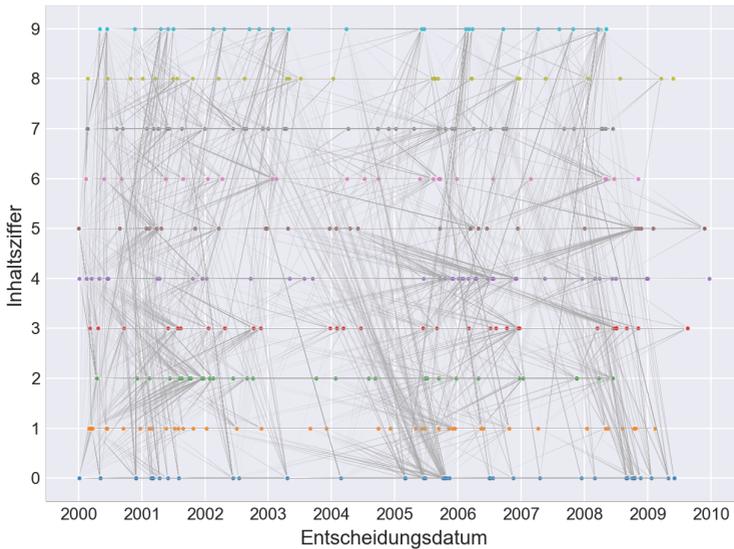


Abbildung 3.8: Zitate zwischen Entscheidungen des *High Court* nach Entscheidungsdatum und Inhaltsziffer<sup>114</sup>

Die bisher geschilderten Fälle laufen alle darauf hinaus, dass beim Übergang von strukturierten Daten zu visuellen Strukturen in Schritt 2 des Visualisierungsprozesses zunächst jedem Objekt ein Punkt und jeder Beziehung eine Linie zugeordnet wird. Dann kodiert man einzelne Objektmerkmale als horizontale und vertikale Komponenten der Punktposition und bestimmt die Linienpositionen in Abhängigkeit von den Punktpositionen (*merkmalsbasierte Positionierung*). Dieses Vorgehen ist allerdings nicht möglich, wenn es keine Merkmale gibt, die sich

*direction* (liberal oder konservativ) und *vote count* (9-0, 8-1, 7-2, 6-3 oder 5-4) aus der *Supreme Court Database* zu kodieren.

<sup>114</sup> Die Abbildung zeigt den Zitiergraphen der flatländischen Rechtsprechung in seiner Version mit Mehrfachkanten, eingeschränkt auf Punkte, die Entscheidungen des hohen Gerichts darstellen; dieser Graph hat 319 Punkte (Entscheidungen) und 4.125 Pfeile (Zitate). Das Entscheidungsdatum (ein diskretes Merkmal an der Grenze zur Quasi-Stetigkeit) ist hier wie im Streudiagramm aus Abbildung 3.7 (S. 108) als Position auf der x-Achse kodiert; die Inhaltsziffer (ein diskretes Merkmal) ist als Position auf der y-Achse kodiert. Die Zitate zwischen den einzelnen Entscheidungen sind als Pfeile zwischen den einzelnen Punkten dargestellt, die von ihrem dickeren Ende wie Nadeln zu ihrem dünneren Ende verlaufen (und deren Richtung sich aus den x-Koordinaten ihrer Endpunkte ergibt, da in Flatland ausschließlich in die Vergangenheit zitiert wird). Trotz der vielen Pfeile ist in dieser Darstellung erkennbar, dass Entscheidungen einzelner Inhaltsziffern vor allem Entscheidungen benachbarter Inhaltsziffern zitieren, wobei die Inhaltsziffer 0 ein Nachbar der Inhaltsziffer 9 zu sein scheint. Inwieweit Entscheidungen einer bestimmten Inhaltsziffer andere Entscheidungen derselben Inhaltsziffer zitieren, ist allerdings kaum erkennbar, da sich die Pfeile überlagern, die parallel zur x-Achse verlaufen.

sinnvoll auf horizontale und vertikale Punktpositionen abbilden lassen. Da das beschriebene Verfahren bei der Punktpositionierung keinerlei Informationen zu den Beziehungen zwischen den Objekten berücksichtigt, ist es ferner ungeeignet, wenn die Beziehungsstruktur des Netzwerks *als solche* (seine Topologie) im Zentrum der Analyse stehen soll.

### b) Kräftebasierte Positionierung

Um die Beziehungsstruktur eines Netzwerks visuell zu erfassen, ist es intuitiv wünschenswert, dass Objekte, die eng miteinander verbunden sind, als Punkte in der Visualisierung nah beieinanderstehen, alle Punkte aber weit genug voneinander entfernt bleiben, um optisch unterscheidbar zu sein.<sup>115</sup> Ein Versuch, diesen Wunsch zu erfüllen, beginnt damit, dass man sich die Punkte als Teilchen mit identischer Ladung vorstellt, die sich grundsätzlich abstoßen (*push*), während die Linien als Federn zwischen ihren Endpunkten fungieren und diese grundsätzlich zueinander ziehen (*pull*).<sup>116</sup> Ausgehend von einer zuvor festgelegten Ausgangsposition wird die endgültige Position eines Punktes dann ermittelt, indem man die beschriebenen „physikalischen Kräfte“ in einer Physiksimulation auf den gesamten Graphen wirken lässt (*kräftebasierte Positionierung, force-directed placement*): Das simulierte physikalische System strebt einem Zustand möglichst geringer Energie entgegen, bis es ein mechanisches Gleichgewicht erreicht, bei dem sich die Punktpositionen von einer Berechnungsrunde zur nächsten nicht mehr signifikant verändern.<sup>117</sup>

Die kräftebasierte Positionierung hat den Vorteil, dass sie über die geschilderte Analogie zu physikalischen Systemen verhältnismäßig anschaulich erklärt werden kann. Außerdem lässt sich auf Basis der Physiksimulation relativ einfach ein Graphlayout errechnen. Welches Graphlayout errechnet wird, hängt allerdings davon ab, wie die Physiksimulation in der verwendeten Software umgesetzt ist (*Implementierung*) und mit welchen Argumenten (Informationen, die bei der Berechnung berücksichtigt werden sollen) die Simulationsfunktion aufgerufen wird. Dabei geht die Software in der Regel von zufälligen Startpositionen der Punkte aus, sofern mit dem Funktionsaufruf keine Ausgangspositionen übergeben

<sup>115</sup> Fruchterman und Reingold, *Software–Practice and Experience* 21 (1991), S. 1129 (1131); Telea, *Data Visualization*, 2015, S. 478.

<sup>116</sup> Telea, *Data Visualization*, 2015, S. 478; Munzner, *Visualization Analysis & Design*, 2015, S. 204; Ward, Grinstein und Keim, *Interactive Data Visualization*, 2015, S. 327.

<sup>117</sup> Das Vorgehen ist auch unter den Namen *spring embedding, energy minimization* und *nonlinear optimization* bekannt; Munzner, *Visualization Analysis & Design*, 2015, S. 204. Ein Überblick zu den physikalischen Kräften, die bei der Simulation eingesetzt werden können, findet sich bei Telea, *Data Visualization*, 2015, S. 478–479. Die in der Realität wirkenden Kräfte werden in der Simulation nur ungenau abgebildet – unter anderem, da die Layout-Algorithmen in der Regel ein Abbruchkriterium festlegen (z.B. eine Maximalzahl von Berechnungsdurchgängen), bei dessen Erreichen die Berechnung abgebrochen und das errechnete Layout als Resultat ausgegeben wird.

werden, sodass mehrere Funktionsaufrufe für denselben Graphen unterschiedliche Layouts errechnen können. Das hat zur Folge, dass die auf kräftebasierter Positionierung beruhenden Visualisierungen nicht ohne Weiteres reproduzierbar sind (*Reproduktionsproblem*, Abb. 3.9, S. 113).<sup>118</sup>

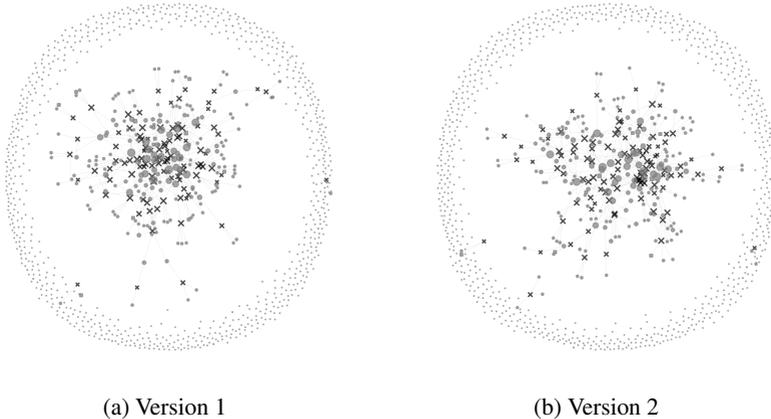


Abbildung 3.9: Kräftebasierte Positionierung der Punkte im Verflechtungsgraphen – Reproduktionsproblem<sup>119</sup>

Bei der Interpretation von Visualisierungen, die auf kräftebasierter Positionierung beruhen, ist außerdem zu beachten, dass eine geringe Distanz zwischen einzelnen Punkten nicht zwangsläufig bedeutet, dass die ihnen entsprechenden Objekte eng miteinander verbunden sind: Die optische Nähe zwischen den Punkten

<sup>118</sup> Die Reproduzierbarkeit lässt sich sicherstellen, indem man den Zufallszahlengenerator im verwendeten Programm mit einer bestimmten Zahl initialisiert (sog. *seed*) und diese Zahl gemeinsam mit den Informationen zur verwendeten Software kommuniziert (Beispiel: Fn. 119, S. 113).

<sup>119</sup> Zur Erstellung der Abbildungen wurden in Schritt 2 des Visualisierungsprozesses folgende Zuordnungen von Merkmalen zu visuellen Strukturen vorgenommen: Der Farbton eines Punktes kodiert den Typ des korrespondierenden Objekts (Schwarz für Unternehmen, Grau für Führungsfiguren), die Form eines Punktes kodiert die Form der Führungsfigur bzw. die Eigenschaft als Unternehmen (ein Kreuz). Die Größe eines Punktes entspricht 0, 5, falls der Punkt keine angrenzende Linien aufweist, anderenfalls entspricht sie der Anzahl der angrenzenden Linien im Graphen, multipliziert mit 7 (die Zahlen haben keine besondere Bedeutung). In den Eigenschaften der Linien wurden keine Merkmale kodiert; die Position von Punkten und Linien folgt einem kräftebasierten Layout-Algorithmus, der in vielen gängigen Softwarepaketen (mit unterschiedlichen Standardeinstellungen) implementiert ist (*Fruchterman-Reingold-Algorithmus*, vorgestellt in *Fruchterman und Reingold*, *Software-Practice and Experience* 21 (1991), S. 1129–1164), hier in der Version von *networkx-1.11*. Der *seed* für Abbildung 3.9 (a) (S. 113) ist 1234, jener für Abbildung 3.9 (b) (S. 113) ist 9876; der Code zur Reproduktion beider Abbildungen ist im Online-Appendix abrufbar.

kann sich auch daraus ergeben, dass sie von anderen Punkten in dieselbe Gegend hinein abgestoßen wurden.<sup>120</sup> Welche der beiden Möglichkeiten einschlägig ist, lässt sich oft nur mit großem Aufwand, manchmal aber auch gar nicht erkennen (*Interpretationsproblem*).<sup>121</sup> Eine weitere Schwierigkeit kräftebasierter Positionierung ist, dass sie (jedenfalls bei einfachen Implementierungen) schnell an ihre Grenzen stößt, wenn es um die Visualisierung großer Graphen geht (*Skalierungsproblem*).<sup>122</sup> In Abbildung 3.10 (S. 115) ist von der Beziehungsstruktur des Zitiernetzwerks nicht viel zu erkennen – es gleicht einer Wollmaus.<sup>123</sup>

Trotz der geschilderten Schwierigkeiten kommt die kräftebasierte Positionierung zur Darstellung von Graphen als Graphdiagramme in der Praxis recht häufig zum Einsatz.<sup>124</sup> Ansätze zur Eindämmung des Wollmauseffekts bei der Visualisierung großer Graphen greifen unter anderem auf Quantifizierungsmethoden der Netzwerkforschung zurück.<sup>125</sup> Daneben fallen die dargestellten Probleme weniger

<sup>120</sup> Munzner, *Visualization Analysis & Design*, 2015, S. 204.

<sup>121</sup> Aus der Kombination von Interpretationsproblem und Nichtdeterminismus resultiert die Versuchung, im Rahmen der analytischen Visualisierung immer neue Darstellungen zu produzieren, bis ein Resultat dem entspricht, was man zu sehen wünscht – und dann genau dieses Resultat bei der Kommunikation der Forschungsergebnisse zur illustrativen Visualisierung einzusetzen. Grundsätzlich ist es aber gerade sinnvoll, einen kräftebasierten Layout-Algorithmus mit verschiedenen Initialisierungen (d.h. verschiedenen Mengen von Startkoordinaten) auszuprobieren, um zu vermeiden, dass die Visualisierung auf einem lokalen Energieminimum aufbaut, das deutlich schlechter ist als das globale Minimum; Ward, *Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 327–328. Vor diesem Hintergrund erscheint es unglücklich, dass nicht jede Visualisierungssoftware dem Nutzer die Restenergie des errechneten Layouts als Indikator für seine Güte mitteilt.

<sup>122</sup> Telea, *Data Visualization*, 2015, S. 483; Munzner, *Visualization Analysis & Design*, 2015, S. 206; Ward, *Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 333.

<sup>123</sup> Im englischen Sprachraum ist die Bezeichnung als *hairball* verbreitet; so etwa bei Munzner, *Visualization Analysis & Design*, 2015, S. 206. Wollmäuse (auf Basis unterschiedlicher Layouts) finden sich auch in der Literatur zur Netzwerkforschung im Recht, ein besonders haariges Exemplar etwa bei Boulet, *Barros-Plataiu und Mazzega*, *Artificial Intelligence and Law* 2016, S. 133 (140); weitere z.B. bei Dibadj, *San Diego Law Review* 46 (2009), S. 1 (17–18); Bommarito, *Katz und Isaacs-See*, *Virginia Tax Review* 30 (2011), S. 523 (543); Boulet, *Mazzega und Bourcier*, *Artificial Intelligence and Law* 19 (2011), S. 333 (338); Li u.a., *Journal of Business and Technology Law* 10 (2015), S. 297 (342, 344, 369, 371, 373); Šadl und Olsen, *Leiden Journal of International Law* 30 (2017), S. 327 (338).

<sup>124</sup> Beispiele aus der Literatur zur Netzwerkforschung im Recht: Porter u.a., *PNAS* 102 (2005), S. 7057 (7058); Fowler, *Social Networks* 28 (2006), S. 454 (463); Paik, *Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (893–894); Dibadj, *San Diego Law Review* 46 (2009), S. 1 (17–18); Katz und Stafford, *Ohio State Law Journal* 71 (2010), S. 457 (489–492); Katz u.a., *Journal of Legal Education* 61 (2011), S. 76 (83); Kim, *Global Environmental Change* 23 (2013), S. 980 (983); Érdi u.a., *Scientometrics* 95 (2013), S. 225 (235, 237); Puig, *The European Journal of International Law* 25 (2014), S. 387 (409–410); Boulet, *Barros-Plataiu und Mazzega*, *Artificial Intelligence and Law* 2016, S. 133 (140–141, 146).

<sup>125</sup> Der Rückgriff kann insbesondere zweierlei Formen annehmen, die auch kombiniert werden können: Einerseits lassen sich Zahlen, die aus netzwerkanalytischer Quantifizierung hervorgehen, in die Berechnung der Punkt- und Linienpositionen einbeziehen, sodass die Beziehungsstrukturen

stark ins Gewicht, soweit eine interaktive Visualisierung erzeugt wird, in der ein Nutzer die Position der Punkte selbst verändern, in den Graphen hineinzoomen, Punkte auf Basis wählbarer Kriterien ein- und ausblenden oder durch Auswahl einzelner Elemente des Graphen zusätzliche Informationen über diese erhalten kann.



Abbildung 3.10: Kräftebasierte Positionierung der Punkte im Zitiergraphen  
– Skalierungsproblem<sup>126</sup>

des zugrundeliegenden Netzwerks deutlicher hervortreten. Andererseits können solche Zahlen genutzt werden, um einen Graphen vor seiner Zeichnung zu reduzieren, indem „unwichtige“ Punkte und Linien eliminiert oder „ähnliche“ Punkte und Linien zusammengefasst werden. Zur Auslegung der Begriffe „Wichtigkeit“ und „Ähnlichkeit“ noch unten S. 129–142 und S. 142–165; allgemein zu den Herausforderungen der Visualisierung großer Graphen und zu möglichen Lösungsansätzen aus der Literatur *Hu und Lei*, *Wiley Interdisciplinary Reviews: Computational Statistics* 7 (2015), S. 107–136.

<sup>126</sup> Zur Erstellung der Abbildung wurde die Position des entscheidenden Gerichts in der Gerichtshierarchie doppelt kodiert – erstens in der Form (nach Maßgabe von Abb. 3.2, S. 89), zweitens in der Größe (1 für Gerichte der unteren, 5 für Gerichte der mittleren und 10 für das Gericht der oberen Stufe). Die Pfeile wurden als gerade Linien und damit insbesondere ohne ihre Richtungsinformationen kodiert. Der verwendete Layout-Algorithmus ist wieder Fruchterman-Reingold (Fn. 119, S. 113) mit *seed 1234*.

### c) Distanzbasierte Positionierung

Eine andere Möglichkeit, die Beziehungsstruktur eines Netzwerks in das Layout seines Graphen einfließen zu lassen, besteht darin, das Layout nicht durch eine Physiksimulation, sondern auf Basis gewünschter Distanzen zwischen den Punkten im Graphen zu ermitteln (*distanzbasierte Positionierung*): Man gibt für jedes Paar von Punkten eine ideale Distanz an und berechnet dann ein Layout, in dem allen Distanzvorgaben möglichst genau entsprochen wird.<sup>127</sup> Diese Form der Visualisierung wird *multidimensionale Skalierung* (*multidimensional scaling*, MDS) genannt.<sup>128</sup> Sie ist ein allgemeines Werkzeug der Datenanalyse und setzt nicht zwingend Netzwerkdaten voraus, da sich für jede beliebige Menge von Daten viele verschiedene Metriken definieren lassen – zum Beispiel, indem aus den Merkmalen der einzelnen Beobachtungen ein Ähnlichkeitsmaß konstruiert wird.<sup>129</sup> Der Name des Verfahrens deutet darauf hin, dass es die Distanzen in einem mehrdimensionalen Raum aus abstrakten Distanzinformationen ermittelt.<sup>130</sup>

Will man die Beziehungsstruktur eines Netzwerks im Graphlayout betonen, so bietet es sich an, die graphentheoretische Distanz zwischen jeweils zwei Punkten als Distanzinformation zu verwenden. Diese gibt an, über wie viele Linien man mindestens laufen müsste, um von einem Punkt zum anderen zu gelangen, wobei Pfeile für die Zwecke des Layouts wie gewöhnliche Linien (also ohne Richtung) behandelt werden.<sup>131</sup> Da die graphentheoretische Distanz zwischen zwei Punkten unendlich ist, wenn sie weder direkt noch indirekt miteinander verbunden sind, berechnet man das Layout eines Graphen separat für seine verbundenen Teile.

Abbildung 3.11 (S. 117) zeigt den größten zusammenhängenden Teil des flatländischen Zitiergraphen in seiner Version ohne Mehrfachkanten – 4.983 Punkte (Entscheidungen) und 43.310 Pfeile (Zitate) – mit zweidimensionalem bzw. dreidimensionalem Layout. Die Punkte sind nach den Inhaltsziffern der Entscheidungen gefärbt, denen sie zugeordnet sind. Anders als bei der kräftebasierten Positionierung ist hier – insbesondere im dreidimensionalen Layout – die Ringstruktur zu erkennen, die den Zitierdaten bewusst gegeben wurde: Entscheidungen der

<sup>127</sup> Eine Möglichkeit, dieses Layout zu berechnen, besteht darin, die Distanzinformationen als ideale Länge von Federn zwischen den einzelnen Punkten darzustellen und wieder physikalische Kräfte zu simulieren, sodass sich distanzbasierte Positionierung technisch auch als kräftebasierte Positionierung umsetzen lässt.

<sup>128</sup> Grundlegend *Torgerson*, *Psychometrika* 17 (1952), S. 401–419.

<sup>129</sup> *Aggarwal*, *Data Mining*, 2015, S. 55–57.

<sup>130</sup> Für Printpublikationen bedeutet „multidimensional“ in der Regel „zweidimensional“. Dreidimensionale Graphlayouts sind (auch merkmals- oder kräftebasiert) möglich, aber ohne Interaktionsmöglichkeiten oft schwer zu interpretieren. Das zeigen etwa die Darstellungen bei *Heinz u.a.*, *The Journal of Politics* 52 (1990), S. 356 (361), und *Heinz u.a.*, *Law & Society Review* 31 (1997), S. 441 (444, 450, 452, 458–459), in denen die dritte Dimension durch Pfeile oder Vorzeichen an den einzelnen Punkten angezeigt wird. Eine Visualisierung, die versucht, eine dreidimensionale MDS-Lösung auf dem Papier in einem Würfel darzustellen, findet sich bei *Laumann und Heinz*, *American Bar Foundation Research Journal* 2 (1977), S. 155 (195).

<sup>131</sup> Zur Bewegung auf Graphen im Einzelnen noch unten S. 123–129.

Inhaltsziffer 1 (orange) zitieren hauptsächlich Entscheidungen der Inhaltsziffern 0 (dunkelblau), 1 und 2 (dunkelgrün); Entscheidungen der Inhaltsziffer 0 zitieren primär Entscheidungen der Inhaltsziffern 9 (hellblau), 0 und 1; abstrakt: Entscheidungen der Inhaltsziffer  $n$  zitieren vor allem Entscheidungen der Inhaltsziffern  $n - 1$ ,  $n$  und  $n + 1$ , wobei  $0 - 1 = 9$  und  $9 + 1 = 0$  gesetzt wurde.

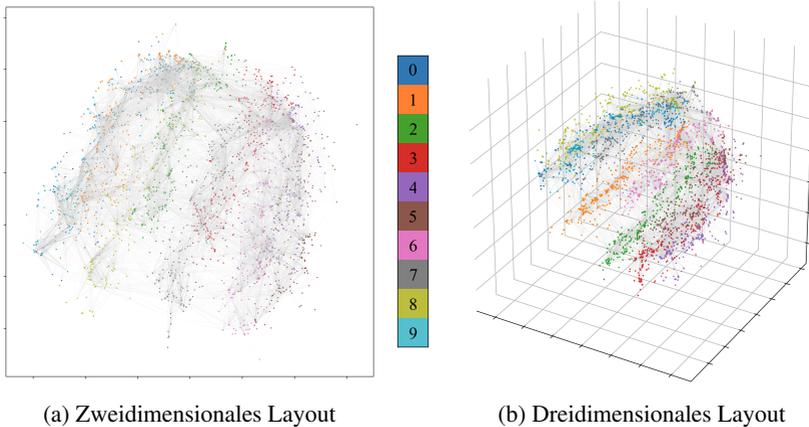


Abbildung 3.11: Layout des Zitiergraphen der flatländischen Rechtsprechung mit multidimensionaler Skalierung (MDS)

Multidimensionale Skalierung taucht in der Literatur zur Netzwerkforschung im Recht immer wieder auf – nicht nur dort, wo Zitate als Beziehungen ersten Grades untersucht werden,<sup>132</sup> sondern auch insbesondere dort, wo es um Ähnlichkeitsbeziehungen geht: Richter, die ähnlich (sonder-)votieren (Beziehungen zweiten Grades);<sup>133</sup> Länder, die in Indizes, die Rechtssysteme zu beschreiben versuchen, ähnliche Werte erzielen (Beziehungen dritten Grades);<sup>134</sup> oder auch Akteure im Rechtssystem, deren Kontakte (Beziehungen ersten Grades) ähnliche Muster aufweisen.<sup>135</sup> Die Ähnlichkeitsbeziehungen selbst werden dabei oft nicht oder nur dann dargestellt, wenn ihre Intensität einen bestimmten Schwellenwert

<sup>132</sup> *Caldeira*, *Social Networks* 10 (1988), S. 29 (43); *Gelter und Siems*, *Utrecht Law Review* 8 (2005), S. 88 (95–96).

<sup>133</sup> *Hook*, *Constitutional Commentary* 24 (2007), S. 221 (254–255); *Engst u.a.*, *JZ* 72 (2017), S. 816 (823–824).

<sup>134</sup> *Cabrelli und Siems*, *The American Journal of Comparative Law* 63 (2015), S. 109 (136, 141, 147); *Siems*, *Journal of Institutional Economics* 12 (2016), S. 579 (596).

<sup>135</sup> *Nelson u.a.*, *Law & Society Review* 22 (1988), S. 237 (289, 292); *Heinz und Manikas*, *Law & Society Review* 26 (1992), S. 831 (842, 847); *Heinz u.a.*, *Law & Society Review* 31 (1997), S. 441 (444, 450, 452, 458–459); *Heinz, Southworth und Paik*, *Law & Society Review* 37 (2003), S. 5 (25, 31); *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (895–897, 901).

(*cut-off*) überschreitet.<sup>136</sup> Die Autoren bieten dann eine Interpretation der erzeugten Bilder an, die dem außenstehenden Betrachter in der Regel plausibel erscheint; mitunter werden auch Zahlen angeführt, an denen sich die Qualität der Anpassung abschätzen lässt.<sup>137</sup>

Je weiter sich die zur Erstellung einer Abbildung genutzte Metrik allerdings von den Beziehungen ersten Grades, auf denen die graphentheoretische Distanz aufbaut, wegbewegt, desto größer ist die Gefahr, dass ein Ähnlichkeitsmaß gerade so konstruiert wird, dass es Bilder liefert, welche die Thesen des Autors stützen. Der Abbildung selbst kann der Betrachter insbesondere dann, wenn die Linien oder Pfeile ganz weggelassen werden, keinerlei Informationen über das verwendete Maß mehr entnehmen, sodass die Darstellung ihre Suggestivkraft frei entfalten kann. So bleibt dem kritischen Leser nur die Möglichkeit, selbst mit den Daten zu arbeiten, um sich von der Plausibilität des gezeichneten Bildes zu überzeugen – was wiederum voraussetzt, dass die verwendeten Daten offengelegt werden.

## 2. Graphmatrizen

Eine andere Möglichkeit, Graphen zu visualisieren, die für größere Graphen (insbesondere solche mit größeren Mengen von Linien) besser geeignet ist als ein Graphdiagramm, beruht auf der Darstellung von Punkten und Linien als *Adjazenzmatrix* (*adjacency matrix*).<sup>138</sup> Eine Adjazenzmatrix ist eine Matrix, die Adjazenzinformationen enthält.<sup>139</sup> Zwei Punkte heißen *adjazent* (*adjacent*, vom lat. *adiacēre*, angrenzen), wenn sie über eine Linie direkt miteinander verbunden sind; man nennt sie dann auch *Nachbarn* (*neighbors*).<sup>140</sup> Matrizen sind Tabellen (ohne Zeilen- und Spaltenüberschriften), mit denen man unmittelbar mathematische Operationen durchführen kann. Die Adjazenzmatrix (abgekürzt **A**) hat für jeden Punkt des Graphen eine Zeile und eine Spalte, sie ist also quadratisch. Die Punkte sind von oben nach unten und von links nach rechts in derselben Reihenfolge angeordnet. Die Werte in den einzelnen Feldern geben an, ob (und

<sup>136</sup> Beispiele für Visualisierungen ohne Darstellung der Ähnlichkeitsbeziehungen: *Nelson u.a.*, *Law & Society Review* 22 (1988), S. 237 (289, 292); *Heinz, Southworth und Paik*, *Law & Society Review* 37 (2003), S. 5 (25, 31); *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (895–896, 901); Beispiele für die Verwendung von Schwellenwerten: *Siems*, *Arizona Journal of International & Comparative Law* 27 (2010), S. 747 (757, 759, 762–763, 766, 768); *Cabrelli und Siems*, *The American Journal of Comparative Law* 63 (2015), S. 109 (136); *Siems*, *Journal of Institutional Economics* 12 (2016), S. 579 (590).

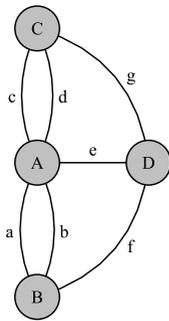
<sup>137</sup> Beispiele: *Heinz u.a.*, *The Journal of Politics* 52 (1990), S. 356 (360); *Paik, Southworth und Heinz*, *Law & Social Inquiry* 32 (2007), S. 883 (897–898, 900–901); *Cabrelli und Siems*, *The American Journal of Comparative Law* 63 (2015), S. 109 (137, 142, 146).

<sup>138</sup> *Munzner*, *Visualization Analysis & Design*, 2015, S. 208; *Ward, Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 333.

<sup>139</sup> In der Sozialen Netzwerkanalyse wird die Adjazenzmatrix auch als Soziomatrix (*sociomatrix*) bezeichnet; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 150.

<sup>140</sup> *Harary*, *Graph Theory*, 1969, S. 9, 82; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 95; *Agnarsson und Greenlaw*, *Graph Theory*, 2007, S. 11.

ggf. wie oft oder wie intensiv) der Punkt in der einschlägigen Zeile mit dem Punkt in der einschlägigen Spalte verbunden ist; bei gerichteten Graphen zeigt der Wert eines Feldes an, ob (und ggf. wie oft oder wie intensiv) ein Pfeil vom Punkt in der jeweiligen Zeile zum Punkt in der jeweiligen Spalte verläuft.<sup>141</sup> Die Graphdarstellung des Königsberger Brückenproblems, die im ersten Kapitel noch vereinfachend als Tabelle bezeichnet wurde (Abb. 1.3, S. 4), ist also genau die Adjazenzmatrix des Graphen, die Zeichnung daneben ist ein konzeptionelles Graphdiagramm (Abb. 3.12, S. 119).



(a) Graphdiagramm

$$A = \begin{matrix} & A & B & C & D \\ A & \begin{pmatrix} 0 & 2 & 2 & 1 \end{pmatrix} \\ B \\ C \\ D \end{matrix}$$

(b) Adjazenzmatrix

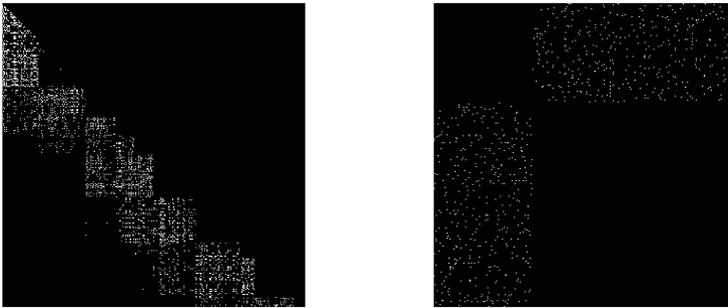
Abbildung 3.12: Darstellungsformen des Graphen zum Königsberger Brückenproblem<sup>142</sup>

Die Repräsentation eines Graphen als Adjazenzmatrix wird auch für große Graphen visuell zugänglich, wenn man die Zahlen in den Feldern durch Farben ersetzt: in Abbildung 3.13 (S. 120) jeweils Schwarz für 0 (keine Verbindung) und Weiß für 1 (Verbindung).<sup>143</sup>

<sup>141</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 152; gegenteilig (und auf die Verwirrungsgefahr der atypischen Definition selbst hinweisend) Newman, *Networks*, 2010, S. 114.

<sup>142</sup> Die Gegenüberstellung der Darstellungsformen macht deutlich, wie sich einzelne Eigenschaften eines Graphen in dessen Adjazenzmatrix niederschlagen: Ist der Graph ungerichtet (sind die dargestellten Beziehungen also symmetrisch), so ist die Adjazenzmatrix symmetrisch (entlang der Diagonale von links oben nach rechts unten). Hat der Graph Mehrfachkanten, so enthält die Adjazenzmatrix Werte, die größer als 1 sind (Newman, *Networks*, 2010, S. 112–113). Hat der Graph keine Schleifen (bestehen also keine Beziehungen eines Objekts zu sich selbst), so befinden sich nur Nullen in der Diagonale (zur Darstellung von Schleifen in Adjazenzmatrizen Newman, *Networks*, 2010, S. 112). Umgekehrt können Werte größer als 1 aber sowohl Kantengewichte als auch Mehrfachkanten zum Ausdruck bringen (zu den Vorteilen dieses Zusammenhangs a.a.O. S. 113), was für die Repräsentation von gewichteten Multigraphen ein Problem darstellt.

<sup>143</sup> Man kann Graphen auch als Adjazenzmatrizen speichern. Das hat gegenüber der hier für die fiktiven Daten gewählten Speicherung als separate Tabellen von Punkten (bzw. Objekten)



(a) Entscheidungen des *High Court* (b) Unternehmen und Führungsfiguren mit Führungspositionen

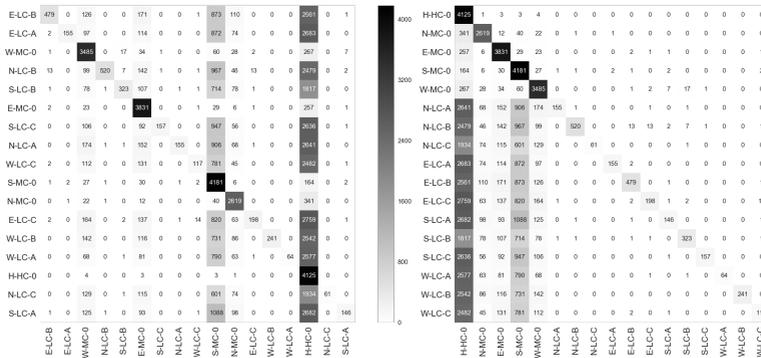
Abbildung 3.13: Matrixdarstellungen zu Rechtsprechungs- und Unternehmensdaten in Flatland<sup>144</sup>

Hier zeigt sich eine Stärke der visuellen Matrixdarstellung: Aus Abbildung 3.13 (a) (S. 120) lässt sich entnehmen, dass Flatlands höchstes Gericht nur eine begrenzte Zeitspanne in die Vergangenheit hinein zitiert, denn unterhalb der Diagonale

und Linien (bzw. Beziehungen) allerdings drei Nachteile. Erstens sind Adjazenzmatrizen zur Speicherung ineffizient, sofern viele potenzielle Verbindungen nicht vorhanden sind, sodass hauptsächlich Nullen gespeichert werden. Zweitens lassen sich in Adjazenzmatrizen keine zusätzlichen Merkmale zu den Objekten und Beziehungen speichern, die durch den Graphen repräsentiert werden. Drittens ist es in der Matrixdarstellung ohne Kontextinformation unmöglich, zwischen gewichteten Kanten und Mehrfachkanten zu unterscheiden, sodass die verlustlose Speicherung eines gewichteten Multigraphen als Adjazenzmatrix unmöglich ist. *Clark, Lauderdale und Katz*, *Political Analysis* 20 (2012), S. 329 (333–334) beschreiben ihre Daten zur Anzahl der Zitate zwischen 18.713 *cases* des *U.S. Supreme Court* als Adjazenzmatrix. Diese hat  $18.713^2 = 350.176.369$  Felder, von denen nach Angaben der Autoren 20.372.788 den Wert 0, 30.702 den Wert 1 und 45.245 einen Wert größer 1 annehmen (a.a.O.). Welche Werte in den verbliebenen 329.727.634 Matrixzellen stehen (Zählzahlen können keine Werte unter Null annehmen), bleibt dabei unklar.

<sup>144</sup> Abbildung 3.13 (a) (S. 120) zeigt, welche Entscheidungen des flatländischen *High Court* einander zitieren. Ein weißes Feld bedeutet, dass die Entscheidung in der jeweiligen Zeile die Entscheidung in der jeweiligen Spalte mindestens einmal zitiert. Die Punkte, welche die Entscheidungen darstellen, sind nach Datum aufsteigend sortiert (von links nach rechts und von oben nach unten). Der so visualisierte Graph hat 319 Punkte und 2.744 Pfeile. Abbildung 3.13 (b) (S. 120) zeigt die Beziehungen zwischen Flatlands Unternehmen und Führungsfiguren, die zum Stichtag mindestens eine Führungsposition innehaben. Die Punkte, die Unternehmen und Führungsfiguren darstellen, sind (von links nach rechts und von oben nach unten) nach ihrer ID im Verflechtungsgraphen aufsteigend alphabetisch sortiert, was dazu führt, dass erst die Unternehmen, dann die Führungsfiguren in der Matrixdarstellung auftauchen. So zeigt sich die für bipartite Graphen charakteristische Aufteilung der Matrix in vier Blöcke, von denen zwei (hier oben links und unten rechts) komplett leer sind, da zwischen Punkten in derselben Teilmenge keinerlei Linien auftreten. Da Führungsfiguren ohne Führungspositionen weggelassen wurden, hat der so visualisierte Graph 306 Punkte (davon 100 für Unternehmen) und 408 Linien.

befindet sich ein weißes Band und die untere linke Ecke der Matrix ist komplett schwarz. Aus Abbildung 3.13 (b) (S. 120) geht hervor, dass der Verflechtungsgraph bipartit ist: Die obere linke und die untere rechte Ecke sind komplett schwarz, eine Bipartition (entsprechend der Aufteilung in Unternehmen und Führungsfiguren) erhält man also, indem man die Punkte, aus deren Zeilen bzw. Spalten sich die zwei schwarzen Blöcke ergeben, in zwei verschiedenen Mengen zusammenfasst.<sup>145</sup> Diese Einsichten sind im hiesigen Kontext relativ unspektakulär, da bereits viele Fakten zu Zitier- und Verflechtungsgraphen eingeführt wurden – und sie sind nur deshalb so offensichtlich, weil die Punkte in Zeilen und Spalten eine bestimmte Reihenfolge haben. Das Erkenntnispotenzial der Matrixdarstellung (vor allem im Rahmen der analytischen Visualisierung) wird daher besonders deutlich, wenn man für einen unbekanntenen Graphen verschiedene Anordnungen der Zeilen und Spalten (*Permutationen*) miteinander vergleicht (Abb. 3.14, S. 121).<sup>146</sup>



(a) Zufällige Anordnung (b) Anordnung nach Hierarchie, Region und Buchstaben

Abbildung 3.14: Relevanz der Zeilen- und Spaltenanordnung am Beispiel der Rechtsprechungsdaten<sup>147</sup>

<sup>145</sup> Dass die Matrix in Abbildung 3.13 (b) (S. 120) symmetrisch ist, jene in Abbildung 3.13 (a) (S. 120) aber nicht, liegt daran, dass der Verflechtungsgraph ungerichtet, der Zitiergraph aber gerichtet ist. Der Entschluss, die Verbindungen im Zitiergraphen als Pfeile, jene im Verflechtungsgraphen aber als einfache Linien zu repräsentieren, wurde bei der Einführung der fiktiven Daten gefasst (oben S. 88–93).

<sup>146</sup> Dieses Potenzial bemerkt bereits *Bertin*, *Sémiologie Graphique*, 2005, S. 270–271 (Erstauffage 1967). Die Umsortierung zielt darauf ab, Strukturen in den *direkten* Beziehungen offenzulegen, die zwischen den dargestellten Objekten bestehen; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 284–287. Sie kann interaktiv oder automatisch erfolgen. Das Aufspüren einer mit Blick auf ein gegebenes Qualitätskriterium „optimalen“ Zeilen- und Spaltenanordnung mithilfe von Algorithmen ist allerdings ein computationell anspruchsvolles Problem; dazu knapp *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 287; *Ward, Grinstein und Keim*, *Interactive Data Visualization*, 2015, S. 333–334.

Gegenüber der Darstellung als Diagramme mit Punkten und Linien hat die Visualisierung eines Graphen als Adjazenzmatrix unter anderem den Nachteil, dass indirekte Beziehungen kaum nachzuverfolgen und Strukturen, die sich aus den Beziehungen zwischen mehr als zwei Objekten ergeben, nur bei adäquater Zeilen- und Spaltenanordnung erkennbar sind.<sup>148</sup> Welche Darstellungsform im konkreten Fall vorzuziehen ist, hängt daher vom Schwerpunkt des Erkenntnisinteresses ab. Im weiteren Verlauf dieser Arbeit wird auf beide Visualisierungsansätze immer wieder zurückgegriffen.

### C. Quantifizierung

Die Quantifizierung ist das methodische Fundament der juristischen Netzwerkforschung. Sie ermöglicht intersubjektiv überprüfbare Aussagen über juristische Netzwerke, indem sie die ihnen entsprechenden Graphen analysiert. Bereits in den Ausführungen zur Visualisierung wurden immer wieder Zahlen produziert, um aus ihnen Bilder zu generieren und aus den Bildern schließlich Erkenntnisse zu gewinnen. Im Rahmen der Quantifizierung stehen nun die Zahlen selbst als Erkenntnismedium im Mittelpunkt. Dabei stellen sich auf Mikro-, Meso- und Makroebene jeweils unterschiedliche Fragen, zu deren Beantwortung verschiedene mathematisch fundierte Konzepte fruchtbar gemacht werden.

Viele dieser Konzepte bedienen sich direkt oder indirekt der Bewegung auf Graphen, deren Grundlagen daher zunächst eingeführt werden (I.). Anschließend werden die spezifischen Probleme und Lösungsansätze der Quantifizierung in der juristischen Netzwerkforschung auf Mikroebene (II.), Mesoebene (III.) und Makroebene (IV.) vorgestellt und anhand der fiktiven Daten aus Flatland illustriert.<sup>149</sup> Hierzu wird für jede Untersuchungsebene die ihr im zweiten Kapitel zugeordnete zentrale Frage (Wichtigkeitsfrage, Gruppierungsfrage oder Charakterisierungsfrage) wieder aufgegriffen (jeweils 1.), zu der verschiedene Antwortstrategien behandelt werden (jeweils 2.).<sup>150</sup> Abschließend wird die Zeit als ebenenübergreifender analytischer Faktor in den Blick genommen (V.).

---

<sup>147</sup> Das hier visualisierte Netzwerk wurde vom Zitiernetzwerk abgeleitet, indem alle Entscheidungen eines Gerichts zu einem Objekt (dem Gericht) zusammengefasst und die ein- und ausgehenden Zitate entsprechend summiert wurden. Dies führt zu einem Zitiergraphen mit siebzehn Punkten (und nach wie vor 65.952 Pfeilen). In der Matrixdarstellung signalisiert die Farbe eines Feldes nun, wie oft Entscheidungen des Gerichts in der jeweiligen Zeile in den Jahren 2000 bis einschließlich 2009 Entscheidungen des Gerichts in der jeweiligen Spalte zitiert haben. Um zu verdeutlichen, dass es sich bei der Visualisierung als Adjazenzmatrix um die optische Aufbereitung einer mathematischen Struktur handelt, werden die Zahlen in den einzelnen Feldern hier ebenfalls gezeigt.

<sup>148</sup> *Munzner*, *Visualization Analysis & Design*, 2015, S. 212.

<sup>149</sup> Zu diesen Daten oben S. 88–99.

<sup>150</sup> Zu Wichtigkeits-, Gruppierungs- und Charakterisierungsfrage oben S. 71–73.

Die Ausführungen zielen darauf ab, die juristische Motivation und die mathematische Intuition der Methoden zu vermitteln, die zur Quantifizierung in der juristischen Netzwerkforschung eingesetzt werden.<sup>151</sup> Welche Schwierigkeiten sich bei der praktischen Anwendung dieser Werkzeuge ergeben können, wird im vierten Kapitel anhand realer Daten zur Rechtsprechung des BVerfG erläutert.<sup>152</sup> Für mathematische oder technische Details, welche die Ergebnisse der quantitativen Analysen zum Teil erheblich beeinflussen, aber das Verständnis der Grundlagen erschweren können, wird an den einschlägigen Stellen auf den Online-Appendix oder auf weiterführende Literatur verwiesen.

### I. Basis: Bewegung auf Graphen

Zum Einstieg in die Quantifizierungsmethoden der juristischen Netzwerkforschung hilft eine erneute Rückkehr nach Königsberg. Am Königsberger Brückenproblem lässt sich nachvollziehen, was Bewegung auf Graphen bedeutet: Angenommen, ein Bewohner von Königsberg befindet sich auf dem Landstück, das in Eulers Zeichnung mit B gekennzeichnet ist, so gelangt er von dort auf die Insel A, indem er entweder über Brücke a oder über Brücke b geht (Abb. 3.15 (a), S. 123). Auf dem Graphen entspricht das einer Bewegung von Punkt B nach Punkt A, wahlweise über Linie a oder über Linie b (Abb. 3.15 (b), S. 123).

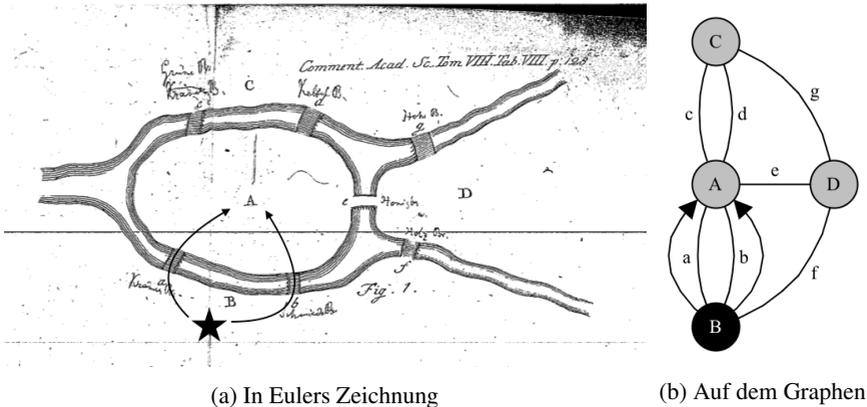


Abbildung 3.15: Bewegung über Königsberger Brücken<sup>153</sup>

<sup>151</sup> Im Folgenden wird oft von schlichten Graphen oder gerichteten Graphen ohne Gewichte und Mehrfachkanten ausgegangen, da die meisten methodischen Werkzeuge zuerst für sie entwickelt wurden und für sie am einfachsten erklärt werden können. Zu den Herausforderungen komplizierterer Graphen unten S. 224–226.

<sup>152</sup> Dazu unten S. 227–319.

<sup>153</sup> Die Position des im Haupttext beschriebenen Einwohners von Königsberg ist in Abbildung 3.15 (a) (S. 123) mit einem Stern, in Abbildung 3.15 (b) (S. 123) durch Schwarzfärbung

Nun gibt es für die Einwohner von Königsberg viele verschiedene Möglichkeiten, sich von einem Landstück zum anderen zu bewegen, wobei sie unter Umständen auch einzelne Landstücke mehrfach besuchen oder einzelne Brücken mehrfach überqueren können. In der Sprache der Graphentheorie werden diese Möglichkeiten als *Wege* (*walks* bzw. *paths*) bezeichnet.<sup>154</sup> Ein Punkt, der mit einem anderen Punkt über einen Weg verbunden ist, heißt von diesem Punkt aus *erreichbar* (*reachable*).<sup>155</sup> Im Graphen zum Königsberger Brückenproblem ist jeder Punkt von jedem anderen Punkt aus erreichbar – Graphen, die diese Eigenschaft aufweisen, werden *verbunden* (*connected*) genannt (andernfalls heißen sie *unverbunden*, *disconnected*).<sup>156</sup>

Jedem Weg lässt sich die Anzahl der Übergänge zwischen zwei Punkten, d.h. die Anzahl der Vorkommen von Linien, als seine *Länge* (*length*) zuordnen, was in Königsberg der Anzahl der Brückenüberquerungen gleichkommt (nicht: der Anzahl der verschiedenen überquerten Brücken, da einzelne Brücken auch mehrfach überquert werden können).<sup>157</sup> Die Bewegung von A nach B über die Brücke a entspricht also einem Weg der Länge 1. Wie viele Wege einer bestimmten Länge *n* es von einem Punkt zu einem anderen Punkt in einem Graphen gibt, lässt sich anhand der Adjazenzmatrix ermitteln, indem man diese *n*-mal mit sich selbst

---

von Punkt B gekennzeichnet. Die beschriebenen Bewegungsmöglichkeiten des Einwohners sind durch schwarze Pfeile indiziert.

<sup>154</sup> Die Terminologie ist uneinheitlich (sowohl im Englischen als auch im Deutschen). Einige Autoren unterscheiden zwischen *walks* (Punkte und Linien können mehrfach vorkommen), *trails* (Punkte können mehrfach vorkommen, aber alle Linien müssen verschieden sein) und *paths* (alle Punkte und alle Linien müssen verschieden sein), beispielsweise *Harary*, *Graph Theory*, 1969, S. 13; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 105; *Estrada und Knight*, *A First Course in Network Theory*, 2015, S. 18. Andere Autoren sprechen einheitlich von *paths* und geben etwaige Zusatzanforderungen gesondert an, so etwa *Newman*, *Networks*, 2010, S. 136; *Barabási*, *Network Science*, 2016, S. 56. In der deutschsprachigen Literatur wie hier *Brandes*, *Graphentheorie*, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 345 (352).

<sup>155</sup> *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 107; *Brandes*, *Graphentheorie*, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 345 (352).

<sup>156</sup> *Harary*, *Graph Theory*, 1969, S. 13; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 109. Gerichtete Graphen können *stark verbunden* (*strongly connected*) oder *schwach verbunden* (*weakly connected*) sein. Sie sind stark verbunden, sofern jeder ihrer Punkte von jedem anderen Punkt aus erreichbar ist, wenn man sich bei der Bewegung über den Graphen an die Pfeilrichtungen hält, und schwach verbunden, sofern jeder ihrer Punkte von jedem anderen Punkt aus erreichbar ist, wenn man bei der Bewegung auf dem Graphen die Pfeilrichtungen ignoriert. Graphen, die stark verbunden sind, sind also immer auch schwach verbunden, während das Gegenteil nicht der Fall ist. Zum Ganzen *Harary*, *Graph Theory*, 1969, S. 199; *Brandes und Erlebach*, *Fundamentals*, in: *Brandes und Erlebach* (Hrsg.), *Network Analysis*, 2005, S. 7 (9).

<sup>157</sup> *Harary*, *Graph Theory*, 1969, S. 13; *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 105; *Brandes*, *Graphentheorie*, in: *Stegbauer und Häußling* (Hrsg.), *Handbuch Netzwerkforschung*, 2010, S. 345 (352). Bei Graphen mit gewichteten Kanten ist es auch möglich, die Gewichtsinformationen in der Längendefinition zu berücksichtigen; dazu *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 144.

multipliziert (Abb. 3.16, S. 125). Die Zeilensummen bzw. Spaltensummen geben dann an, wie viele Wege der Länge  $n$  es von dem Punkt der betrachteten Zeile aus bzw. zu dem Punkt der betrachteten Spalte hin gibt (beispielsweise existieren von Punkt A aus 15 Wege der Länge 2).

$$\mathbf{A}^2 = \begin{pmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} = \begin{matrix} \text{A} \\ \text{B} \\ \text{C} \\ \text{D} \end{matrix} \begin{pmatrix} 9 & 1 & 1 & 4 \\ 1 & 5 & 5 & 2 \\ 1 & 5 & 5 & 2 \\ 4 & 2 & 2 & 3 \end{pmatrix}$$

Abbildung 3.16: Ermittlung der Anzahl der Wege mit Länge 2 für den Graphen zum Königsberger Brückenproblem<sup>158</sup>

Die Wege, die in einem Graphen möglich sind, lassen sich nach verschiedenen Kriterien kategorisieren, beispielsweise:

- nach Vorgaben zu den besuchten Punkten (oder auch Linien):  
Wege, die *von* einem bestimmten Punkt (z.B. Punkt B in Abb. 3.15 (a), S. 123) ausgehen; Wege, die *zu* einem bestimmten Punkt verlaufen;<sup>159</sup> Wege, die *durch* einen bestimmten Punkt verlaufen;
- nach Vorgaben zu ihrer Länge:  
Wege der Länge 1, Wege der Länge 2, ..., Wege unendlicher Länge;
- nach besonderen strukturellen Eigenschaften:  
Wege, die keine *Linie* mehrfach überqueren; Wege, auf denen kein *Punkt* mehrfach vorkommt;
- nach ihrem Berechnungsprozess:  
durch eindeutige Festlegung (deterministisch) oder als Zufallsläufe (*random walks*) (nichtdeterministisch);<sup>160</sup> oder

<sup>158</sup> Der Wert einer Zelle  $a_{ij}$  von  $\mathbf{A}^2$ , die in Zeile  $i$  und Spalte  $j$  steht, lässt sich errechnen, indem man die Elemente in Zeile  $i$  der ersten Matrix paarweise mit den Elementen in Spalte  $j$  der zweiten Matrix multipliziert und die jeweiligen Ergebnisse addiert. So ergibt sich beispielsweise der Wert 9 für Zelle  $a_{11}$ :  $0 \cdot 0 + 2 \cdot 2 + 2 \cdot 2 + 1 \cdot 1 = 9$ ; und der Wert 1 für Zelle  $a_{21}$ :  $2 \cdot 0 + 0 \cdot 2 + 0 \cdot 2 + 1 \cdot 1 = 1$ .

<sup>159</sup> Diese Kategorie ist neben der erstgenannten nur für gerichtete Graphen relevant, deren Pfeile lediglich in Pfeilrichtung überschritten werden können (Analogie aus dem Straßenverkehr: Einbahnstraßen). In ungerichteten Graphen ist die Menge dieser Wege identisch mit den Wegen, die von dem bezeichneten Punkt ausgehen, da die Linien in beide Richtungen überschritten werden können.

<sup>160</sup> Bei einem *Zufallslauf* (*random walk*) wird die Bewegung eines Zufallsläufers (*random walker*) über den Graphen dadurch simuliert, dass in jedem Schritt mit einer bestimmten Wahrscheinlichkeit eine Linie zufällig ausgewählt wird, die der Zufallsläufer dann überschreitet. Der Zufallslauf wird im Deutschen auch als Irrfahrt oder Zufallsbewegung bezeichnet.

- nach einer Kombination von mehreren Vorgaben:  
 Wege, die von einem bestimmten Punkt ausgehen (z.B. Punkt A in Abb. 3.15 (a), S. 123) und zu einem bestimmten Punkt verlaufen (z.B. Punkt B in Abb. 3.15 (a), S. 123); Wege, die von einem bestimmten Punkt ausgehen und eine bestimmte Länge haben; Wege, die in demselben Punkt enden, in dem sie anfangen (Zyklen, *cycles*).<sup>161</sup>

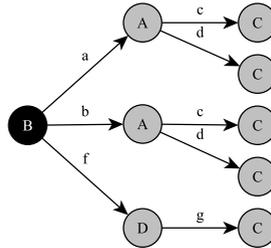


Abbildung 3.17: Kürzeste Wege von B nach C als Kürzeste-Wege-Baum<sup>162</sup>

Aus Effizienzgesichtspunkten besonders interessant sind die jeweils *kürzesten Wege* (*shortest paths, geodesics*) zwischen zwei Punkten: Wege, deren Länge der minimalen Länge aller existierenden Wege zwischen ihren Endpunkten entspricht.<sup>163</sup> Oft gibt es in einem Graphen mehrere kürzeste Wege zwischen zwei Punkten. In Königsberg ist beispielsweise der Weg von B nach C, der über die Linien a und c verläuft und die Länge 2 hat, ein kürzester Weg (es gibt keinen Weg von B nach C von geringerer Länge), ebenso der Weg, der über die Linien f und g

<sup>161</sup> *Estrada und Knight*, A First Course in Network Theory, 2015, S. 18; *Barabási*, Network Science, 2016, S. 60; *Zweig*, Network Analysis Literacy, 2016, S. 64. Manchmal wird für Zyklen zusätzlich gefordert, dass die Länge des Wegs  $\geq 3$  ist, so etwa *Harary*, Graph Theory, 1969, S. 13 und *Wasserman und Faust*, Social Network Analysis, 1994, S. 108, oder dass kein Punkt oder keine Linie mehrfach vorkommt, so etwa *Harary*, Graph Theory, 1969, S. 13 (für Punkte) und *Berge*, Graphes et Hypergraphes, 1970, S. 11 (für Linien). In der deutschsprachigen Literatur werden Zyklen oft als *Kreise* bezeichnet, so etwa bei *Volkman*, Fundamente der Graphentheorie, 1996, S. 10; *Brandes*, Graphentheorie, in: *Stegbauer und Häußling* (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 345 (352); *Krumke und Noltemeier*, Graphentheoretische Konzepte und Algorithmen, 2012, S. 31. Graphen ohne Zyklen heißen *azyklisch* (dazu bereits Kapitel 1 Fn. 18, S. 8).

<sup>162</sup> Dargestellt sind nur Wege der Länge 2, die im Punkt C enden. Insgesamt gibt es 15 Wege der Länge 2, die von Punkt B ausgehen, wie sich aus der Zeilensumme der zweiten Zeile in  $A^2$  ergibt.

<sup>163</sup> *Brandes*, Graphentheorie, in: *Stegbauer und Häußling* (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 345 (352); *Newman*, Networks, 2010, S. 139. Die minimale Länge eines Wegs zwischen zwei Punkten nennt man auch ihre Distanz (*[geodesic] distance*) bzw. ihren Abstand; *Wasserman und Faust*, Social Network Analysis, 1994, S. 110; *Brandes*, Graphentheorie, in: *Stegbauer und Häußling* (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 345 (352); *Zweig*, Network Analysis Literacy, 2016, S. 64. Auf einem kürzesten Weg kommt kein Punkt (und damit auch keine Linie) mehrfach vor.

verläuft, nicht aber der Weg über die Linien f, e und d (mit Zwischenstopps bei D und A), der die Länge 3 hat. Insgesamt gibt es von B nach C fünf verschiedene kürzeste Wege, wie man sich an einem Kürzeste-Wege-Baum (*shortest-path tree*) verdeutlichen kann (Abb. 3.17, S. 126).<sup>164</sup>

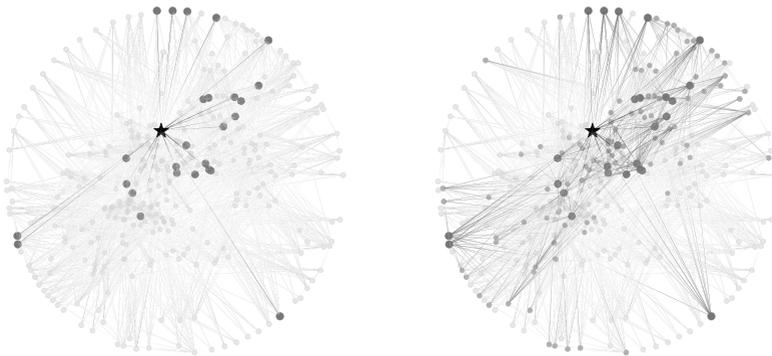
In vielen Situationen ist es sinnvoll, die Informationen zu den Wegen in einem Graphen anhand der oben angesprochenen Unterscheidungskriterien zusammenzufassen. Hierzu muss man zählen. Je nachdem, welche Eigenschaften die Wege aufweisen, die gezählt werden sollen, sind dabei verschiedene Vorgehensweisen denkbar. Für den Graphen im Königsberger Brückenproblem kann man beispielsweise fragen, wie viele verschiedene Wege der Länge 1 von B ausgehen. Zur Beantwortung dieser Frage ermittelt man schlicht die *Anzahl* der verschiedenen Wege, die in diesem Fall gleich der Anzahl der von B ausgehenden Linien ist: 3. Man kann aber auch kompliziertere Fragen stellen, etwa nach der Gesamtlänge der kürzesten Wege von B zu allen anderen Punkten. Dann ermittelt man für eine Antwort zunächst die *Länge* der jeweils kürzesten Wege und bildet aus den Längen die *Summe* – hier lautet die Antwort: 4. Ähnlich kann man die Frage nach der durchschnittlichen Länge der kürzesten Wege von B zu allen anderen Punkten beantworten, indem man die Längen nicht addiert, sondern ihren *Durchschnitt* bildet:  $\frac{4}{3}$ . Die maximale Anzahl von Brücken, die ein Bürger vom Landstück B aus überqueren muss, um zu einem beliebigen anderen Landstück zu gelangen, erhält man als *Maximum* der Längen aller kürzesten Wege, die von B ausgehen: 2. Die Komplexität lässt sich beliebig erhöhen. Für die Frage, wie viel *Prozent* aller kürzesten Wege zwischen zwei beliebigen Punkten durch einen bestimmten dritten Punkt oder über eine bestimmte Linie verlaufen, wird man (schon bei kleinen Graphen) regelmäßig einen Rechner bemühen. Dass Antworten auf ähnliche Fragen durchaus hilfreich sein können, wird sich im weiteren Verlauf dieser Arbeit noch zeigen. Insgesamt lassen sich die möglichen Differenzierungen bei der Wegerfassung in vier Dimensionen zusammenfassen, auf die im Folgenden immer wieder zurückgegriffen wird:<sup>165</sup>

- Typ der erfassten Wege (Wegtyp):  
Wege von bestimmter Länge; Wege, die sich nicht überschneiden; kürzeste Wege; ...
- Eigenschaft der Wege, die erfasst wird (Wegeigenschaft):  
Anzahl der Wege oder Länge der Wege;
- Position des bei der Erfassung betrachteten Punktes (Punktposition):  
Ausgangs- bzw. Endpunkt (radiale Erfassung) oder Durchgangspunkt (mediale Erfassung);
- Zusammenfassung der erfassten Wege (Zusammenfassungsmodus):  
keine Zusammenfassung, Summe, Durchschnitt, Maximum, ...

<sup>164</sup> Dieses Ergebnis lässt sich auch der Zelle  $a_{23}$  von  $A^2$  entnehmen (Abb. 3.16 (b), S. 125).

<sup>165</sup> Diese Unterscheidung orientiert sich an *Borgatti und Everett*, *Social Networks* 28 (2006), S. 466 (476).

Dass es von Königsberg zum Recht nur ein kleiner Schritt ist, verdeutlicht ein Blick auf das Zitiernetzwerk der flatländischen Rechtsprechung und den ihm entsprechenden Zitiergraphen: Angenommen, ein Jurist in Flatland benutzt zur Recherche von Fallrecht die Suchfunktion einer juristischen Datenbank, die ihm (nur) die höchstrichterliche Rechtsprechung zugänglich macht, wobei zu jeder höchstrichterlichen Entscheidung die zitierten und zitierenden höchstrichterlichen Entscheidungen annotiert sind. Diese Annotationen eröffnen dem Juristen die Möglichkeit, sich auf einem Teil des gesamten Graphen, genannt Subgraph (*subgraph*),<sup>166</sup> zu bewegen – und zwar so, als wenn der Zitiergraph ungerichtet wäre, also unabhängig von der Richtung der Zitate.<sup>167</sup> Ausgehend von einer beliebigen Entscheidung kann dieser Jurist mit wenigen (Maus-)Klicks eine Vielzahl von Entscheidungen erreichen – jedenfalls dann, wenn die Entscheidung, die er betrachtet, in das Fallrecht hinreichend eingebettet ist (Abb. 3.18, S. 128). Der Datenbankanbieter könnte den Juristen bei der Recherche noch weiter unterstützen, indem er die Quantifizierungsmethoden der juristischen Netzwerkforschung fruchtbar macht (dazu sogleich).



(a) Reichweite mit einem Klick

(b) Reichweite mit zwei Klicks

Abbildung 3.18: Recherche in der Rechtsprechung des *High Court* als Bewegung über den Zitiergraphen<sup>168</sup>

<sup>166</sup> Harary, Graph Theory, 1969, S. 11; Brandes und Erlebach, Fundamentals, in: Brandes und Erlebach (Hrsg.), Network Analysis, 2005, S. 7 (9). Das Gegenstück eines Subgraphen in der realen Welt wird hier als *Subnetzwerk* (*subnetwork*) bezeichnet. Die Definition von Subgraphen und Subnetzwerken setzt stets bei einem umschließenden Graphen bzw. Netzwerk an; sie ist also relativ.

<sup>167</sup> Mit anderen Worten: Der Jurist bewegt sich auf dem Graphen, den man erhält, wenn man beim Zitiergraphen die Pfeilrichtungen weglässt und diesen so in einen ungerichteten Graphen umwandelt.

## II. Mikroebene

### 1. Problemstellung: Was ist wichtig?

Wie im zweiten Kapitel bereits ausgeführt, geht es bei der Untersuchung eines Netzwerks auf der Mikroebene vor allem darum, einzelne Netzwerkelemente zu identifizieren, die besonders wichtig sind (Wichtigkeitsfrage).<sup>169</sup> Hierzu muss man einerseits eine Definition von Wichtigkeit festlegen, die auf das betrachtete Netzwerk und das Erkenntnisinteresse des Untersuchenden passt (*Definitionsproblem*). Andererseits muss man die gewählte Wichtigkeitsdefinition operationalisieren, so dass die Wichtigkeit der einzelnen Netzwerkelemente definitionsgemäß gemessen werden kann (*Messproblem*).<sup>170</sup>

Bei der Wichtigkeitsdefinition kann sich der Untersuchende in zwei verschiedenen Positionen befinden. Einerseits kann er das Netzwerk situationsunabhängig als Außenstehender analysieren und ohne zusätzliche Informationen möglichst allgemeine Aussagen zur Wichtigkeit einzelner Netzwerkelemente treffen wollen (*abstraktes Erkenntnisinteresse*, Beobachterperspektive). Andererseits kann er das Netzwerk in einer bestimmten Situation betrachten, in der ihm zusätzliche Informationen vorliegen, die in die Bestimmung der Wichtigkeit einfließen sollen (*konkretes Erkenntnisinteresse*, Teilnehmerperspektive). Den Unterschied zwischen abstraktem und konkretem Erkenntnisinteresse verdeutlicht ein Blick auf das Zitiernetzwerk der flatländischen Rechtsprechung (S. 88–91): Der Forscher, der versucht, die Rechtsentwicklung in Flatland in der ersten Dekade des 21.

<sup>168</sup> Die Abbildungen zeigen alle Entscheidungen des flatländischen *High Court* (von 2000 bis einschließlich 2009) und die Zitierbeziehungen zwischen diesen Entscheidungen (ohne Differenzierung nach Zitierhäufigkeit) grundsätzlich als kleine hellgraue Punkte bzw. dünne hellgraue Pfeile. Entscheidungen, die von der mit dem schwarzen Stern markierten Entscheidung aus mit einem Klick (Abb. 3.18 (a), S. 128) bzw. mit zwei Klicks (Abb. 3.18 (b), S. 128) erreicht werden können, sind dunkelgrau gekennzeichnet; Entscheidungen, die mit einem Klick erreichbar sind, zusätzlich als größere Punkte dargestellt. Zitierbeziehungen, welche zur Erreichbarkeit mit dem jeweiligen Klick führen, sind dunkelgrau hervorgehoben. Das Layout wurde wieder mit dem kräftebasierten Fruchterman-Reingold-Algorithmus (Fn. 119, S. 113) und *seed 1234* errechnet; die mit dem Stern markierte Ausgangsentscheidung (Aktenzeichen H-HC-0:2-2005-181) mit demselben *seed* zufällig ausgewählt. Die Pfeile verlaufen (wie Nadeln) von ihrem dickeren, dunkleren Ende zu ihrem dünneren, helleren Ende – eine Darstellung, die ungewohnt sein mag, aber der Zeichnung „klassischer“ Pfeile jedenfalls dann vorzuziehen ist, wenn deren Spitzen viele andere Elemente des Graphen verdecken würden (*occlusion*). Siehe dazu die Ergebnisse der Experimente von *Holten und van Wijk*, *Proceedings of SIGCHI 2009*, S. 2299–2308; aufgegriffen von *Ware*, *Information Visualization*, 2013, S. 226.

<sup>169</sup> Dazu oben S. 72.

<sup>170</sup> Aus Sicht des einzelnen Forschers ist auch das umgekehrte Vorgehen grundsätzlich denkbar: Er kann zuerst ein bereits von anderen definiertes Maß anwenden und sich dann fragen, welcher Definition von Wichtigkeit die erhaltenen Messwerte für das untersuchte Netzwerk entsprechen könnten. Das birgt allerdings die Gefahr, dass man die Wichtigkeitsdefinition, die dem verwendeten Maß zugrunde liegt, übernimmt, ohne ihre Eignung im vorliegenden Fall kritisch zu prüfen.

Jahrhunderts für die Nachwelt zusammenzufassen, wird sich für andere Entscheidungen interessieren als der Richter, dem ein Fall vorgelegt wird, den er der Inhaltsziffer 3 zuordnet (ihn interessieren möglicherweise vor allem Entscheidungen mit derselben Inhaltsziffer), oder der Anwalt, der für das Rechtsproblem seines Mandanten bereits eine Entscheidung des *High Court* als Leitentscheidung identifiziert hat. Richter und Anwalt verfügen über kontextspezifische Informationen, die sie in ihrer Definition von Wichtigkeit berücksichtigen können. Ihre konkrete Definition wird daher besser auf ihre jeweilige Situation zugeschnitten, gleichzeitig aber für andere Situationen unter Umständen weniger brauchbar sein als die abstrakte Definition des Forschers.

Aus juristischer Perspektive scheint es daher sinnvoll, bei der Wichtigkeitsdefinition zwischen *abstrakter Wichtigkeit* (Wichtigkeit bei abstraktem Erkenntnisinteresse) und *konkreter Wichtigkeit* (Wichtigkeit bei konkretem Erkenntnisinteresse) zu differenzieren. Die Diskussion zur Wichtigkeit in der juristischen Netzwerkforschung bezieht sich größtenteils auf die Definition und Messung von abstrakter Wichtigkeit; wie konkrete Wichtigkeit sich definieren und messen lässt, wird kaum analysiert. Die nachfolgende Darstellung konzentriert sich daher zunächst auf diejenigen Lösungsansätze zum Wichtigkeitsproblem, die ein abstraktes Erkenntnisinteresse zugrunde legen. Dabei steht die Definition und Messung der Wichtigkeit von Objekten anhand von Beziehungen im Mittelpunkt, die im Schrifttum auf der Mikroebene sehr viel mehr Aufmerksamkeit erfährt als die Definition und Messung der Wichtigkeit von Beziehungen. Auf die Messung der Wichtigkeit von Beziehungen wird auf der Mesoebene zurückzukommen sein.<sup>171</sup> Möglichkeiten zur Wichtigkeitsdefinition bei Annahme eines konkreten Erkenntnisinteresses werden im vierten Kapitel anhand realer Daten zur Rechtsprechung des BVerfG thematisiert.<sup>172</sup>

## 2. Lösungsansätze

Die Maße, die abstrakte Wichtigkeitsdefinitionen operationalisieren, bezeichnet man als Zentralitätsmaße (*centrality measures*), kurz Zentralitäten (*centralities*).<sup>173</sup> Sie ordnen den einzelnen Punkten in einem Graphen Zahlen als *Zentralitätswerte* (*centrality values*) zu, anhand derer die Punkte in eine Reihenfolge gebracht und die „wichtigsten“ Objekte des dargestellten Netzwerks identifiziert werden können.<sup>174</sup> Dabei hat das Größenverhältnis der Werte zueinander in der

<sup>171</sup> Zur Analyse eines Netzwerks auf der Mesoebene S. 142–165.

<sup>172</sup> Dazu S. 278–281.

<sup>173</sup> *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 169; *Newman*, *Networks*, 2010, S. 168.

<sup>174</sup> Die Maße sind darauf ausgerichtet, im oberen Bereich der Wichtigkeit scharf zu messen, womit verbunden ist, dass ihre Werte im mittleren und unteren Bereich in der Regel wenig aussagekräftig sind. Mithilfe der meisten Zentralitäten lassen sich daher beispielsweise die definitionsgemäß „unwichtigsten“ Objekte in einem Netzwerk nicht zuverlässig identifizieren.

Regel keine Bedeutung.<sup>175</sup> Wie die Zentralitätswerte berechnet werden, hängt von der Wichtigkeitsdefinition ab, die einem Maß zugrunde gelegt wird. Dennoch hat sich ein gewisses Standardrepertoire von Maßen herausgebildet, die mit gängigen Softwarelösungen errechnet werden können und die in Publikationen zur Netzwerkforschung im Recht immer wieder auftauchen. Diese Maße wurden zwar in anderen Kontexten geboren; sie lassen sich aber auch anhand rechtsspezifischer Überlegungen zur Wichtigkeitsdefinition entwickeln.

#### a) Gradzentralitäten

Das einfachste der gängigen Zentralitätsmaße ist ohne nähere Erläuterung bereits mehrfach zum Einsatz gekommen – unter anderem, als ein mögliches Vorgehen bei der juristischen Recherche als Bewegung auf einem Zitiergraphen dargestellt wurde.<sup>176</sup> Dieses Maß heißt Gradzentralität (*degree centrality*), kurz Grad (*degree*), und entspricht der Anzahl der (direkten) Verbindungen, an denen ein Punkt beteiligt ist.<sup>177</sup> Die in Abbildung 3.18 (a) (S. 128) mit dem Stern markierte Ausgangsentscheidung (Aktenzeichen H-HC-0:2-2005-181) ist im dargestellten Subnetzwerk an insgesamt 26 Beziehungen beteiligt – sie hat also den Grad 26.<sup>178</sup> Die 26 Beziehungen der betrachteten Entscheidung lassen sich weiter in eingehende Zitate (die Entscheidung wird zitiert, Passiv) und ausgehende Zitate (die Entscheidung zitiert, Aktiv) unterscheiden, woraus man für Netzwerke, die als gerichtete Graphen dargestellt werden können, zwei weitere Zentralitätsmaße ableiten kann: den Innengrad, im Deutschen auch Eingangsgrad (*indegree*; Innengradzentralität, *indegree centrality*) für die Anzahl der eingehenden Zitate und den Außengrad, im Deutschen auch Ausgangsgrad (*outdegree*; Außengradzentralität, *outdegree centrality*) für die Anzahl der ausgehenden Zitate.<sup>179</sup> Die im Mittelpunkt von Abbildung 3.18 (a) (S. 128) stehende Entscheidung zitiert selbst 16 höchstrichterliche Entscheidungen und wird von 10 höchstrichterlichen Entscheidungen zitiert. Sie hat (im auf die höchstrichterlichen Entscheidungen eingeschränkten Zitiernetzwerk) also einen Innengrad von 10 und einen Außengrad von 16 – vorausgesetzt, dass man sich (wie für Abb. 3.18 [S. 128] angenommen)

<sup>175</sup> Mit anderen Worten: Zentralitätswerte sollten grundsätzlich als ordinale und nicht als metrische Merkmale der Objekte behandelt werden, denen sie zugeordnet sind.

<sup>176</sup> Dazu S. 123–129.

<sup>177</sup> *Borgatti* und *Everett*, *Social Networks* 28 (2006), S. 466 (468); *Newman*, *Networks*, 2010, S. 169. Sofern man mit der Adjazenzmatrix des Graphen arbeitet, kann man den Grad eines Punktes ermitteln, indem man die Werte der Zeile aufsummiert, in der er steht.

<sup>178</sup> Der dort abgebildete Graph ist ein auf die höchstrichterlichen Entscheidungen eingeschränkter Subgraph des flatländischen Zitiergraphen (S. 88–91), bei dem die gerichteten Mehrfachkanten zu einem ungewichteten Pfeil zusammengefasst wurden.

<sup>179</sup> *Newman*, *Networks*, 2010, S. 169. Arbeitet man auf der Adjazenzmatrix des gerichteten Graphen, die in der Regel nicht symmetrisch ist (dazu bereits oben bei Abb. 3.13, S. 120), so erhält man den Innengrad eines Punktes als Summe der Werte in der *Spalte*, den Außengrad eines Punktes als Summe der Werte in der *Zeile*, in welcher der Punkt steht.

nur dafür interessiert, *ob* eine Entscheidung zitiert wird, und nicht registriert, *wie häufig* dies geschieht.

Hier zeigt sich die bereits angesprochene Differenzierung zwischen *binärer Zählung* und *gewichteter Zählung* von Zitaten (*Zählmodus*).<sup>180</sup> Bei binärer Zählung von Zitaten ermittelt man für jede Entscheidung lediglich, *ob* diese eine andere Entscheidung zitiert bzw. von dieser zitiert wird. Bei gewichteter Zählung hält man zusätzlich fest, *wie häufig* eine Entscheidung eine andere zitiert bzw. von einer anderen Entscheidung zitiert wird. Letzteres führt dazu, dass sich jede Entscheidung sowohl aktiv als auch passiv in jeweils zwei Dimensionen evaluieren lässt: Sie kann von wenigen oder von vielen verschiedenen Entscheidungen zitiert werden (*passive Zitierbreite*) und sie kann von zitierenden Entscheidungen jeweils wenig oder viel zitiert werden (*passive Zitiertiefe*); ebenso kann sie wenige oder viele verschiedene Entscheidungen zitieren (*aktive Zitierbreite*) und zitierte Entscheidungen jeweils wenig oder viel zitieren (*aktive Zitiertiefe*).

So kann man für jede Entscheidung ein detailliertes *Zitierprofil* erstellen, das unter anderem die Verhältnisse zwischen den verschiedenen Gradzentralitäten bei binärer und gewichteter Zählung sowie verschiedene Übersichten zu Zitierbreite und Zitiertiefe enthalten kann (Abb. 3.19, S. 133). Möglicherweise gehen unterschiedliche Zitierprofile mit unterschiedlichen Entscheidungstypen oder -inhalten einher, sodass man von Ersteren auf Letztere schließen kann. Untersuchen lässt sich diese Hypothese nur, wenn man Zitierdaten mit gewichteter Zählung erfasst und Zitiernetzwerke als Graphen mit Mehrfachkanten beschreibt, wie es für die fiktiven Rechtsprechungsdaten aus Flatland angenommen wurde.

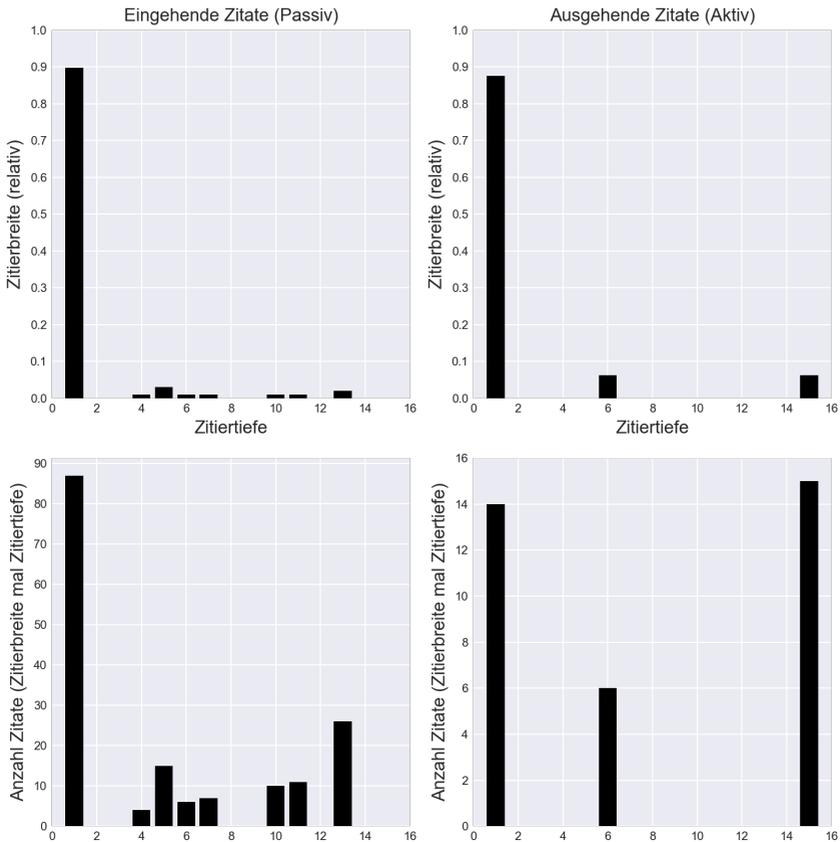
Innengrad, Außengrad und Grad lassen sich zusammenfassend als *Gradzentralitäten* bezeichnen. In der Terminologie der Bewegung auf Graphen (oben S. 123–129) kann man sie ermitteln als Anzahl (Wegeigenschaft) der Wege mit Länge 1 (Wegtyp), bei denen ein Punkt Ausgangs- oder Endpunkt ist (Punktposition), ohne weitere Zusammenfassung (Zusammenfassungsmodus).<sup>181</sup>

<sup>180</sup> Dazu schon S. 76–79. Die gewichtete Zählung wird in der juristischen Netzwerkforschung unter anderem genutzt von Clark, *Lauderdale und Katz*, *Political Analysis* 20 (2012), S. 329 (333–334) und Charlotin, *Journal of International Economic Law* 20 (2017), S. 279 (286–287). Zahlreiche Beispiele für die binäre Zählung finden sich in Fn. 11 (S. 78).

<sup>181</sup> Anhand dieser Terminologie lassen sich auch mögliche Modifikationen identifizieren. Beispielsweise kann man statt Wegen der Länge 1 Wege der Länge 2 erfassen, sodass der Zentralitätswert eines Punktes nicht nur von seinem Grad, sondern auch vom Grad seiner Nachbarn abhängt.

	<i>binär</i>	<i>gewichtet</i>	<i>Ratio</i>
<i>Innengrad</i>	97	166	0,58
<i>Außengrad</i>	16	35	0,46
<i>Grad</i>	113	201	0,48

(a) Binäre und gewichtete Zählung



(b) Aktive und passive Zitierbreite und Zitertiefe

Abbildung 3.19: Zitierprofil der Entscheidung H-HC-0:2-2005-181 auf Basis des gesamten Zitiergraphen

Dass die einzelnen Gradzentralitäten ganz unterschiedliche Entscheidungen als „die Wichtigsten“ erfassen und die Ergebnisse auch davon abhängen, ob man binär oder gewichtet zählt, zeigt exemplarisch ein Blick in die Rechtsprechungsdaten: Schränkt man das Zitiernetzwerk auf höchstrichterliche Entscheidungen der Inhaltsziffern 0, 1 und 2 ein und berechnet für die übrig gebliebenen 108

Entscheidungen Innengrad, Außengrad und Grad jeweils auf dem Graphen mit Mehrfachkanten (gewichtete Zählung) und auf dem Graphen ohne Mehrfachkanten (binäre Zählung), so offenbaren sich deutliche Differenzen (Abb. 3.20, S. 134) – sowohl beim Vergleich der einzelnen Maße (Innengrad, Außengrad und Grad) als auch beim Vergleich der Zählmodi (binäre und gewichtete Zählung). Nun betrifft das Beispiel zwar Flatland, doch ein ähnlicher Befund wäre auch in einem echten Staat nicht verwunderlich: Entscheidungen, die allgemein viel zitieren, sind vielleicht „wichtig“ im Sinne von „schwer“ oder „kontrovers“, Entscheidungen, die allgemein viel zitiert werden, eher „wichtig“ im Sinne von „einflussreich“. Entscheidungen, die eher breit zitieren oder zitiert werden, behandeln möglicherweise Querschnittsfragen (z.B. Prozedurales); Entscheidungen, die eher tief zitieren oder zitiert werden, sind inhaltlich vielleicht besonders fokussiert.<sup>182</sup>

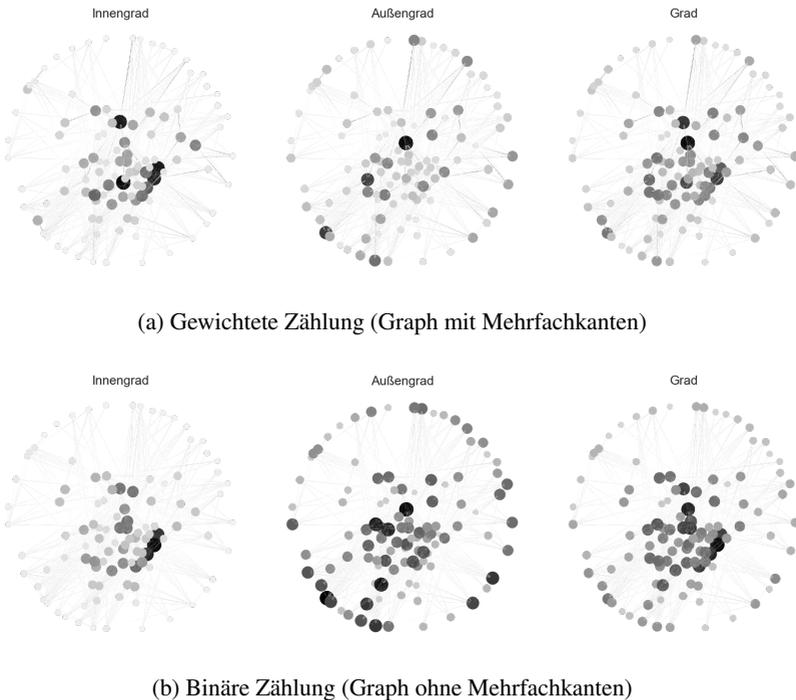


Abbildung 3.20: Innengrad, Außengrad und Grad für höchstrichterliche Entscheidungen mit Inhaltswerten 0, 1 oder 2<sup>183</sup>

<sup>182</sup> Die tiefe Zitierung birgt auch Potenzial im Rahmen eines konkreten Erkenntnisinteresses; dazu noch S. 278–281. Zu möglichen Unterschieden zwischen binärer und gewichteter Zählung bereits *Coupette und Fleckner*, JZ 73 (2018), S. 379 (384).

<sup>183</sup> Der in Abbildung 3.20 (a) (S. 134) dargestellte Graph wurde aus dem Zitiergraphen produziert, indem dieser auf Punkte eingeschränkt wurde, die höchstrichterliche Entscheidungen mit

## b) Rückkopplungszentralitäten

Was an den Gradzentralitäten unbefriedigend erscheint, ist die Tatsache, dass alle Beziehungen gleich behandelt werden. Soweit sich die Unzufriedenheit mit der Gleichbehandlung auf nicht-relationale Attribute der Objekte bezieht (Beispiel: „Soll das Zitat eines Untergerichts wirklich so viel wiegen wie das des höchsten Gerichts?“), kann man das Problem lösen, indem man die Zitate in Abhängigkeit von den jeweiligen Merkmalen ihrer Urheber gewichtet (Beispiel: Zitate des höchsten Gerichts werden dreifach gezählt, jene der mittleren Gerichte zweifach und jene der Untergerichte einfach). Doch was, wenn man die Wichtigkeit einer Entscheidung in Abhängigkeit von der Wichtigkeit der Entscheidungen bestimmen will, die sie zitieren (Passiv) oder die sie zitiert (Aktiv)? Dann wünscht man sich wahrscheinlich ein Maß, nach dem der Zentralitätswert eines Punktes proportional zu den Zentralitätswerten seiner Nachbarn ist.<sup>184</sup> Das lässt sich durch einschrittige Berechnungen, wie sie den zuerst vorgestellten Gradzentralitäten zugrunde liegen, nicht erreichen. Stattdessen bedarf es eines Rechenmechanismus, bei dem der Zentralitätswert der Nachbarn eines Punktes auf den Zentralitätswert dieses Punktes zurückwirkt. Maße, die einen solchen Mechanismus beinhalten, werden hier als *Rückkopplungszentralitäten* (*feedback centralities*) bezeichnet.<sup>185</sup>

Der einfachste Ansatz zur Konstruktion einer Rückkopplungszentralität besteht darin, mit irgendeinem Zentralitätswert (zum Beispiel 1) für jeden Punkt zu beginnen und diesen Wert in mehreren Berechnungsrunden jeweils in Abhängigkeit von den Zentralitätswerten der benachbarten Punkte aus der vorangehenden Runde zu aktualisieren.<sup>186</sup> Wenn man das häufig genug tut, so die Idee, pendeln sich die so erhaltenen Zentralitätswerte der Punkte irgendwann ein und weisen dann die gewünschte Proportionalitätseigenschaft auf. Dass das zumeist der Fall ist, kann man mithilfe der Darstellung von Graphen als Adjazenzmatrix mathematisch beweisen.<sup>187</sup> Das Zentralitätsmaß, das man so erhält, heißt *Eigenvektorzentralität*

Inhaltsziffer 0, 1 oder 2 darstellen (also indem alle Punkte, welche die Selektionskriterien nicht erfüllen, aus dem Graphen entfernt wurden); er enthält 108 Punkte und 984 Linien (darunter auch Mehrfachkanten, was der gewichteten Zählung von Zitaten entspricht). Der in Abbildung 3.20 (b) (S. 134) dargestellte Graph wurde analog aus dem Zitiergraphen produziert, bei dem zuvor die Mehrfachkanten durch einfache Linien ersetzt wurden, was der binären Zählung von Zitaten entspricht. Das Layout für beide Graphen wurde mit dem Fruchterman-Reingold-Algorithmus (Fn. 119, S. 113) und *seed 1234* errechnet. Dabei wurde jeweils der in Abbildung 3.20 (a) (S. 134) abgebildete Graph zugrunde gelegt, um die Vergleichbarkeit von (a) und (b) sicherzustellen. Die Werte der Gradzentralitäten wurden mit *igraph-0.7.0* für *python-2.7* ermittelt, normalisiert (d.h. auf Werte zwischen 0 und 1 abgebildet) sowie in Farbe und Größe der einzelnen Punkte kodiert (zu Einzelheiten siehe den Online-Appendix). Die Entscheidungen mit den relativ höchsten Zentralitätswerten sind dabei als (relativ) große schwarze Punkte dargestellt.

<sup>184</sup> Mathematisch ausgedrückt sucht man nach einem Maß, sodass gilt:  $x_i = k \cdot \sum_j (a_{ij} \cdot x_j)$ , wobei  $x_i$  bzw.  $x_j$  Zentralitätswerte der Punkte  $i$  bzw.  $j$  sind,  $k$  eine Konstante ist und  $a_{ij}$  das Element in Zeile  $i$  und Spalte  $j$  der Adjazenzmatrix darstellt ( $\sum_j (a_{ij} \cdot x_j)$  entspricht also der Summe der Zentralitäten aller Nachbarn von Punkt  $i$ ).

<sup>185</sup> Von *feedback centralities* sprechen unter anderem Koschützki u.a., *Centrality Indices*, in: Brandes und Erlebach (Hrsg.), *Network Analysis: Methodological Foundations*, 2005, S. 16 (46); von *Rückkopplungszentralitäten* auch Brandes, Kosub und Nick, *Was messen Zentralitätsindizes?*, in: Hennig und Stegbauer (Hrsg.), *Die Integration von Theorie und Methode in der Netzwerkforschung*, 2012, S. 33 (42, 44).

<sup>186</sup> Newman, *Networks*, 2010, S. 169.

<sup>187</sup> Dazu Newman, *Networks*, 2010, S. 170.

(*eigenvector centrality*),<sup>188</sup> was damit zusammenhängt, dass seine Werte proportional zum dominanten Eigenvektor der Adjazenzmatrix sind.<sup>189</sup> Grundsätzlich kann man die Eigenvektorzentralität sowohl für ungerichtete Graphen als auch für gerichtete Graphen berechnen. Bei gerichteten Graphen muss man sich entscheiden, ob die Zentralität eines Punktes proportional zu den Werten der Punkte sein soll, die auf ihn zeigen (linke Eigenvektorzentralität), oder zu jenen der Punkte, auf die er zeigt (rechte Eigenvektorzentralität).<sup>190</sup> In einem Rechtsprechungsnetzwerk entspricht dies der Festlegung, ob die Wichtigkeit einer Entscheidung proportional zur Wichtigkeit der Entscheidungen sein soll, von denen sie zitiert wird (dann linke Eigenvektorzentralität), oder proportional zur Wichtigkeit der Entscheidungen, die durch sie zitiert werden (dann rechte Eigenvektorzentralität). Bei näherer Betrachtung erweist sich die Eigenvektorzentralität aber insbesondere für gerichtete Graphen ohne Zyklen als unbrauchbar: Unabhängig davon, ob man linke oder rechte Eigenvektorzentralität zugrunde legt, erhalten dort alle Punkte den Zentralitätswert 0.<sup>191</sup> Da juristische Zitiernetzwerke in der Regel als gerichtete Graphen repräsentiert werden, die aufgrund der Zeitabhängigkeit von Zitaten annähernd azyklisch sind,<sup>192</sup> ist so ein wesentlicher Teil der juristi-

<sup>188</sup> Newman, Networks, 2010, S. 169–170 (mit Verweis auf Bonacich, American Journal of Sociology 92 (1987), S. 1170–1182 als Erstpublikation); entwickelt bereits in Bonacich, Sociological Methodology 4 (1972), S. 176 (180–181); Bonacich, Journal of Mathematical Sociology 2 (1972), S. 113 (113–115).

<sup>189</sup> Zur Darstellung von Graphen als Adjazenzmatrix bereits oben bei Abb. 3.12 (S. 119) und Abb. 3.16 (S. 125). Ein Eigenvektor einer quadratischen Matrix  $\mathbf{A}$  mit  $n$  Zeilen und  $n$  Spalten ist ein Spaltenvektor  $v$  (eine Matrix mit  $n$  Zeilen und exakt einer Spalte), der nicht nur Nullen enthält und die Gleichung  $\mathbf{A} \cdot v = \lambda \cdot v$  für eine Konstante  $\lambda$  erfüllt, die als Eigenwert zum Eigenvektor  $x$  bezeichnet wird. Der betragsmäßig größte Eigenwert ( $\lambda_1$ ) heißt dominanter Eigenwert und der dominante Eigenvektor ist der Eigenvektor, der dem dominanten Eigenwert zugeordnet ist. Die Werte der Eigenvektorzentralität ergeben sich dann als  $x_i = \frac{1}{\lambda_1} \cdot \sum_j (a_{ij} \cdot x_j)$ ; siehe auch Newman, Networks, 2010, S. 170 (mit etwas anderer Notation).

<sup>190</sup> Da die (quadratische) Adjazenzmatrix  $\mathbf{A}$  (mit  $n$  Zeilen und Spalten) eines gerichteten Graphen im Allgemeinen nicht symmetrisch ist, gibt es neben Spaltenvektoren  $v$ , welche die Gleichung  $\mathbf{A} \cdot v = \lambda \cdot v$  für eine Konstante  $\lambda$  erfüllen, auch Zeilenvektoren (Matrizen mit einer Zeile und  $n$  Spalten)  $u$ , welche die Gleichung  $u \cdot \mathbf{A} = \kappa \cdot u$  für eine Konstante  $\kappa$  erfüllen. Diese Vektoren nennt man linke Eigenvektoren (weil sie von links mit  $\mathbf{A}$  multipliziert werden), während die Vektoren  $v$  rechte Eigenvektoren heißen. Für Zentralitätswerte, die zu den Werten der Vorgänger eines Punktes proportional sind, verwendet man die linken Eigenvektoren, im gegenteiligen Fall die rechten Eigenvektoren; Estrada und Knight, A First Course in Network Theory, 2015, S. 162 (gegenteilig Newman, Networks, 2010, S. 171, auf Basis der gegenteiligen Definition der Adjazenzmatrix, a.a.O. S. 114).

<sup>191</sup> Der Grund ist, dass Punkte mit Innengrad (rechts) bzw. Außengrad (links) 0 nach der Proportionalitätsanforderung den Zentralitätswert 0 erhalten müssen, womit deren Nachbarn ebenfalls eine Zentralität von 0 erhalten müssen, und ebenso die Nachbarn dieser Nachbarn, ...

<sup>192</sup> Ein gerichteter Graph ist azyklisch, wenn er keine gerichteten Zyklen enthält, d.h. wenn es in ihm keine Folge von Pfeilen gibt, durch deren Nachzeichnen entlang der Pfeilrichtungen man am Ende wieder zum Ausgangspunkt gelangen könnte; Harary, Graph Theory, 1969, S. 200. Zur Definition für ungerichtete Graphen bereits in Kapitel 1 Fn. 18 (S. 8).

schen Netzwerkforschung aus dem Anwendungsbereich der Eigenvektorzentralität ausgeschlossen.

Die Schwierigkeiten der Eigenvektorzentralität für gerichtete Graphen lassen sich abmildern, indem man jedem Punkt einen kleinen Betrag zusätzlicher Zentralität „schenkt“, der von seiner Einbindung in den Graphen unabhängig ist.<sup>193</sup> Wendet man das so erhaltene Maß (in seiner linken Version) auf den Graphen eines juristischen Zitiernetzwerks an, so zeigt sich allerdings, dass eine einzige „wichtige“ Entscheidung ihre hohe Zentralität auf alle Entscheidungen überträgt, die sie zitiert – unabhängig davon, wie viele das sind. Folgt man daher der Auffassung, dass das Zitat einer Entscheidung „mehr wert“ ist, wenn weniger andere Entscheidungen zitiert werden, so kann man fordern, dass die Zentralität eines Punktes proportional sein soll zur Summe der Zentralitäten aller Punkte, die auf ihn zeigen, *jeweils geteilt durch deren Außengrad* (bzw. durch 1, falls der Außengrad 0 ist).<sup>194</sup> Dies führt auf den *PageRank*, den die Entwickler von Google aufgrund ähnlicher Erwägungen zu Verlinkungen auf Webseiten ihrem Ranking von Suchergebnissen zugrunde legten.<sup>195</sup> Der *PageRank* lässt sich auch als Modell eines Zufallssurfers (*random surfer*) interpretieren, der sich durch die Linkstruktur des *World Wide Web* klickt und mit einer gewissen Wahrscheinlichkeit eine Seite ansteuert, die mit der zuvor besuchten nicht über einen Link verbunden ist.<sup>196</sup> Man kann sich daher fragen, inwieweit dieses Modell im juristischen Kontext angepasst werden kann, um das juristische Rechercheverhalten besser abzubilden, ohne die Orientierung an der Struktur des Zitiernetzwerks aufzugeben.<sup>197</sup>

In der Literatur zur Netzwerkforschung im Recht wird der *PageRank* vor allem bei der Messung der Wichtigkeit von Gerichtsentscheidungen in juristischen Zitier-

<sup>193</sup> Man gelangt so zu einem Maß, das (nach seinem Urheber) *Katz centrality* genannt wird und das zeitlich vor der Eigenvektorzentralität vorgeschlagen wurde (*Katz*, *Psychometrika* 18 (1953), S. 39–43); dazu zusammenfassend *Newman*, *Networks*, 2010, S. 172–174.

<sup>194</sup> Mathematisch:  $x_i = \alpha \cdot \sum_j (a_{ij} \frac{x_j}{d_j^{out}}) + \beta$ . Dabei sind  $x_i$  und  $x_j$  die Zentralitäten der Punkte  $i$  bzw.  $j$ ,  $a_{ij}$  ist das Element in Zeile  $i$ , Spalte  $j$  der Adjazenzmatrix,  $d_j^{out}$  der Außengrad des Punktes  $j$  (bzw. 1, falls der Außengrad 0 ist).  $\alpha$  und  $\beta$  sind Konstanten, die – in Grenzen, die unter anderem vom dominanten Eigenwert abhängen – frei gewählt werden können.

<sup>195</sup> Der *PageRank* wird erstmals beschrieben in *Brin und Page*, *Computer Networks and ISDN Systems* 30 (1998), S. 107–117, insbes. S. 109–110; heute ist er nur noch einer von vielen Faktoren, die das Ranking der Suchmaschine beeinflussen.

<sup>196</sup> Die Wahrscheinlichkeit, dass sich der Zufallssurfer langweilt, kommt in der Formel von Fn. 194 (S. 137) im Parameter  $\alpha$  (bei *Brin und Page*, *Computer Networks and ISDN Systems* 30 (1998), S. 107 (110):  $d$ ) zum Ausdruck;  $\beta$  entspricht dann  $1 - \alpha$  (bei *Brin und Page*, *Computer Networks and ISDN Systems* 30 (1998), S. 107 (110):  $1 - d$ ). Der *PageRank* definiert eine Wahrscheinlichkeitsverteilung, sodass sich seine Werte auf 1 summieren. Das bedingt, dass die *PageRanks* für einzelne Punkte sehr klein sind (für die Identifikation „wichtiger“ Punkte kommt es aber ohnehin nur auf ihre Reihenfolge an). Zusammenfassend zum *PageRank* und seinen möglichen Erweiterungen *Easley und Kleinberg*, *Networks, Crowds, and Markets*, 2010, S. 358–363. v

<sup>197</sup> Dazu noch S. 212–216.

netzwerken verwendet.<sup>198</sup> Zur Analyse von juristischen Netzwerken mit gerichteten Beziehungen kommen aber noch zwei weitere Rückkopplungszentralitäten zum Einsatz, die ebenfalls für das *World Wide Web* entwickelt wurden und untrennbar miteinander verbunden sind: *Hub Scores* und *Authority Scores* (als Methode oft auch bezeichnet mit *Hubs & Authorities*).<sup>199</sup> Die Idee hinter diesen Maßen ist (in der Sprache des *Web*), dass ein „guter *Hub*“ von vielen „guten *Authorities*“ verlinkt wird, während eine „gute *Authority*“ auf viele „gute *Hubs*“ verlinkt.<sup>200</sup> Dementsprechend wird verlangt, dass der *Hub Score* eines Punktes proportional zur Summe der *Authority Scores* der Punkte ist, die auf ihn zeigen, und dass der *Authority Score* eines Punktes proportional zur Summe der *Hub Scores* der Punkte ist, auf die er zeigt.<sup>201</sup> Inwieweit sich die Originalmotivation von *Hubs & Authorities* auf juristische Zitiernetzwerke übertragen lässt, hängt davon ab, welches konkrete Netzwerk man betrachtet und welche Annahmen über Zitierprozesse man zugrunde legt. Interessant ist an den *Hub Scores* und den *Authority Scores* jedenfalls, dass sie mathematisch eng mit bibliographischer Kopplung und Co-Zitation zusammenhängen, die sich (als ungerichtete Beziehungen zweiten Grades) gerade von den Zitierbeziehungen ableiten: Der *Hub Score* entspricht in etwa der Eigenvektorzentralität für das Netzwerk der bibliographischen Kopplung und der *Authority Score* entspricht in etwa der Eigenvektorzentralität für das Co-Zitationsnetzwerk.<sup>202</sup>

Abbildung 3.21 (S. 139) visualisiert die Werte der hier vorgestellten Rückkopplungszentralitäten für die in Abbildung 3.20 (S. 134) bereits betrachteten Subnetzwerke des flatländischen Zitiernetzwerks. Hier zeigt sich unter anderem, dass diese Maße sehr viel selektiver sind als die Gradzentralitäten

<sup>198</sup> Beispiele: *Winkels, de Ruyter und Kroese*, Determining Authority of Dutch Case Law, in: *Atkinson* (Hrsg.), *Frontiers in Artificial Intelligence and Applications*, 2011, S. 103–112; *Derlén und Lindholm*, *European Law Journal* 20 (2014), S. 667–687; *Derlén und Lindholm*, *Journal of International Economic Law* 20 (2017), S. 257–277.

<sup>199</sup> Erstmals vorgeschlagen in *Kleinberg*, *Journal of the ACM* 46 (1999), S. 604–632; zusammenfassend *Easley und Kleinberg*, *Networks, Crowds, and Markets*, 2010, S. 356–358. In der Literatur zur Netzwerkforschung im Recht werden *Hubs & Authorities* eingesetzt etwa von *Leicht u.a.*, *The European Physical Journal B* 59 (2007), S. 75–83; *Fowler und Jeon*, *Social Networks* 30 (2008), S. 16–30; *Bommarito, Katz und Isaacs-See*, *Virginia Tax Review* 30 (2011), S. 523–557; *Lupu und Voeten*, *British Journal of Political Science* 42 (2012), S. 413–439; *Derlén und Lindholm*, *German Law Journal* 18 (2017), S. 647–686.

<sup>200</sup> *Kleinberg*, *Journal of the ACM* 46 (1999), S. 604–632, S. 611.

<sup>201</sup> *Easley und Kleinberg*, *Networks, Crowds, and Markets*, 2010, S. 358; *Newman*, *Networks*, 2010, S. 179–180.

<sup>202</sup> Zu Co-Zitation und bibliographischer Kopplung bereits S. 37–55. Das oben angesprochene Problem der Eigenvektorzentralität bei gerichteten Graphen stellt sich hier nicht, da Co-Zitation und bibliographische Kopplung symmetrische Beziehungen sind, die als Linien dargestellt werden, sodass der resultierende Graph ungerichtet ist. Der Zusammenhang lässt sich mit Matrizenrechnung herleiten, dazu *Newman*, *Networks*, 2010, S. 180–181, allerdings auf Basis der umgekehrten (und unüblichen) Definition der Adjazenzmatrix (a.a.O. S. 114).

und dass der *PageRank* und der *Authority Score* andere Strukturinformationen hervorheben als der *Hub Score*.

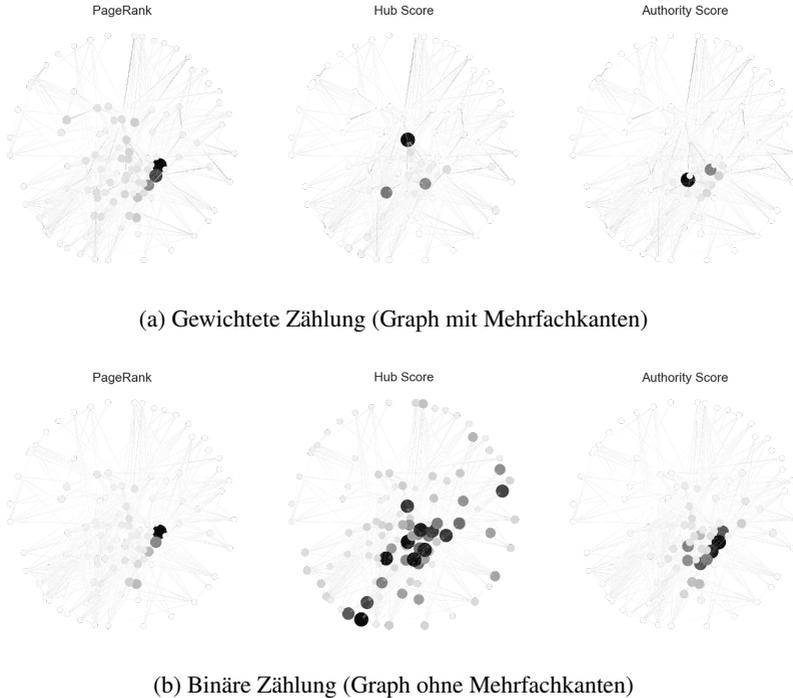


Abbildung 3.21: Hub Score, Authority Score und PageRank für höchstrichterliche Entscheidungen mit Inhaltsziffern 0, 1 oder 2<sup>203</sup>

### c) Positionszentralitäten

Grad- und Rückkopplungszentralitäten haben gemeinsam, dass sie auf Basis *lokaler* Informationen ermittelt werden: nur Informationen zu Linien oder Punkten in der unmittelbaren Nachbarschaft eines Punktes fließen direkt in seinen Zentralitätswert ein. Rückkopplungszentralitäten führen zwar dazu, dass sich lokale Informationen über mehrere Berechnungsschritte im Graphen verteilen, doch sie berücksichtigen die Informationen über die Struktur des gesamten Graphen und die Position eines Punktes relativ zu allen anderen Punkten nur indirekt. Anders arbeiten Zentralitätsmaße, die hier als *Positionszentralitäten* bezeichnet werden und die nicht nur lokale, sondern auch *globale* Informationen unmittelbar

<sup>203</sup> Zur Konstruktion der abgebildeten Graphen bereits in Fn. 183 (S. 134).

in die Ermittlung der Zentralität eines Punktes einbeziehen, wenngleich in sehr unterschiedlicher Weise.<sup>204</sup>

Betrachtet man das Verflechtungsnetzwerk der Unternehmen in Flatland aus regulatorischer Perspektive, so kann man beispielsweise fragen, welche Unternehmen bei der Ausbreitung bestimmter (erwünschter oder unerwünschter) Führungspraktiken in der Unternehmenslandschaft insgesamt eine zentrale Rolle spielen. Nun ist anzunehmen, dass die Führungsfiguren auf die Führungspraktiken der Unternehmen einen gewissen Einfluss haben und dass sie die Erfahrungen, die sie bei einem Unternehmen machen, in ihre Führungstätigkeit bei anderen Unternehmen einfließen lassen.<sup>205</sup> Geht man zudem davon aus, dass die Regulierer nur die Umsetzung oder Nicht-Umsetzung von Führungspraktiken durch die Unternehmen direkt beobachten und nur die Unternehmen, nicht aber die Führungsfiguren unmittelbar regulieren können, so scheint es nicht fernliegend, zur Untersuchung der Wichtigkeit von Unternehmen für die Ausbreitung von Führungspraktiken das Unternehmensnetzwerk heranzuziehen, in dem zwei Unternehmen miteinander verbunden sind, wenn sie eine Führungsfigur teilen. Dabei kann ein Unternehmen zwei grundlegend verschiedene Rollen spielen: Es kann Ursprung einer Führungspraxis sein oder aber eine Führungspraxis von einem anderen Unternehmen übernehmen und dann an weitere Unternehmen übertragen.

Der erste Fall ist aus Regulierungssicht besonders interessant für Interventionen, bei denen es darum geht, dafür zu sorgen, dass sich als positiv bewertete Praktiken in der Unternehmenslandschaft durchsetzen (*positive Interventionen*). Hier kann man fragen: Wie schnell kann sich eine Führungspraxis, ausgehend von einem betrachteten Unternehmen, in andere Unternehmen ausbreiten? Eine mögliche Antwort bietet die Nähezentralität (*closeness centrality*), die misst, wie nah ein Punkt sich an allen anderen Punkten im Graphen befindet. Demgegenüber interessiert der zweite Fall vor allem für Interventionen, durch welche die Weiterverbreitung einer Praxis gerade verhindert werden soll (*negative Interventionen*). Hier stellt sich vor allem die Frage: Wie sehr hängt die Ausbreitung einer Führungspraxis davon ab, dass das betrachtete Unternehmen eine Vermittlerrolle übernimmt? In diesem Kontext hilft unter anderem die Zwischenzentralität (*betweenness centrality*), die misst, inwieweit sich ein Punkt zwischen allen anderen Punkten im Graphen befindet.

Sowohl Nähe- als auch Zwischenzentralität beziehen globale Weginformationen zu allen Punkten des untersuchten Graphen unmittelbar in die Ermittlung der Zentralitätswerte einzelner Punkte ein.<sup>206</sup> Sie unterscheiden sich aber in der

<sup>204</sup> Eine Unterscheidung zwischen lokalen und globalen Zentralitätsmaßen findet sich auch bei Koschützki u.a., Centrality Indices, in: Brandes und Erlebach (Hrsg.), Network Analysis: Methodological Foundations, 2005, S. 16 (20).

<sup>205</sup> Diese Annahme liegt unter anderem der im Literaturüberblick eingeführten Forschung zu Personalverflechtungen zugrunde, durch welche die fiktiven Unternehmensdaten inspiriert sind, dazu S. 22–28.

<sup>206</sup> Die folgenden Ausführungen sind auf ungerichtete Graphen zugeschnitten, zu denen auch der Unternehmens- und der Personalgraph aus Flatland gehören. Die diskutierten Maße lassen

Position, die ein Punkt bei der Erfassung dieser Informationen einnimmt: Die Nähezentralität ist ein *radiales* Maß, bei dem der betrachtete Punkt den Ausgangs- bzw. Endpunkt aller erfassten Wege bildet, während die Zwischenzentralität ein *mediales* Maß darstellt, bei dem der betrachtete Punkt Durchgangspunkt aller erfassten Wege ist.<sup>207</sup>

In ihrer Standardformulierung ist die Nähezentralität eines Punktes  $i$  definiert als das Inverse des arithmetischen Mittels der Distanzen von  $i$  zu allen anderen Punkten im Graphen, d.h.  $n - 1$  multipliziert mit dem Inversen der Summe (Zusammenfassungsmodus) der Längen  $d_{ij}$  (Weegeigenschaft) der kürzesten Wege (Wegtyp), die von  $i$  zu allen Punkten  $j$  ausgehen (Punktposition):

$(\frac{1}{n-1} \cdot \sum_j d_{ij})^{-1} = \frac{n-1}{\sum_j d_{ij}}$ .<sup>208</sup> Demgegenüber ist die Zwischenzentralität eines Punktes  $i$  definiert als der Anteil der Anzahl aller kürzesten Wege zwischen zwei vom betrachteten Punkt verschiedenen Punkten, die gerade durch  $i$  verlaufen ( $\frac{g_{jik}}{g_{jk}}$ ), summiert über alle Paare von Punkten:

$$\sum_i \sum_j \frac{g_{jik}}{g_{jk}}, j, k \neq i.$$
<sup>209</sup>

Die Definitionen von Nähe- und Zwischenzentralität sind sperrig und werden in der Literatur weder sprachlich noch mathematisch einheitlich formuliert; welche Definitionen in der Praxis zum Einsatz kommen, hängt von der verwendeten Software ab. Hier interessieren daher vor allem die grundsätzlichen Einschränkungen dieser Maße: Beide haben mathematische Probleme mit unverbundenen Graphen (also Graphen, bei denen nicht alle Punkte von allen anderen Punkten aus über Linien erreichbar sind).<sup>210</sup> Noch schwerer wiegt allerdings, dass sowohl Nähe-

---

sich aber auch für gerichtete Graphen definieren, dazu *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 200–201; *Estrada und Knight*, *A First Course in Network Theory*, 2015, S. 147, 152.

<sup>207</sup> *Borgatti und Everett*, *Social Networks* 28 (2006), S. 466 (476).

<sup>208</sup> Grundlegend *Beauchamp*, *Behavioral Science* 10 (1965), S. 161 (161) in Reaktion auf *Bavelas*, *The Journal of the Acoustical Society of America* 22 (1950), S. 725 (726); aufgegriffen unter anderem von *Freeman*, *Social Networks* 1 (1979), S. 215 (226); *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 185; *Knoke und Yang*, *Social Network Analysis*, 2008, S. 66; *Estrada und Knight*, *A First Course in Network Theory*, 2015, S. 147. Dass die Summe im Nenner des Bruchs steht, rührt daher, dass man sich ein Maß wünscht, bei dem diejenigen Punkte die höchsten Werte erhalten, die geringe Entfernungen zu allen anderen Punkten aufweisen. Stünde die Summe im Zähler, würde man stattdessen die Distanzen direkt messen und dadurch hohe Werte gerade für Punkte erhalten, die besonders weit von allen anderen entfernt sind.

<sup>209</sup> Grundlegend *Freeman*, *Sociometry* 40 (1977), S. 35 (37); aufgegriffen unter anderem von *Koschützki u.a.*, *Centrality Indices*, in: *Brandes und Erlebach* (Hrsg.), *Network Analysis: Methodological Foundations*, 2005, S. 16 (29–30); *Borgatti und Everett*, *Social Networks* 28 (2006), S. 466 (474); *Estrada und Knight*, *A First Course in Network Theory*, 2015, S. 152 (jeweils mit leichten Notationsunterschieden).

<sup>210</sup> Die Probleme mit unverbundenen Graphen rühren daher, dass es keine kürzesten Wege zwischen zwei unverbundenen Punkten gibt. Die Schwierigkeiten lassen sich praktisch durch verschiedene Anpassungen der Definitionen oder durch bestimmte Berechnungstechniken umgehen, was aber die Semantik der umgesetzten Wichtigkeitsdefinition beeinträchtigen kann. Als Lösungsansatz kann man beispielsweise in Betracht ziehen, die Zentralitätswerte für jeden Teil

als auch Zwischenzentralität auf Basis *kürzester Wege* errechnet werden, womit für die Interpretation der Zentralitätswerte die Annahme verbunden ist, dass auf dem als Graph dargestellten Netzwerk Prozesse ablaufen und der Ablauf auch noch über die kürzesten Wege erfolgt.<sup>211</sup> Dem entspricht für das eingangs betrachtete Beispiel des Unternehmensnetzwerks die Annahme, dass einzelne Unternehmenspraktiken zielgerichtet und über die kleinstmögliche Anzahl vermittelnder Führungsfiguren von Unternehmen zu Unternehmen übertragen werden. Diese Annahme ist in etwa so (wenig) realistisch wie die Annahme, dass ein Virus die Individuen in einer Population auf kürzesten Wegen infiziert. Auf Basis der eingeführten Differenzierungen bei der Wegerfassung (S. 123–129) sind aber zahlreiche Modifikationen denkbar, die insbesondere mit Diffusionsprozessen besser vereinbar sind als die oben wiedergegebenen klassischen Definitionen;<sup>212</sup> in der Literatur zur Netzwerkforschung im Recht dominieren dennoch die klassischen Formulierungen.<sup>213</sup> Ihr größtes Potenzial entfalten die Positionszentralitäten denn auch jenseits des Lokalen – bei der Beantwortung der Gruppierungsfrage (hierzu sogleich).

### III. Mesoebene

#### 1. Problemstellung: Was gehört zusammen?

Bei der Untersuchung eines Netzwerks auf der Mesoebene geht es darum, seine regionalen Strukturen zu ermitteln und zu beschreiben. Hierzu muss festgestellt werden, welche Elemente des betrachteten Netzwerks zusammengehören, im Mittelpunkt steht also die Zuordnung von Netzwerkelementen zu Gruppen (Gruppierungsfrage).<sup>214</sup> Ebenso wie bei der Beantwortung der Wichtigkeitsfrage auf der Mikroebene sind dabei ein Definitionsproblem und ein Messproblem zu lösen. Einerseits muss präzisiert werden, was mit Zusammengehörigkeit gemeint ist. An-

---

des untersuchten Graphen, auf dem sie wohldefiniert sind, separat zu berechnen (von *Koschützki u.a.*, Centrality Indices, in: *Brandes und Erlebach* (Hrsg.), *Network Analysis: Methodological Foundations*, 2005, S. 16 (56–57) als „very naïve approach“ bezeichnet). Dann erhält man allerdings für jeden dieser Teile eine separate Antwort auf die Wichtigkeitsfrage und es ist unklar, wie die erhaltenen Zentralitätswerte zueinander ins Verhältnis zu setzen sind.

<sup>211</sup> *Borgatti*, *Social Networks* 27 (2005), S. 55 (59–63).

<sup>212</sup> Dazu beispielsweise *Koschützki u.a.*, Centrality Indices, in: *Brandes und Erlebach* (Hrsg.), *Network Analysis: Methodological Foundations*, 2005, S. 16 (43–46).

<sup>213</sup> Beispiele: *Fowler*, *Political Analysis* 14 (2006), S. 454–465; *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457–509; *Boulet, Mazzeo und Bourcier*, *Artificial Intelligence and Law* 19 (2011), S. 333–355; *Winkels, Boer und Plantevin*, *Creating Context Networks in Dutch Legislation*, in: *Ashley* (Hrsg.), *Legal Knowledge and Information Systems*, 2014, S. 155–164; *Derlén und Lindholm*, *European Law Journal* 20 (2014), S. 667–687. Über die Schwierigkeiten, die sich bei der Anwendung der Maße auf gerichtete Graphen ergeben, wird dabei in der Regel hinweg gegangen.

<sup>214</sup> Dazu bereits S. 73.

dererseits ist zu bestimmen, wie die so definierte Zusammengehörigkeit gemessen werden soll.<sup>215</sup>

Definitionen von Zusammengehörigkeit können auf der Mesebene im Wesentlichen drei Stoßrichtungen haben: Sie können erstens darauf angelegt sein, Subnetzwerke zu erzeugen, deren Zusammensetzung dann im Einzelnen näher untersucht wird – das Ziel ist dann *Fokussierung*.<sup>216</sup> Zweitens können sie aber auch bezwecken, aus dem untersuchten Netzwerk ein reduziertes Netzwerk zu erstellen, bei dem alles Zusammengehörige durch *ein* Objekt repräsentiert wird und die Beziehungen entsprechend zusammengefasst werden – hier ist das Ziel *Komprimierung*.<sup>217</sup> Schließlich können sie drittens dazu dienen, dem Netzwerk (ohne seine Veränderung) neue Informationen hinzuzufügen, die einen Überblick über seine Organisation ermöglichen – das Ziel ist hier *Strukturierung*. Während Komprimierung prinzipiell eine *Zerlegung (partition)* erfordert, die jedes Objekt *genau einer* Gruppe zuordnet,<sup>218</sup> damit eindeutig feststellbar ist, *welcher* Gruppe es angehört,<sup>219</sup> wird zur Fokussierung lediglich eine Gruppierung benötigt, anhand derer eindeutig feststellbar ist, *ob* ein Objekt die definitionsgemäßen Anforderungen an die Gruppenzugehörigkeit erfüllt, sodass es auch mehreren oder gar keinen Gruppen angehören kann; zur Strukturierung ist sogar Eindeutigkeit nicht zwingend erforderlich.

Außerdem können sich Definitionen von Zusammengehörigkeit sowohl an den Merkmalen der Objekte oder Beziehungen als auch an der Beziehungsstruktur eines Netzwerks orientieren. Im ersten Fall lassen sie sich als *semantische Definitionen*, im zweiten Fall als *strukturelle Definitionen* bezeichnen. Beide Definitionstypen sind in dieser Arbeit bereits zum Einsatz gekommen: So wurde zur Veranschaulichung der Zentralitätsmaße im vorigen Abschnitt aus dem Zitiernetz-

<sup>215</sup> Wie bei der Wichtigkeitsfrage ist auch hier problematisch, inwieweit die Definition der Messung vorausgehen muss. Symptomatisch dafür ist die Formulierung am Ende des Überblicksbeitrags von *Fortunato*, *Physics Reports* 486 (2010), S. 75 (163): „It is quite possible that the scientific community will converge sooner or later to a definition *a posteriori* of community.“ (Hervorhebung im Original). Zu jedem Maß muss aber zumindest ein Berechnungsverfahren (*Algorithmus*) angegeben werden, das (auch dann, wenn es auf abstrakten mathematischen Kriterien beruht) kaum ohne ein gewisses Vorverständnis formuliert werden kann. Bei der Anwendung des Maßes wird dann jedenfalls dieses Vorverständnis (und damit eine „implizite Definition“ des gemessenen Konzepts) übernommen.

<sup>216</sup> Beispiel: Abb. 3.13 (S. 120).

<sup>217</sup> Beispiel: Abb. 3.14 (S. 121).

<sup>218</sup> Formal ist eine Zerlegung (auch im Deutschen teilweise *Partition* genannt) einer Menge  $M$  definiert als eine Menge von nichtleeren Teilmengen von  $M$  (d.h. jede Teilmenge enthält mindestens ein Element), die paarweise disjunkt sind (jedes Element von  $M$  gehört maximal einer Teilmenge an, sodass sich keine zwei Teilmengen überschneiden) und deren Vereinigung  $M$  ist (d.h. jedes Element von  $M$  gehört mindestens einer Teilmenge an); *Jänich*, *Lineare Algebra*, 2013, S. 228; *Beutelspacher*, *Lineare Algebra*, 2014, S. 6.

<sup>219</sup> Lediglich *prinzipiell*, weil es durchaus denkbar ist, einzelne Objekte zunächst mehreren Gruppen zuzuordnen und dann Regeln zu formulieren, wie Mehrfachzuordnungen bei der Konstruktion des reduzierten Graphen zu behandeln sind.

werk ein Subnetzwerk gewonnen, das nur höchstrichterliche Entscheidungen der Inhaltsziffern 0, 1 oder 2 enthält.<sup>220</sup> Das entspricht einer semantischen Definition von Zusammengehörigkeit mit dem Ziel, das Netzwerk durch Filterung (bzw. umgekehrt: Selektion) so zu verkleinern, dass sich sein zugehöriger Graph ohne Wollmauseffekt visualisieren lässt.<sup>221</sup> Im Rahmen der Diskussion zur Bewegung auf Graphen wurde demgegenüber eine (partiell) strukturelle Definition eingesetzt, um die juristische Recherche als Bewegung auf einem Zitiergraphen darstellen zu können: In das Netzwerk, das Abbildung 3.18 (S. 128) zugrunde liegt, wurden nur solche Entscheidungen des *High Court* aufgenommen, die maximal zwei Klicks von einer bestimmten Ausgangsentscheidung (Aktenzeichen H-HC-0:2-2005-181) entfernt sind (deren Punkte im Graphen also eine Distanz von maximal 2 zum Ausgangspunkt haben). Eine strukturelle Definition von Zusammengehörigkeit ermöglicht hier die Konzentration auf den Kontext der Ausgangsentscheidung.

Semantische Definitionen von Zusammengehörigkeit haben zwar den Vorteil, dass mit ihrer Hilfe viele verschiedene Fokussierungen, Komprimierungen oder Strukturierungen erlangt werden können, indem man in unterschiedlicher Weise auf die Gleichheit oder (relative) Ähnlichkeit der Ausprägungen diverser Merkmale abstellt. Um diese Flexibilität sinnvoll zu nutzen, muss man allerdings bereits einige Informationen über das betrachtete Netzwerk besitzen – wo es keine inhaltliche Klassifikation von Gerichtsentscheidungen gibt, kann diese nicht zur Filterung eines Zitiernetzwerks eingesetzt werden – und man läuft in besonderer Weise Gefahr, Opfer der eigenen Erwartungen zu werden.<sup>222</sup>

Während semantische Definitionen *konkret* sind und vor allem praktische Schwierigkeiten mit sich bringen, sind strukturelle Definitionen *abstrakt*, da sich Beziehungsstrukturen in einem Netzwerk direkt (und damit insbesondere ohne Kenntnis der Netzwerksemantik) auf Verbindungsstrukturen in einem Graphen abbilden lassen. Mit strukturellen Definitionen gehen allerdings spezifische methodische Schwierigkeiten einher, denn es ist unklar, wie die definitionsbestimmenden Beziehungsstrukturen aussehen können und wie sie im Netzwerk gefunden werden sollen. Die folgenden Ausführungen konzentrieren sich daher auf strukturelle Definitionen von Zusammengehörigkeit, während ausgewählte semantische Definitionen im vierten Kapitel anhand realer Daten zur Rechtsprechung des BVerfG behandelt werden.<sup>223</sup>

<sup>220</sup> Siehe z.B. Abb. 3.20 (S. 134).

<sup>221</sup> Filterung und Selektion unterscheiden sich (nur) darin, wie das Subnetzwerk definiert wird. Filterung beginnt mit dem Gesamtnetzwerk und schließt Netzwerkelemente aus, weil sie bestimmte Kriterien nicht erfüllen, Selektion beginnt mit einem leeren Netzwerk und schließt Netzwerkelemente aus dem Gesamtnetzwerk ein, weil sie bestimmte Kriterien erfüllen. Es handelt sich also um gegenläufige Ansätze zur Lösung desselben Auswahlproblems.

<sup>222</sup> Das Problem potenziert sich noch, wenn bereits in die Netzwerkdefinition zahlreiche individuelle Wertungsentscheidungen einfließen, wie es bei der Arbeit mit Beziehungen dritten Grades der Fall ist; dazu bereits im zweiten Kapitel (S. 38–39).

<sup>223</sup> Dazu S. 301–312.

## 2. Lösungsansätze

Wie lässt sich Zusammengehörigkeit strukturell definieren und messen? Die allgemeine Literatur zur Netzwerkforschung bietet eine Vielzahl von Antworten, die mit zwei unterschiedlichen Vorstellungen von Zusammengehörigkeit operieren.<sup>224</sup> Zum einen kann man Zusammengehörigkeit als (relativ) *engen Zusammenhalt*, als Kohäsion (*cohesion*) begreifen: Objekte mit hoher Beziehungsdichte untereinander gehören eher zusammen als Mengen von Objekten mit niedriger Beziehungsdichte.<sup>225</sup> Mit der Beziehungsdichte einer Gruppe kann Verschiedenes gemeint sein; insbesondere kann man zwischen der *internen* Beziehungsdichte einer Gruppe (Verhältnis der realisierten zu den möglichen Beziehungen innerhalb einer Gruppe) und ihrer *externen* Beziehungsdichte (Verhältnis der realisierten zu den möglichen Beziehungen von Mitgliedern der Gruppe zu Objekten, die der Gruppe nicht angehören) unterscheiden.<sup>226</sup> Ansätze, die sich zur Gruppierung von Objekten an der Beziehungsdichte orientieren, werden hier als *dichtebasierte Ansätze* bezeichnet (Abb. 3.22 (a), S. 146).<sup>227</sup>

Zum anderen kann man Zusammengehörigkeit aber auch als (relativ) *ähnliche Einbettung* verstehen: Mengen von Objekten, deren Beziehungen zu anderen Objekten ähnlich strukturiert sind, gehören eher zusammen als Mengen von Objekten, deren Beziehungen zu anderen Objekten sehr unterschiedlich strukturiert sind.<sup>228</sup> Ähnliche Einbettung kann ebenfalls Unterschiedliches bedeuten; stets geht es dabei aber um Symmetrien in den beobachteten Beziehungen zwischen den zu einer Gruppe zusammengefassten Objekten, sodass Ansätze, die sich zur Gruppierung von Objekten an ihrer Einbettung im Netzwerk orientieren, hier als *symmetriebasierte Ansätze* bezeichnet werden (Abb. 3.22 (b), S. 146).

Während es bei dichtebasierten Ansätzen prinzipiell unerwünscht ist, dass Objekte in eine Gruppe fallen, die nicht mindestens indirekt miteinander verbunden sind, ist dies bei symmetriebasierten Ansätzen durchaus möglich (s. Abb. 3.22, S. 146). Welches Konzept zur Gruppenbildung bei der Analyse eines Netzwerks auf der Mesoebene zugrunde gelegt werden kann und wie diese Vorstellung formal umzusetzen ist, hängt vom Erkenntnisinteresse ab: Eine Orientierung an enger

<sup>224</sup> So auch Wasserman und Faust, *Social Network Analysis*, 1994, S. 348.

<sup>225</sup> Beispiel: Sucht man nach Gerichtsentscheidungen, die ähnliche Themen betreffen, so gehört eine Menge von Gerichtsentscheidungen, von denen sich viele untereinander (im Rahmen des zeitlich möglichen) zitieren, eher zusammen als eine Menge von Gerichtsentscheidungen, die sich untereinander kaum zitieren – jedenfalls dann, wenn man davon ausgeht, dass Zitate zumindest auch inhaltlich bedingt sind.

<sup>226</sup> Fortunato, *Physics Reports* 486 (2010), S. 75 (84).

<sup>227</sup> Von *density-based approaches* spricht auch Gaertler, Clustering, in: Brandes und Erlebach (Hrsg.), *Network Analysis*, 2005, S. 178 (178).

<sup>228</sup> Beispiel: Sucht man nach Akteuren, die ähnliche Funktionen in einem hierarchischen Rechtssystem übernehmen, so gehören mehrere Gerichte einer Ebene, die demselben Gericht untergeordnet sind, eher in eine Kategorie als Gerichte einer Ebene, die unterschiedlichen Gerichten untergeordnet sind, oder Gerichte, die in einem Über-Unterordnungsverhältnis zueinander stehen.

Verbindung zielt auf die Entdeckung von *Gemeinschaften (communities)*, auch genannt *Module (modules)* oder *Cluster (clusters)*, jene an ähnlicher Einbettung auf die Erkennung von *Positionen (positions)* und *Rollen (roles)*.<sup>229</sup>

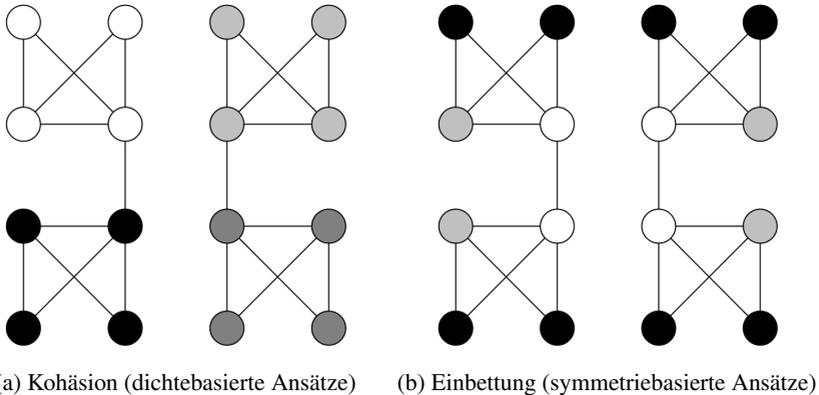


Abbildung 3.22: Konzepte von Zusammengehörigkeit<sup>230</sup>

<sup>229</sup> Die in der Literatur anzutreffende Terminologie ist sehr heterogen, was auch damit zusammenhängt, dass sie sich in vielen verschiedenen Wissenschaftszweigen mit unterschiedlichen Begriffstraditionen (zum Teil unkoordiniert) entwickelt hat und teils mathematische Operationen auf Graphen, teils Interpretationen von Netzwerken betont. Das erschwert weniger das Verständnis einzelner Texte, in denen die Begriffsverwendung in der Regel aus dem Kontext deutlich wird, als die Suche nach einschlägiger Literatur. Im Folgenden wird von Gemeinschaften einerseits sowie von Positionen und Rollen andererseits gesprochen. Sie alle sind Ergebnis von Prozessen der Gruppierung, sodass „Gruppen“ als Oberbegriff verwendet wird. Das entspricht einem Mittelweg zwischen den Terminologien der Natur- und Sozialwissenschaften, von denen die Erstere bisweilen stark mathematisch, die Letztere stark soziologisch aufgeladen ist.

<sup>230</sup> Unterschiedliche Farben stehen für unterschiedliche Gruppen. Der Graph in Abbildung 3.22 (a) (S. 146) ist so gefärbt, dass ein möglichst gutes Verhältnis der internen zur externen Beziehungsdichte der einzelnen Gruppen erreicht wird. Dabei geht es nicht ausschließlich darum, die Anzahl der Linien zu minimieren, die zwischen einzelnen Gruppen verlaufen, denn dann dürften hier lediglich zwei Gruppen gebildet werden (Weiß-Schwarz und Hellgrau-Dunkelgrau); zum Vorgehen im Einzelnen unten S. 147–159. Der Graph in Abbildung 3.22 (b) (S. 146) ist so gefärbt, dass die gleichfarbigen Punkte in dem Sinne gegeneinander austauschbar sind, dass jeder von ihnen mit derselben Anzahl von Punkten anderer Farbe verbunden ist. Dadurch fallen auch Punkte in eine Gruppe, die nicht direkt miteinander verbunden sind (z.B. die weißen Punkte in der „linken“ und „rechten“ Hälfte des Graphen), was bei symmetriebasierten Ansätzen keine Seltenheit ist; zum Vorgehen im Einzelnen unten S. 159–165.

a) *Gemeinschaften*

Der gemeinsame Ausgangspunkt aller Methoden zur Gemeinschaftserkennung in Netzwerken (*community detection*) ist die Erwartung, dass in einem Netzwerk mit Gemeinschaftsstruktur (*community structure*)<sup>231</sup> relativ viele Beziehungen *innerhalb* einzelner Gemeinschaften und relativ wenige Beziehungen *zwischen* verschiedenen Gemeinschaften verlaufen sollten. Im Fokus steht also die (relative) Beziehungsdichte. Die Kriterien, nach denen Gemeinschaften in einem Netzwerk identifiziert werden, nehmen unterschiedliche Formen an. Sie können als mathematische Formeln ausgedrückt werden (*direkte Definitionen*) oder aber algorithmisch formuliert sein (*indirekte Definitionen*).<sup>232</sup> Mathematische Formeln kommen vor allem dann in Betracht, wenn unmittelbar auf graphentheoretischen Definitionen aufgebaut wird.<sup>233</sup> Die Methoden, mit denen direkte und indirekte Definitionen umgesetzt werden, lassen sich ferner danach unterscheiden, welche Informationen über den Graphen sie zur Gruppenbildung benötigen: *Lokale Methoden* kommen mit Informationen zu einzelnen Punkten und deren jeweiligen Nachbarschaften aus, ohne stets Berechnungen auf dem gesamten Graphen durchzuführen; *globale Methoden* hingegen benötigen Informationen zum gesamten Graphen in ihren Kalkulationen.<sup>234</sup>

<sup>231</sup> Nicht jedes Netzwerk muss Gemeinschaftsstruktur aufweisen. Der Ausdruck *community structure* wurde in der Physik von *Girvan und Newman*, PNAS 99 (2002), S. 7821–7826, populär gemacht; in der Soziologie wird das so bezeichnete Phänomen (seit deutlich längerer Zeit) unter dem Stichwort *cohesive subgroups* diskutiert, so etwa bei *Wasserman und Faust*, Social Network Analysis, 1994, S. 249–290.

<sup>232</sup> Von indirekten Definitionen spricht in diesem Kontext auch *Fortunato*, Physics Reports 486 (2010), S. 75 (85). Sie haben an Bedeutung gewonnen, seitdem der Beitrag von *Girvan und Newman*, PNAS 99 (2002), S. 7821–7826, das Forschungsfeld der Gemeinschaftserkennung für die Physik geöffnet hat. Die Soziale Netzwerkanalyse scheint traditionell direkte Definitionen zu betonen. Das kann historisch bedingt sein (die Soziale Netzwerkanalyse ist älter als der Computer), aber auch damit zusammenhängen, dass indirekte Definitionen hinsichtlich der Semantik des untersuchten Netzwerks das Risiko eines Theoriedefizits bergen (es ist oft unklar, welches inhaltliche Konzept von Zusammengehörigkeit einem Algorithmus zugrunde liegt). Das stört die Physik mit ihrem Interesse am Verallgemeinerbaren naturgemäß weniger als die Soziologie (dazu bereits S. 62–71); das Theoriedefizit indirekter Methoden bemerkt allerdings auch *Fortunato*, Physics Reports 486 (2010), S. 75 (161).

<sup>233</sup> Ein Beispiel ist die Fokussierung auf den Kontext einer ausgewählten Entscheidung im Zitiernetzwerk: „Kontext einer Entscheidung“ lässt sich übersetzen in „das Egonetzwerk der Entscheidung mit Radius 1“ (zu Egonetzwerken bereits in Kapitel 2 bei Fn. 138 [S. 60]) – was graphentheoretisch nichts anderes ist als der Subgraph, den man erhält, wenn man den Zitiergraphen auf den (der Entscheidung entsprechenden) Punkt und dessen Nachbarschaft einschränkt.

<sup>234</sup> Die Unterscheidung zwischen globalen und lokalen Methoden findet sich auch bei *Porter, Onnela und Mucha*, Notices of the American Mathematical Society 56 (2009), S. 1082 (1088); globale und lokale Definitionen unterscheidet *Fortunato*, Physics Reports 486 (2010), S. 75 (84–86).

Die höchste Beziehungsdichte wird in Netzwerken erreicht, deren Objekte sämtlich mit allen anderen Objekten in Beziehung stehen, in deren Graphen also alle Punkte mit allen anderen Punkten direkt verbunden sind. Solche Graphen heißen *vollständig* (*complete*).<sup>235</sup> Diese Eigenschaft lässt sich auch auf Subgraphen beziehen. Vollständige Subgraphen heißen *maximal*, wenn man keinen Punkt aus dem umgebenden Graphen hinzunehmen kann, ohne dass sie ihre Vollständigkeits-eigenschaft verlieren.<sup>236</sup> Dies führt zum striktesten dichte-basierten Ansatz: dem Konzept der *Clique* (*clique*), auf deren direkter Definition viele lokale Methoden aufbauen. Cliques sind maximale vollständige Subgraphen und grundsätzlich (nur) für schlichte Graphen (also ungerichtete, ungewichtete Graphen ohne Schleifen oder Mehrfachkanten) definiert.<sup>237</sup> Ein Punkt kann ohne weiteres mehreren Cliques angehören;<sup>238</sup> eine Clique mit einer Anzahl von  $k$  Punkten wird  $k$ -Clique genannt.<sup>239</sup> Eine naheliegende Frage ist dann, welche Punkte zu den *größten* (*maximum*) Cliques eines Graphen gehören – also zu jenen Cliques, welche die größtmögliche Anzahl von Punkten enthalten.<sup>240</sup>

<sup>235</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 102. Die Anzahl der Linien eines solchen vollständigen Graphen ergibt sich zwingend aus der Anzahl seiner Punkte: Ein vollständiger Graph mit  $n$  Punkten hat  $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n \cdot (n-1)}{2}$  Linien (das ist der Binomialkoeffizient, welcher der Anzahl aller möglichen zweielementigen Kombinationen von  $n$  verschiedenen Elementen entspricht); Harary, *Graph Theory*, 1969, S. 16; Krischke und Röpcke, *Graphen und Netzwerktheorie*, 2015, S. 24.

<sup>236</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 254; Newman, *Networks*, 2010, S. 193–194. Generell bezeichnet man eine Einheit, die eine bestimmte Eigenschaft erfüllen soll, immer dann als maximal, wenn man sie nicht erweitern kann, ohne dass sie diese Eigenschaft verliert; dazu auch Wasserman und Faust, *Social Network Analysis*, 1994, S. 109.

<sup>237</sup> Harary, *Graph Theory*, 1969, S. 20. In manchen Definitionen wird die Menge der Punkte (und nicht der durch sie bestimmte Subgraph) als Clique bezeichnet, so etwa bei Newman, *Networks*, 2010, S. 193. Zum Teil wird die Clique auch ohne die Maximalitätsanforderung definiert (man unterscheidet dann Cliques und maximale Cliques); so beispielsweise Fortunato, *Physics Reports* 486 (2010), S. 75 (84). Der Begriff wurde im soziologischen Kontext eingeführt von Luce und Perry, *Psychometrika* 14 (1949), S. 95 (97–98), die zusätzlich fordern, dass der Subgraph mindestens drei Punkte hat, um einfache Verbindungen zweier Punkte nicht als Cliques auffassen zu müssen; dem folgen Wasserman und Faust, *Social Network Analysis*, 1994, S. 254. Hier wird die Größeneinschränkung weggelassen, um später einen Zusammenhang zu einem weiteren nützlichen Konzept aufzeigen zu können. Verbindungen zweier Punkte sind als Gruppen zwar nicht interessant, müssen aber auch nicht zwingend analysiert werden.

<sup>238</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 254; Newman, *Networks*, 2010, S. 194.

<sup>239</sup> Porter, Onnela und Mucha, *Notices of the American Mathematical Society* 56 (2009), S. 1082 (1088). Die Terminologie ist uneinheitlich und die  $k$ -Cliques im hiesigen Sinne sind nicht zu verwechseln mit den  $n$ -Cliques bei Wasserman und Faust, *Social Network Analysis*, 1994, S. 258 oder den  $k$ -Cliques bei Newman, *Networks*, 2010, S. 195–196. Zur Terminologie auch Fortunato, *Physics Reports* 486 (2010), S. 75 (130).

<sup>240</sup> Die Suche nach Cliques einer bestimmten Mindestgröße (und damit auch nach den größten Cliques) ist allerdings algorithmisch anspruchsvoll – das entsprechende Berechnungsproblem (genannt *Clique*) gehört zur Klasse der NP-vollständigen Probleme, die in der theoretischen

Für die juristische Netzwerkforschung sind unter anderem Cliques in den Graphen von Rechtsprechungsziernetzwerken interessant, die man für die Zwecke der Analyse als ungerichtet behandeln kann.<sup>241</sup> Eine Menge von  $k$  Entscheidungen, die sich alle untereinander zitieren (im Rahmen des zeitlich Möglichen), hängt vielleicht auch inhaltlich miteinander zusammen. So ein Zusammenhang scheint umso wahrscheinlicher, je größer  $k$  ist. Das ist unter anderem für Gerichtssysteme hilfreich, in denen es keine inhaltliche Klassifizierung von Entscheidungen gibt.

Gerade für Zitiergraphen scheint die ursprüngliche Cliquendefinition jedoch zu eng.<sup>242</sup> Ein einziges „fehlendes“ Zitat führt dazu, dass eine Menge eng miteinander verwobener Entscheidungen nicht als zusammengehörig identifiziert wird, da die entsprechenden Punkte im Zitiergraphen keine Clique bilden. In solchen Fällen kann es dem Konzept von Zusammengehörigkeit als Gemeinschaft besser entsprechen, mehrere  $k$ -Cliques zu vereinigen, sofern diese ihrerseits benachbart, das heißt über  $k - 1$  Punkte miteinander verbunden sind.<sup>243</sup> Fasst man schließlich alle  $k$ -Cliques zusammen, die über benachbarte  $k$ -Cliques voneinander erreichbar sind, so erhält man *Gemeinschaften von  $k$ -Cliques* (*k-clique communities*),<sup>244</sup> die inklusiver sind als die herkömmlichen Cliques (Abb. 3.23, S. 150).<sup>245</sup> Setzt man  $k = 2$ , so erhält man als Gemeinschaften von  $k$ -Cliques alle maximalen verbundenen Subgraphen eines Graphen, die als dessen *Komponenten* (*components*)

---

Informatik eine große Rolle spielt; dazu grundlegend *Karp*, *Reducibility Among Combinatorial Problems*, in: *Miller und Thatcher* (Hrsg.), *Complexity of Computer Computations*, 1972, S. 85–103; einführend *Schöning*, *Theoretische Informatik – kurz gefasst*, 2008, S. 148–176.

<sup>241</sup> Das ist unproblematisch, da Rechtsprechungszitate (fast) ausschließlich in die Vergangenheit gerichtet sind, sodass die Zitiergraphen (fast) azyklisch sind. Im Rahmen der Analyse kann man also davon ausgehen, dass es für jede Menge von Punkten genau eine Möglichkeit gibt, vollständig verbunden zu sein, womit man die Pfeilrichtungen ohne Informationsverlust weglassen kann.

<sup>242</sup> Die Enge der Definition bemerken schon ihre Urheber: *Luce und Perry*, *Psychometrika* 14 (1949), S. 95 (98).

<sup>243</sup> *Derényi, Palla und Vicsek*, *Physical Review Letters* 94 (2005), S. 160202-1 (2).

<sup>244</sup> *Palla u.a.*, *Nature* 435 (2005), S. 814 (815); *Porter, Onnela und Mucha*, *Notices of the American Mathematical Society* 56 (2009), S. 1082 (1048). Dieses Konzept ist der Grund dafür, dass hier von  $k$ -Cliques im oben angegebenen Sinn gesprochen wird. Die Methode, mit der man Gemeinschaften von  $k$ -Cliques ermittelt, wird *Cliquenperkolation* (*clique percolation*) genannt. Zu den zahlreichen weiteren Möglichkeiten, die Cliquendefinition realitätstauglicher zu machen, beispielsweise *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 256–262; *Newman*, *Networks*, 2010, S. 194–198. Es gibt auch noch weitere lokale Methoden zur Operationalisierung von Zusammengehörigkeit als Gemeinschaft, die nicht vom Konzept der Clique abgeleitet sind; ein Beispiel hierfür findet sich bei *Bagrow und Bollt*, *Physical Review E* 72 (2005), S. 046108-1–046108-10.

<sup>245</sup> Das zeigt sich etwa, wenn man die ungerichtete Version des Zitiergraphen ohne Mehrfachkanten zur Rechtsprechung in Flatland, eingeschränkt auf die höchstrichterliche Rechtsprechung, untersucht. Wie im Online-Appendix nachvollzogen werden kann, gibt es dort neben den in Abbildung 3.23 (b) (S. 150) dargestellten größten Cliques zwölf 9-Cliques, die sich zu zwei Gemeinschaften von 9-Cliques (ohne Mitgliedschaftsüberschneidungen) zusammenfassen lassen, von denen eine 15 und die andere 12 Mitglieder hat.

bezeichnet werden.<sup>246</sup> Die Größe der Komponenten im Verhältnis zur Größe des Gesamtgraphen ist bei der Charakterisierung eines Netzwerks auf der Makroebene relevant.<sup>247</sup>

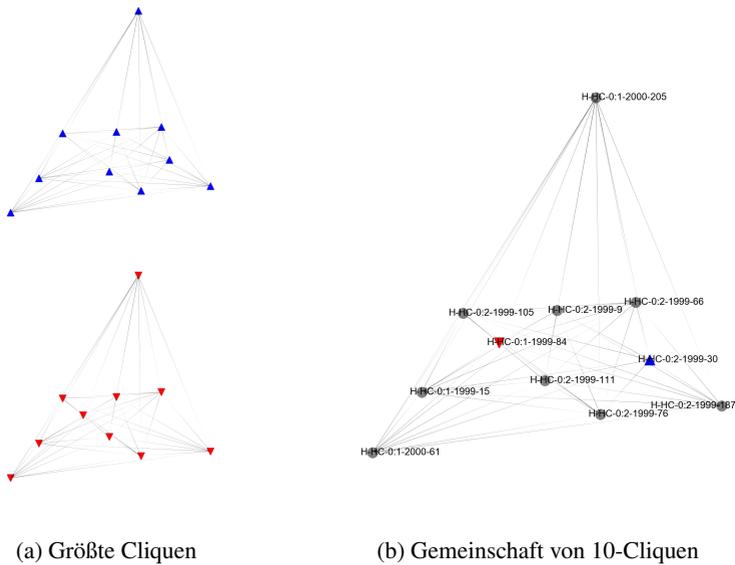


Abbildung 3.23: Direkte Definitionen zur Identifikation von Gemeinschaften im Zitiernetzwerk der höchstrichterlichen Entscheidungen<sup>248</sup>

Aufgrund ihres Aufbaus aus Cliques ist es unter anderem möglich, über Gemeinschaften von  $k$ -Cliques die *überlappende Gemeinschaftsstruktur (overlapping)*

<sup>246</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 109; Barabási, *Network Science*, 2016, S. 61.

<sup>247</sup> Dazu unten S. 167.

<sup>248</sup> Abbildung 3.23 (a) (S. 150) zeigt die zwei größten Cliques im auf höchstrichterliche Entscheidungen eingeschränkten Zitiergraphen der flatländischen Rechtsprechung in seiner Version ohne Mehrfachkanten. Sie enthalten jeweils zehn Punkte. Diese Cliques sind so eng miteinander verbunden, dass sie sich zu einer Gemeinschaft von 10-Cliques mit elf Punkten zusammenfassen lassen. In Abbildung 3.23 (b) (S. 150) sind die Entscheidungen, die beiden in Abbildung 3.23 (a) (S. 150) dargestellten Cliques angehören, als schwarze Punkte dargestellt; Entscheidungen, die lediglich einer von beiden Cliques angehören, haben die Farbe und Form aus Abbildung 3.23 (a) (S. 150). Die Darstellung der Pfeile entspricht der in Fn. 168 (S. 129) beschriebenen Konvention, wenngleich die Gemeinschaftsstruktur selbst auf der ungerichteten Version des Graphen ermittelt wurde. Die Layouts sind wieder kräftebasiert (Fn. 119, S. 113) mit *seed 1234*.

*community structure*) eines Netzwerks zu erkunden.<sup>249</sup> Hierzu sind viele Methoden, die mit weniger restriktiven, indirekten Definitionen arbeiten, nicht geeignet, da sie Zerlegungen von Graphen produzieren.<sup>250</sup> Einige indirekte Methoden erzeugen allerdings mehrere, ineinander verschachtelte Zerlegungen, womit sie sich wiederum anbieten, um die *hierarchische Gemeinschaftsstruktur* (*hierarchical community structure*) von Netzwerken zu untersuchen.<sup>251</sup> Die Vielfalt möglicher Herangehensweisen, die mit indirekten Definitionen arbeiten, ist groß, und der Methodenzeitraum der allgemeinen Netzwerkforschung bringt stets neue Algorithmen hervor.<sup>252</sup> Hier werden daher lediglich einige grundlegende Ideen und Differenzierungen eingeführt, die zur Orientierung in der Literatur und bei der Methodenwahl im Einzelfall nützlich sein können.

Um die Motivation hinter den weniger restriktiven Ansätzen zur Gemeinschaftserkennung zu verstehen, helfen einige Überlegungen zur grundsätzlichen Struktur des zu lösenden Problems. Übergeordnetes Ziel ist, eine Zuordnung von Objekten zu Gemeinschaften zu finden, die der intuitiven Forderung nach möglichst hoher Beziehungsdichte innerhalb und möglichst niedriger Beziehungsdichte zwischen Gemeinschaften in einem Netzwerk gerecht wird (wobei die Anforderungen an die Beziehungsdichte gegenüber den cliquenbasierten Ansätzen abgesenkt werden).<sup>253</sup> So kann man beispielsweise versuchen, die Gerichtsentscheidungen in Flatland so zu kategorisieren, dass relativ viele Zitate innerhalb einzelner Kategorien und relativ wenige Zitate zwischen verschiedenen Kategorien verlaufen. Das ist ein *kombinatorisches Optimierungsproblem* mit mehreren Variablen, die sich gegenseitig beeinflussen: Wie viele Kategorien soll es geben und welche Größe sollen diese haben? Wie vielen verschiedenen Kategorien sollen die einzelnen Entscheidungen angehören können? Wie sollen die einzelnen Entscheidungen den Kategorien zugeordnet werden?

Bei dem Versuch, das Problem in den Griff zu bekommen, stößt man an die Grenzen dessen, was heute mit Computern berechenbar ist. Das hängt auch damit zusammen, dass die Anzahl der möglichen Lösungen ohne jegliche Ein-

---

<sup>249</sup> Palla u.a., *Nature* 435 (2005), S. 814–818; Porter, Onnela und Mucha, *Notices of the American Mathematical Society* 56 (2009), S. 1082 (1088).

<sup>250</sup> Zur Erinnerung: Eine Zerlegung eines Graphen ist die Einteilung seiner Punktmenge in nichtleere Teilmengen so, dass jeder Punkt genau einer Teilmenge angehört.

<sup>251</sup> Palla u.a., *Nature* 435 (2005), S. 814–818.

<sup>252</sup> Mittlerweile werden diese Algorithmen auch getestet und miteinander verglichen; dazu etwa Lancichinetti, Fortunato und Radicchi, *Physical Review E* 78 (2008), S. 046110-1–046110-5; Lancichinetti und Fortunato, *Physical Review E* 80 (2009), S. 056117-1–056117-11.

<sup>253</sup> Diese Forderung ist nicht gleichzusetzen mit der Forderung danach, dass *möglichst viele* Beziehungen innerhalb von Gemeinschaften und *möglichst wenige* Beziehungen zwischen Gemeinschaften verlaufen sollten (dazu bereits Abb. 3.22 [S. 146] und begleitende Fußnote). Das Gemeinschaftskonzept ist vielmehr auch intuitiv „relativ“, etwa stören Beziehungen, die zwischen Gemeinschaften verlaufen, den wahrgenommenen Zusammenhalt umso weniger, je mehr Mitglieder diese Gemeinschaften haben, wie auch Newman und Girvan, *Physical Review E* 69 (2004), S. 026113-1 (1) bemerken.

schränkungen bereits für kleine Netzwerke sehr groß ist und (selbst mit Einschränkungen) im Verhältnis zur Größe der Eingabe – hier: zur Anzahl der Entscheidungen, die kategorisiert werden sollen – sehr schnell wächst (*kombinatorische Explosion*): Fordert man nur, dass jede Entscheidung *mindestens* einer Kategorie angehört, und formuliert keine weiteren Einschränkungen, so gibt es schon für ein Zitiernetzwerk mit fünf Entscheidungen 2.147.321.017 Zuordnungsmöglichkeiten.<sup>254</sup> Legt man fest, dass jede Entscheidung *genau* einer Kategorie angehören soll (fordert man also eine Zerlegung), lässt die Zahl der gesuchten Kategorien aber offen, gibt es für fünf Entscheidungen nur noch 52 Möglichkeiten; für zehn sind es allerdings schon wieder 115.975.<sup>255</sup> Selbst, wenn man zusätzlich vorgibt, genau zwei Kategorien zu suchen, gibt es für ein Netzwerk mit zehn Entscheidungen immer noch 511 Aufteilungsmöglichkeiten, mit fünfzehn Entscheidungen schon wieder 16.383.<sup>256</sup> Mit anderen Worten: Die optimale Lösung des Problems lässt sich zu den meisten Netzwerken, die für die juristische Netzwerkforschung interessant sind, nicht durch Evaluation aller Gruppierungsmöglichkeiten berechnen. Man trifft daher Festlegungen, um die Anzahl der Lösungsmöglichkeiten zu verringern, und sucht in ihrem Rahmen nach Näherungslösungen (*approximations*).<sup>257</sup> Dass viele Methoden, die mit indirekten Definitionen arbeiten, jedes Objekt genau einer Gemeinschaft (im Beispiel also: jede Entscheidung genau einer Kategorie) zuordnen, beruht daher unter anderem auf Berechenbarkeitserwägungen – der Suchraum (*search space*) wird bereits im Voraus auf Zerlegungen eingeschränkt. Die Anzahl der Gemeinschaften (hier: Kategorien) wird dann in der Regel offen gelassen, weil man meist zu wenig

<sup>254</sup> Das ist die Anzahl der möglichen *Überdeckungen* (*covers*) einer Menge mit fünf Elementen; siehe dazu den Eintrag in der Online-Enzyklopädie für Zahlenfolgen, [oeis.org/A003465](http://oeis.org/A003465) m.w.N. Eine Überdeckung unterscheidet sich von einer Zerlegung dadurch, dass ein Element in mehreren Teilmengen vorkommen kann (d.h. die Teilmengen müssen nicht paarweise disjunkt sein); siehe *Fortunato*, *Physics Reports* 486 (2010), S. 75 (87).

<sup>255</sup> Allgemein ist die Anzahl möglicher Zerlegungen einer  $n$ -elementigen Menge gegeben durch die  $n$ -te *Bellzahl*,  $B_n$ , auch bekannt als *Exponentialzahl*. Namensgebend ist die Veröffentlichung von *Bell*, *The American Mathematical Monthly* 41 (1934), S. 411–419; siehe auch [oeis.org/A000110](http://oeis.org/A000110) m.w.N. Zu den Bellzahlen im Kontext von Netzwerken u.a. *Fortunato*, *Physics Reports* 486 (2010), S. 75 (87).

<sup>256</sup> Die Anzahl der möglichen Zerlegungen einer  $n$ -elementigen Menge in *genau*  $k$  Teilmengen ist gegeben durch die *Stirlingzahlen zweiter Art*,  $S_{n,k}$ , deren Name auf eine Publikation zurückgeht, die deutlich älter ist als die namensgebende Publikation für die Bell-Zahlen: *Stirling*, *Methodus Differentialis: Sive Tractatus de Summatione et Interpolatione Serierum Infinitarum*, 1730, S. 7–8; dazu [oeis.org/A008277](http://oeis.org/A008277) m.w.N. Zu den Stirlingzahlen zweiter Art im Kontext von Netzwerken u.a. *Newman*, *Physical Review E* 69 (2004), S. 066133-1 (2); *Fortunato*, *Physics Reports* 486 (2010), S. 75 (87).

<sup>257</sup> Zusammenfassend dazu *Fortunato*, *Physics Reports* 486 (2010), S. 75 (83); allgemein zur Klassifizierung der Berechnungsschwierigkeit und zu Approximationsverfahren für algorithmisch anspruchsvolle Probleme *Schöning*, *Theoretische Informatik – kurz gefasst*, 2008, S. 143–176; *Cormen u.a.*, *Introduction to Algorithms*, 2009, S. 1048–1140.

über das untersuchte Netzwerk weiß, um *a priori* die *optimale* Zahl angeben zu können.<sup>258</sup>

Manche Algorithmen erzeugen per Konstruktion in jedem Durchlauf nur eine einzige Zerlegung. Ein Beispiel ist die *Label-Propagierung* (*label propagation*), die mit lokalen Informationen auskommt und auch auf großen Graphen eingesetzt werden kann.<sup>259</sup> Dabei erhalten alle Punkte zu Beginn ein eigenes Label, das in mehreren Berechnungsrunden verändert werden kann. In jeder dieser Runden werden die Punkte in einer zufälligen Reihenfolge abgearbeitet und jeder Punkt erhält das Label, das in seiner Nachbarschaft am häufigsten vorkommt.<sup>260</sup> Das Verfahren endet, wenn in einer Runde keine Änderungen mehr erfolgen. Die gesuchte Zerlegung ergibt sich aus den finalen Labels: Punkte mit dem gleichen Label gehören derselben Gemeinschaft an.<sup>261</sup> Problematisch an dieser Vorgehensweise ist unter anderem, dass die erhaltene Zerlegung von der (dem Zufall überlassenen) Reihenfolge abhängt, in der die Punkte abgearbeitet werden.<sup>262</sup> Da die Methode in jedem Durchlauf nur eine einzige Zerlegung hervorbringt, eignet sie sich außerdem nicht zur Aufdeckung hierarchischer Strukturen.<sup>263</sup>

Algorithmen, welche die Untersuchung hierarchischer Strukturen in Netzwerken ermöglichen, müssen mehrere, ineinander verschachtelte Zerlegungen erzeugen; man spricht auch von *hierarchischem Clustering* (*hierarchical clustering*).<sup>264</sup> Dabei bieten sich zwei entgegengesetzte Vorgehensweisen an, von denen eine als *abwärts* (*top-down*) und die andere als *aufwärts* (*bottom-up*) bezeichnet werden kann. Ansätze, die abwärts arbeiten, beginnen mit der *größten* (*coarsest*) denkbaren Zerlegung, in der alle Punkte in eine einzige Gruppe fallen, und teilen diese Gruppe dann schrittweise in immer mehr Gruppen auf, bis schließlich jeder Punkt seine eigene Gruppe hat – sie sind also *spaltend* (*divisive*). Ansätze, die aufwärts operieren, fangen demgegenüber mit der *feinsten* (*finest*) denkbaren Zerlegung an, in der jeder Punkt seine eigene Gruppe hat, und führen dann einzelne Gruppen sukzessive zusammen, bis es nur noch eine einzige Gruppe gibt – sie sind also *ver-*

<sup>258</sup> Fortunato, Physics Reports 486 (2010), S. 75 (92).

<sup>259</sup> Die einschlägige Publikation ist Raghavan, Albert und Kumara, Physical Review E 76 (2007), S. 036106-1–036106-11.

<sup>260</sup> Falls es mehrere „häufigste“ Label gibt, wird eines dieser Label zufällig ausgewählt; Raghavan, Albert und Kumara, Physical Review E 76 (2007), S. 036106-1 (3).

<sup>261</sup> Man kann sich fragen, ob das Verfahren immer zu einem Abschluss kommt. Raghavan, Albert und Kumara, Physical Review E 76 (2007), S. 036106-1 (9) schreiben dazu: „[A]lthough one can observe the algorithm beginning to converge significantly after about five iterations, the mathematical convergence is hard to prove.“

<sup>262</sup> Diesen Nachteil kann man ausgleichen, indem man die Ergebnisse mehrerer Durchläufe aggregiert; dazu Raghavan, Albert und Kumara, Physical Review E 76 (2007), S. 036106-1 (7–8); man kann ihn sogar als Vorteil ansehen, a.a.O. S. 10.

<sup>263</sup> Zu weiteren Problemen des Verfahrens und möglichen Lösungsansätzen zusammenfassend Fortunato, Physics Reports 486 (2010), S. 75 (127) m.w.N.

<sup>264</sup> Hierarchisches Clustering ist eine Technik der Datenanalyse, die nicht auf Netzwerkdaten beschränkt ist; zu hierarchischem Clustering allgemein Aggarwal, Data Mining, 2015, S. 166–173; speziell im netzwerkanalytischen Kontext Newman, Networks, 2010, S. 386–391.

einend (*agglomerative*).<sup>265</sup> Die ineinander verschachtelten Zerlegungen, die aus beiden Vorgehensweisen resultieren, lassen sich visuell als hierarchischer Baum darstellen, den man Dendrogramm (*dendrogram*) nennt (Abb. 3.24, S. 154).<sup>266</sup>

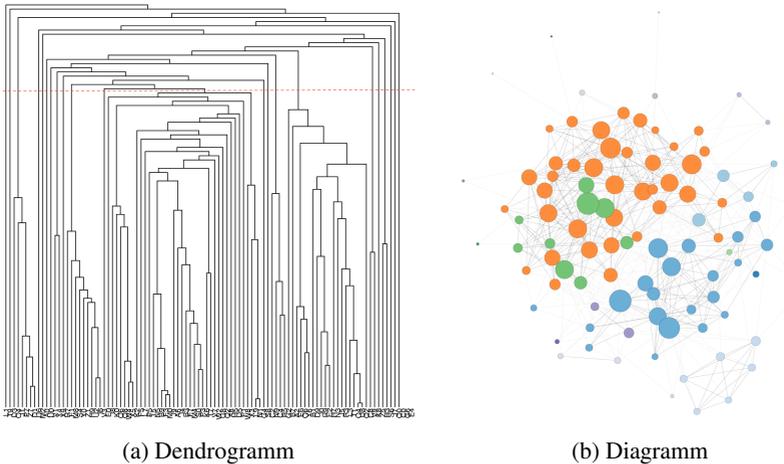


Abbildung 3.24: Hierarchisches Clustering zur Strukturierung des Unternehmensnetzwerks<sup>267</sup>

<sup>265</sup> Die Terminologie ist sowohl in der Netzwerkforschung als auch in der allgemeinen Datenanalyse etabliert; aus Letzterer etwa Aggarwal, *Data Mining*, 2015, S. 167; aus Ersterer Fortunato, *Physics Reports* 486 (2010), S. 75 (93).

<sup>266</sup> Newman, *Networks*, 2010, S. 383–384. Die Bezeichnung als Dendrogramm leitet sich ab von griech. δένδρον (Baum) und γράμμα (Zeichnung). Dendrogramme können horizontal, vertikal oder auch radial dargestellt werden, wobei Letzteres sinnvoll sein kann, um den Eindruck einer Reihenfolge zu vermeiden. Beispiele für den Einsatz von Dendrogrammen mit unterschiedlichen Layouts in der Literatur zur Netzwerkforschung im Recht: *Laumann und Heinz*, *American Bar Foundation Research Journal* 2 (1977), S. 155 (198); *Laumann und Heinz*, *American Bar Foundation Research Journal* 4 (1979), S. 217 (236); *Caldeira*, *Social Networks* 10 (1988), S. 29 (41, 47); *Porter u.a.*, *PNAS* 102 (2005), S. 7057 (7059); *Porter u.a.*, *Physica A* 386 (2007), S. 414 (421–422, 424, 432); *Zhang u.a.*, *Physica A* 387 (2008), S. 1705 (1707, 1710); *Siems*, *Arizona Journal of International & Comparative Law* 27 (2010), S. 747 (779, 781); *Boulet, Barros-Plataiu und Mazzeo*, *Artificial Intelligence and Law* 2016, S. 133 (142).

<sup>267</sup> Abbildung 3.24 (a) (S. 154) zeigt mögliche Strukturierungen des Unternehmensnetzwerks (Personalverflechtungen zwischen Unternehmen in Flatland) in der für hierarchische Methoden gängigen Darstellung als Dendrogramm, basierend auf einem abwärts arbeitenden Verfahren, das die Zwischenzentralität für Linien nutzt. Die horizontale Linie indiziert die (relativ) „beste“ Zerlegung, gemessen an ihrer Modularität (dazu sogleich), in diesem Fall eine Zerlegung in 25 Gemeinschaften. Diese Zerlegung liegt auch der Färbung der Punkte in Abbildung 3.24 (b) (S. 154) zugrunde, wobei allerdings einige Farben mehrfach verwendet wurden, da die zugrundeliegende Farbkarte – eine Standardkarte zur Kodierung qualitativer Merkmale – nur zwanzig Farben hat (man kann sie erweitern, was allerdings dem Sinn standardisierter Farbkarten widerspricht). Den acht Gemeinschaften mit mehreren Mitgliedern (der Rest sind *Singletons*) wurden

Die abwärts und aufwärts arbeitenden Methoden unterscheiden sich vor allem darin, wie sie die Spaltung oder Vereinigung durchführen. Globale Methoden, die abwärts arbeiten, können unter anderem das auf der Mikroebene eingeführte Konzept der Zentralität in seiner Version für Linien fruchtbar machen.<sup>268</sup> Bewährt hat sich insbesondere die Zwischenzentralität der kürzesten Wege (*shortest-path betweenness centrality*), welche auch Abbildung 3.24 (S. 154) zugrunde liegt, aber andere Ansätze sind ebenfalls denkbar (beispielsweise Zwischenzentralität auf Basis von Zufallsläufen, *random-walk betweenness centrality*).<sup>269</sup> Die Verfahren berechnen in mehreren Runden zuerst die Zentralität jeder (verbleibenden) Linie im Graphen und entfernen dann diejenige Linie, die den höchsten Zentralitätswert aufweist.<sup>270</sup> So ergeben sich – wenngleich mit verhältnismäßig großem Berechnungsaufwand<sup>271</sup> – immer feinere, ineinander verschachtelte Zerlegungen, bis schließlich jeder Punkt seine eigene Gruppe hat.

Unabhängig davon, auf welche Weise mehrere verschiedene (nicht notwendig ineinander verschachtelte) Zerlegungen erzeugt werden, stellt sich die Frage, welche Zerlegung des Graphen diejenige ist, die der Gemeinschaftsstruktur des Netzwerks am besten entspricht. Anstatt die Frage intuitiv zu beantworten, kann man die Zerlegungen mithilfe einer Qualitätsfunktion in eine Reihenfolge bringen und anhand dieser Reihenfolge eine „beste“ Zerlegung auswählen.<sup>272</sup> Zur Konstruktion einer Qualitätsfunktion sind verschiedene Ansätze denkbar. Man kann etwa die Beziehungen zählen, die *innerhalb* der identifizierten Gemeinschaften verlaufen, und sie ins Verhältnis setzen zur Anzahl der Beziehungen, die *zwischen* den Gemeinschaften verlaufen (Abdeckung, *coverage*).<sup>273</sup> Oder man vergleicht die beobachtete Dichte der Beziehungen innerhalb der identifizierten Gemeinschaften mit der Dichte der Beziehungen, die man in einem Netzwerk mit ähnlichen Eigenschaften erwarten würde, das keine Gemeinschaftsstruktur

---

allerdings unterschiedliche Farben zugeordnet. Die drei isolierten Punkte (im Dendrogramm unten rechts dargestellt) wurden in Abbildung 3.24 (b) (S. 154) weggelassen, um den Rest des Graphen größer darstellen zu können (sie befanden sich im Layout abseits der anderen Punkte am unteren linken Rand). Die Größe der Punkte, die jeweils für ein Unternehmen stehen, ist zu ihrem Grad proportional. Das Layout ist kräftebasiert, hier allerdings nach dem Algorithmus von *Kamada und Kawai* (vorgestellt in *Kamada und Kawai*, Information Processing Letters 31 (1989), S. 7–15) in der Implementierung von *igraph-0.7.0* für *python-2.7* (Einzelheiten dazu im Online-Appendix).

<sup>268</sup> *Girvan und Newman*, PNAS 99 (2002), S. 7821 (7822–7823); *Newman und Girvan*, Mixing Patterns and Community Structure in Networks, in: *Pastor-Satorras, Rubi und Diaz-Guilera* (Hrsg.), Statistical Mechanics of Complex Networks, 2003, S. 66 (69–70). Zu Zentralitäten bereits S. 129–142.

<sup>269</sup> *Newman und Girvan*, Physical Review E 69 (2004), S. 026113-1 (3).

<sup>270</sup> *Newman und Girvan*, Physical Review E 69 (2004), S. 026113-1 (3).

<sup>271</sup> Dazu *Fortunato*, Physics Reports 486 (2010), S. 75 (97–99).

<sup>272</sup> *Fortunato*, Physics Reports 486 (2010), S. 75 (88–89).

<sup>273</sup> *Fortunato*, Physics Reports 486 (2010), S. 75 (89).

aufweist (Nullmodell, *null model*), und sucht nach derjenigen Zerlegung, welche die Differenz beider Dichten maximiert (Modularität, *modularity*).<sup>274</sup>

Modularität als Funktion, deren Wert es zu maximieren gilt, ist mittlerweile in zahlreiche Optimierungsalgorithmen integriert, darunter auch solche, die aufwärts operieren und ineinander verschachtelte Zerlegungen hervorbringen.<sup>275</sup> Sie hat allerdings Probleme bei der Identifikation von Gemeinschaften, die im Verhältnis zum Gesamtnetzwerk klein sind (Auflösungsschranke, *resolution limit*), liefert zum Teil auch für Netzwerke ohne Gemeinschaftsstruktur hohe Werte, und ihre Höchstwerte sind praktisch unerreichbar, während ganz unterschiedliche Zerlegungen Werte knapp unterhalb des Maximums erzielen können.<sup>276</sup> Alternative Ansätze zur Gemeinschaftserkennung basieren unter anderem auf der Simulation physikalischer Prozesse,<sup>277</sup> statistischer Inferenz<sup>278</sup> oder informationstheoretischen Erwägungen.<sup>279</sup> Für die Erforschung juristischer Zitiernetzwerke erscheinen dabei unter anderem informationstheoretische Ansätze interessant, welche die Gemeinschaftsstruktur eines Netzwerks aus der Bewegung eines Zufallsläufers ableiten, indem sie versuchen, dessen erwartete Schritte möglichst kompakt zu beschreiben.<sup>280</sup>

<sup>274</sup> *Newman und Girvan*, *Physical Review E* 69 (2004), S. 026113-1 (7). Dabei ist insbesondere die Auswahl eines geeigneten Vergleichsnetzwerks problematisch; dazu knapp *Fortunato*, *Physics Reports* 486 (2010), S. 75 (89). Zu den Eigenschaften von Netzwerken, auf deren Ähnlichkeit es dabei ankommen kann, unten S. 165–186; zur Modellierung von Netzwerken S. 193–223.

<sup>275</sup> Beispiel: *Newman*, *Physical Review E* 69 (2004), S. 066133-1–066133-5.

<sup>276</sup> Zu den Limitationen von Modularität als Qualitätskriterium für Zerlegungen zusammenfassend *Fortunato*, *Physics Reports* 486 (2010), S. 75 (111–114). Speziell zur Erreichbarkeit der Höchstwerte im Rahmen der Maximierung von Modularität *Brandes u.a.*, *IEEE Transactions on Knowledge and Data Engineering* 20 (2008), S. 172–188.

<sup>277</sup> Beispiele: *Reichardt und Bornholdt*, *Physical Review Letters* 93 (2004), S. 218701-1–218701-4; *Reichardt und Bornholdt*, *Physical Review E* 74 (2006), S. 016110-1–016110-14.

<sup>278</sup> Beispiel: *Hastings*, *Physical Review E* 74 (2006), S. 035102-1–035102-4.

<sup>279</sup> Beispiel: *Rosvall und Bergstrom*, *PNAS* 104 (2007), S. 7327–7331.

<sup>280</sup> Ein Beispiel für einen solchen Ansatz, der auch auf gerichtete und gewichtete Graphen angewendet werden kann, ist *Rosvall und Bergstrom*, *PNAS* 105 (2008), S. 1118–1123; detailliert beschrieben in *Rosvall, Axelsson und Bergstrom*, *The European Physical Journal Special Topics* 178 (2009), S. 13–23. Die Autoren stellen einen Zusammenhang zwischen dem Problem der Gemeinschaftserkennung und dem Problem der Datenkomprimierung her, bei dem es darum geht, viele Informationen mit möglichst wenigen Bits (die Basisrecheneinheit eines Computers) zu beschreiben und dazu Strukturen in den Daten zu nutzen. Sie motivieren ihren Vorschlag mit kartographischen Erwägungen und nennen die Funktion, die sie zu minimieren versuchen, daher *Map Equation* (*Rosvall und Bergstrom*, *PNAS* 105 (2008), S. 1118 (1120); *Rosvall, Axelsson und Bergstrom*, *The European Physical Journal Special Topics* 178 (2009), S. 13 (19)). Diese Erwägungen sind auf Rechtstexte übertragbar, wie auch Publikationen mit rechtswissenschaftlicher Beteiligung zeigen, die auf der *Map Equation* aufbauen und im Umfeld ihrer Urheber entstanden sind; beispielsweise *Derlén u.a.*, *Europarättslig Tidskrift* 2013, S. 517–535. Die Gemeinschaftserkennung für Netzwerke aus juristischen Texten kann also auch eine Kartographie des Rechts ermöglichen, welche die Rechtswissenschaft in der Vergangenheit (aus sehr

Allen Methoden zur Gemeinschaftserkennung liegt die Annahme zugrunde, dass es im untersuchten Netzwerk eine *Grundwahrheit* (*ground truth*) gibt, die es aufzudecken gilt.<sup>281</sup> Die beschriebenen Verfahren wehren sich aber auch nicht, wenn man sie auf Netzwerke anwendet, in denen eine solche gerade fehlt, und die hierarchische Gemeinschaftsstruktur, die viele Zerlegungsmethoden ermitteln, kann auch allein dem Berechnungsverfahren geschuldet sein.<sup>282</sup> Beispielsweise wurden die Unternehmensdaten für die Zwecke dieser Arbeit bewusst *ohne* Gemeinschaftsstruktur generiert, sodass die in Abbildung 3.24 (S. 154) dargestellte „hierarchische Gemeinschaftsstruktur“ keinerlei Bedeutung hat. Demgegenüber wurden die Rechtsprechungsdaten zwar grundsätzlich *mit* Gemeinschaftsstruktur erzeugt.<sup>283</sup> Die hier konkret vorliegenden Daten sind allerdings nur *eine* mögliche Realisierung des zugrundeliegenden Modells, sodass unklar ist, inwieweit die ihnen vermeintlich innewohnende Struktur dem Zufall geschuldet ist.<sup>284</sup> In der Realität ist in der Regel nicht einmal das Modell bekannt, welches den beobachteten Daten zugrunde liegt (falls es ein solches überhaupt gibt); wäre es anders, bräuchte man die Daten kaum zu untersuchen. Man kann daher analysieren, ob die gefundene Gemeinschaftsstruktur *statistisch signifikant* (*statistically significant*) ist, wobei die Frage danach, was das im hiesigen Kontext bedeuten soll, nicht leicht zu beantworten ist.<sup>285</sup> Oder man versucht, Modelle zu bauen, welche die gefundene Struktur hervorbringen, in der Hoffnung, das untersuchte Netzwerk dadurch besser zu verstehen.<sup>286</sup>

---

verschiedenen Blickwinkeln) immer wieder beschäftigt hat – ein Beispiel ist *Twining*, Northern Ireland Legal Quarterly 50 (1999), S. 12–49, in ganz anderem Kontext (Verhältnis der Rechtswissenschaft zur Geographie) *Frölich*, Vierteljahresschrift für Sozial- und Wirtschaftsgeschichte 27 (1934), S. 40–64.

<sup>281</sup> Zu den Problemen dieser Annahme *Hric, Darst und Fortunato*, Physical Review E 90 (2014), S. 062805-1–062805-19.

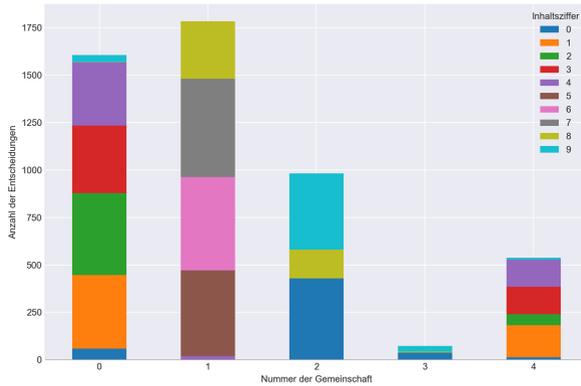
<sup>282</sup> *Fortunato*, Physics Reports 486 (2010), S. 75 (137). Allerdings weisen viele reale Netzwerke jedenfalls intuitiv hierarchische Gemeinschaftsstrukturen auf, wie auch *Porter, Onnela und Mucha*, Notices of the American Mathematical Society 56 (2009), S. 1082 (1084) bemerken.

<sup>283</sup> Dazu unten Fn. 480 (S. 211) und begleitender Text.

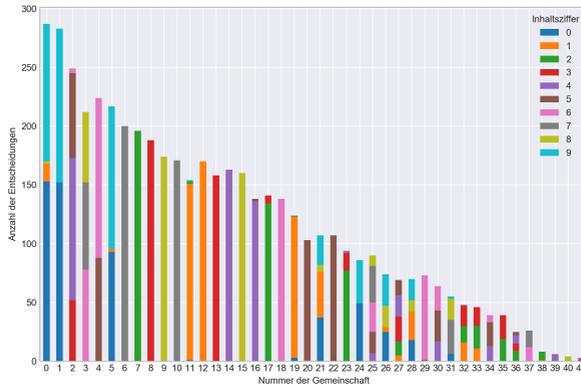
<sup>284</sup> Das Problem lässt sich lösen, indem man aus dem Modell mit unterschiedlichen *seeds* des Zufallszahlengenerators mehrere Graphen erzeugt und dann die Eigenschaften dieses *statistischen Ensembles* (*statistical ensemble*) untersucht.

<sup>285</sup> Beispiele: *Lancichinetti, Radicchi und Ramasco*, Physical Review E 81 (2010), S. 046110-1–046110-9; *Lancichinetti u.a.*, PLoS ONE 6 (2011), e18961-1–e18961-18; *Mirshahvalad u.a.*, PLoS ONE 7 (2012), e33721-1–e33721-7.

<sup>286</sup> Beispiel: *Clauset, Moore und Newman*, Structural Inference of Hierarchies in Networks, in: *Airoldi u.a.* (Hrsg.), Statistical Network Analysis: Models, Issues, and New Directions, 2007, S. 1–13. Zur Modellierung in der juristischen Netzwerkforschung noch unten S. 193–223.



(a) Maximierung von Modularität



(b) Kompression der Beschreibung eines Zufallslaufs

Abbildung 3.25: Inhaltliche Zusammensetzung der von verschiedenen Algorithmen identifizierten Gemeinschaften<sup>287</sup>

<sup>287</sup> Abbildung 3.25 (a) und (b) (S. 158) zeigen die inhaltlichen Zusammensetzungen der Gemeinschaften, die von zwei verschiedenen, mit indirekten Definitionen von Zusammengehörigkeit operierenden Verfahren bei der Analyse des Zitiergraphen in seiner simpelsten Form (ungerichtet und ohne Mehrfachkanten) identifiziert werden; isolierte Punkte sind von der Analyse ausgenommen. Die Nummern der einzelnen Gemeinschaften werden von den Algorithmen vergeben und haben keine Bedeutung; bei der Interpretation ist auf die unterschiedliche Skalierung der y-Achse in den einzelnen Abbildungen zu achten. Das Abbildung 3.25 (a) (S. 158) zugrundeliegende Verfahren (Kurzbezeichnung: *fast greedy*) arbeitet aufwärts und generiert bei dem Versuch, Modularität zu maximieren, ineinander verschachtelte Zerlegungen (Algorithmus von *Newman*, *Physical Review E* 69 (2004), S. 066133-1–066133-5, in der Implementierung von *igraph-0.7.0* für *python-2.7*). Es ist schnell (daher der Namensteil *fast*) und gierig (*greedy*), weil in jeder Berechnungsrunde lokal optimale Teillösungen ausgewählt werden, um eine global optimale Gesamtlösung zu ermitteln. Das hier abgebildete Resultat entspricht der gefundenen Zerlegung

Dass verschiedene indirekte Definitionen von Zusammengehörigkeit für Daten, denen eine Grundwahrheit mit Gemeinschaftsstruktur innewohnt, ganz unterschiedliche Resultate hervorbringen können, verdeutlicht Abbildung 3.25 (S. 158) für das Zitiernetzwerk der flatländischen Rechtsprechung. Der modularitätsoptimierende Algorithmus findet im hier dargestellten Durchlauf fünf Gemeinschaften (Abb. 3.25 (a), S. 158), jener, der den Weg eines Zufallsläufers komprimiert, dagegen 42 (Abb. 3.25 (b), S. 158). Nimmt man allein die zehn Inhaltsziffern als Grundwahrheit an, die ihrerseits eine Zerlegung der Punktmenge des Zitiergraphen definieren, so ist die Aufteilung im ersten Fall deutlich gröber, im zweiten Fall sehr viel feiner als erhofft. Die Schwierigkeiten der Algorithmen, verschiedene Inhaltsziffern in verschiedene Gruppen einzuordnen, hängen auch damit zusammen, dass die Inhaltsziffern zyklisch miteinander verknüpft sind, wodurch sich eine zyklisch überlappende Gemeinschaftsstruktur ergibt.<sup>288</sup> Hinzu kommt, dass das Zitiernetzwerk Entscheidungen verschiedener Gerichte enthält, die auf unterschiedlichen Ebenen in der Gerichtshierarchie stehen, sodass die Vermutung hierarchischer Gemeinschaftsstruktur ebenfalls naheliegt. Nicht nur die Entscheidungsinhalte, sondern auch die Position der entscheidenden Gerichte in der Gerichtshierarchie müssten dann Teil einer vorausgesetzten Grundwahrheit sein. Ob dichte-basierte Ansätze allein ausreichen, um Gruppierungen hervorzu- bringen, die hierarchischer und überlappender Gemeinschaftsstruktur zugleich Rechnung tragen, ist derzeit unklar. Ein anderer Weg, solche Gruppierungen zu erreichen, könnte darin bestehen, dichte-basierte Ansätze mit symmetrie-basierten Ansätzen zu kombinieren (dazu sogleich).

### b) Positionen und Rollen

Symmetrie-basierte Ansätze zur Definition von Zusammengehörigkeit fassen mehrere Elemente in eine Gruppe, sofern sie ähnlich in das Netzwerk eingebettet sind. Was „ähnliche Einbettung“ bedeuten kann, zeigt ein erneuter Blick auf das Gerichtssystem in Flatland. Dieses ist so angelegt, dass in jeder der vier Regionen einem Gericht auf der mittleren Ebene jeweils drei Gerichte auf der unteren Ebene untergeordnet sind.<sup>289</sup> Die vier Gerichte auf der mittleren Ebene sind zwar demselben Gericht unter-, nicht aber denselben Gerichten übergeordnet. Sowohl

---

mit maximaler Modularität. Der Abbildung 3.25 (b) (S. 158) zugrundeliegende Algorithmus ermittelt die Gemeinschaften, indem er versucht, die Ausbreitung von Informationen anhand des Wegs eines Zufallsläufers möglichst komprimiert zu beschreiben (Algorithmus von *Rosvall und Bergstrom*, PNAS 105 (2008), S. 1118–1123, in der Implementierung von *igraph-0.7.0* für *python-2.7*, siehe schon oben Fn. 280 [S. 156]); daher auch die Kurzbezeichnung *infomap*. Dieses Verfahren lässt sich zwar auch auf gerichtete und gewichtete Graphen anwenden; hier wurde auf dem schlichten Zitiergraphen gearbeitet, um den direkten Vergleich mit der modularitätsoptimierenden Methode aus Abbildung 3.25 (a) (S. 158) zu ermöglichen. Die verwendeten Algorithmen sind beide grundsätzlich nichtdeterministisch; die hier dargestellten Resultate wurden mit dem *seed 1234* erzielt.

<sup>288</sup> Dazu Abb. 3.11 (S. 117) und begleitender Text.

<sup>289</sup> Siehe Abb. 3.2 (S. 89).

die Gerichte auf der unteren Ebene als auch jene auf der mittleren Ebene sind in gewisser Hinsicht einander ähnlich, aber unter den unteren Gerichten sind sich manche ähnlicher als andere. Doch wo genau liegt der Unterschied – und wie verhält sich die Ähnlichkeit der Gerichte auf der mittleren Ebene zur Ähnlichkeit der Gerichte auf der unteren Ebene?

Der Versuch, solche Ähnlichkeitsunterschiede formal zu beschreiben, führt zu drei direkten Definitionen von Zusammengehörigkeit, die unterschiedlich streng sind. Die unteren Gerichte innerhalb einer Region stehen mit exakt denselben anderen Gerichten (genauer: mit demselben mittleren Gericht) in Beziehung – sie sind sich strukturell im strengsten denkbaren Sinne ähnlich, was man als *strukturelle Äquivalenz* (*structural equivalence*) bezeichnet.<sup>290</sup> Bei den mittleren Gerichten und den unteren Gerichten aus verschiedenen Regionen ist die Situation etwas komplizierter. Hier beruht die wahrgenommene Ähnlichkeit darauf, dass die Gerichte in ihren jeweiligen Regionen ähnlich in die Gerichtshierarchie eingebunden sind. Diese Gerichtshierarchie ist nichts anderes als die Zerlegung des Graphen, der das Gerichtssystem darstellt, in drei Punktmenge (eine Menge für jede Ebene). Bei der Erfassung der Ähnlichkeitsintuition kann man dann mehr oder weniger streng sein. Verlangt man, dass die Nachbarschaft eines Gerichts der Nachbarschaft eines anderen Gerichts so genau entspricht, dass man die eine auf die andere abbilden könnte (was impliziert, dass die Gerichte gleich viele Nachbarn auf den anderen Ebenen haben), so wird *automorphe Äquivalenz* (*automorphic equivalence*) gefordert.<sup>291</sup> Setzt man hingegen lediglich voraus, dass die Gerichte zu anderen Gerichten auf denselben Ebenen in Beziehung stehen, so wird *reguläre Äquivalenz* (*regular equivalence*) verlangt.<sup>292</sup>

An einem Alltagsbeispiel: Mit Blick auf die Beziehung zwischen Elternteil und Kind sind Eltern derselben (und nur dieser) Kinder strukturell äquivalent, Eltern mit derselben Anzahl von Kindern automorph äquivalent und alle Eltern regulär äquivalent.<sup>293</sup> In der Gerichtslandschaft Flatlands wird der Unterschied zwischen den verschiedenen Äquivalenzen deutlich, wenn man sich vorstellt, Gericht C in Region W würde eingespart (Abb. 3.26 (a), S. 161): Die (verbleibenden) unteren Gerichte innerhalb einzelner Regionen sind dann jeweils strukturell äquivalent. Untere Gerichte in verschiedenen Regionen sind automorph äquivalent, ebenso die mittleren Gerichte in Regionen N, E und S. Das mittlere Gericht in der Region W hingegen ist zu den anderen Gerichten auf der mittleren Ebene lediglich

<sup>290</sup> Vorgeschlagen von Lorrain und White, *The Journal of Mathematical Sociology* 1 (1971), S. 49 (63); zusammenfassend Wasserman und Faust, *Social Network Analysis*, 1994, S. 356–357.

<sup>291</sup> Eingeführt von Winship und Mandel, *Sociological Methodology* 14 (1983), S. 314 (324) auf Basis eines unveröffentlichten Arbeitspapiers von 1974 (dazu Winship, *Social Networks* 10 (1988), S. 209–231) sowie von Everett und Borgatti, *Social Networks* 10 (1988), S. 77–91 (79) auf Basis von Everett, *Social Networks* 7 (1985), S. 353 (355); zusammenfassend Wasserman und Faust, *Social Network Analysis*, 1994, S. 470–471.

<sup>292</sup> Vorgeschlagen von White und Reitz, *Social Networks* 5 (1983), S. 193 (200); zusammenfassend Wasserman und Faust, *Social Network Analysis*, 1994, S. 474–475.

<sup>293</sup> Das Beispiel stammt von Everett, *Social Networks* 7 (1985), S. 353 (355).



erneuter Blick auf die Rechtsprechung in Flatland: Hier liegen Informationen vor sowohl zur formalen Position der Gerichte im Gerichtssystem als auch zu einer Facette ihrer tatsächlichen Position, nämlich zu ihrem Zitierverhalten (Abb. 3.26, S. 161). Wenngleich die mittleren Gerichte *theoretisch* alle auf einer Stufe stehen, scheint das mittlere Gericht in S sich *praktisch* von den anderen mittleren Gerichten zu unterscheiden – eine Positionsanalyse des Zitiernetzwerks würde das mittlere Gericht in S wahrscheinlich nicht mit den anderen mittleren Gerichten gruppieren.

Mit anderen Worten: Die mittleren Gerichte haben im Hierarchienetzwerk zwar dieselbe Position, aber dem mittleren Gericht in S kommt als dasjenige Gericht, an dem sich die Untergerichte im Zitiernetzwerk offenbar regionsunabhängig orientieren, eine andere Rolle zu. In dieser Beobachtung (die man auch quantitativ fassen kann) werden Informationen aus mehreren Netzwerken (hier dem Hierarchienetzwerk und dem Zitiernetzwerk, aggregiert auf die Gerichte) zusammengefasst, die dieselbe Objektmenge, aber unterschiedliche Beziehungsarten beinhalten.<sup>298</sup> So kann man davon sprechen, dass zwei Objekte dieselbe Rolle haben, wenn sie in allen betrachteten Netzwerken eine ähnliche Position einnehmen.<sup>299</sup> Was „ähnliche Position“ bedeutet, ist dann wiederum zu konkretisieren – die feinste Unterteilung der Objektmenge erhält man, wenn (wie im obigen Beispiel für die Positionen der Gerichte) vollständige Übereinstimmung gefordert wird.<sup>300</sup>

Nun ist das deutsche Gerichtssystem sehr viel komplizierter als dasjenige in Flatland, und zwischen Gerichten gibt es noch ganz andere Beziehungen als Zitate und formale Hierarchien, doch macht dies eine Positions- und Rollenanalyse aus juristischer Perspektive jedenfalls für eine Rechtswissenschaft, die sich auch als Justizwissenschaft versteht, gerade interessant. Positionen und Rollen bergen aber möglicherweise noch für ganz andere Fragestellungen Erkenntnispotenziale. Denn fast alle Ausdrucksformen des Rechts (allen voran Gesetze, Verordnungen und Verträge) definieren Netzwerke: Vater, Mutter, Kind, Ehegatte, Elternteil, Verwandter — Verdächtiger, Beschuldigter, Angeklagter, Geschädigter — Verbraucher, Unternehmer, Arbeitgeber, Arbeitnehmer — Eigentümer, Eigentum,

---

differenziert (Beispiel: *Lerner*, Role Assignments, in: *Brandes und Erlebach* (Hrsg.), *Network Analysis*, 2005, S. 216 (216)). Rollen lassen sich außerdem wiederum auf Mikro-, Meso- und Makroebene analysieren (dazu *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 464–465); im Folgenden wird lediglich die Konzeption auf der Mesoebene (*local roles*) eingeführt.

<sup>298</sup> Diese Netzwerke lassen sich zu einem Netzwerk mit mehreren Beziehungsarten zusammenfassen; dazu noch unten S. 224–226.

<sup>299</sup> Man könnte davon sprechen, dass die betroffenen Objekte in diesem Fall *rollenäquivalent* (*role equivalent*) sind, wiche damit aber von der Begriffsverwendung bei *Winship und Mandel*, *Sociological Methodology* 14 (1983), S. 314–344 (aufgegriffen von *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 487–488) ab.

<sup>300</sup> Wird nur ein Netzwerk mit *einer* Beziehungsart betrachtet, fallen Rollen mit Positionen zusammen, sodass sich für die entsprechend vorgehenden Autoren auch die begriffliche Differenzierung erübrigt.

Besitzer, Besitz — Käufer, Verkäufer, Kaufgegenstand, Gegenleistung, Lieferant, Konkurrent — Gesellschaft, Tochtergesellschaft, Konzern — Antragsteller, Antragsgegner — ... Mit all diesen Bezeichnungen gehen abstrakte Bündel von Rechten und Pflichten einher, die gelegentlich neu spezifiziert und täglich im Rechtsverkehr konkretisiert werden. Sie lassen sich als *Rollen* auffassen. Welche Beziehungen zwischen welchen Positionen in welchen realen Netzwerken sind für die Zuschreibung dieser Rollen normativ ausschlaggebend? Wie verhalten sie sich zu den faktischen Beziehungsstrukturen, die in diesen Netzwerken beobachtet werden können? Welches „Beziehungsmodell“ liegt dem Recht zugrunde? Diese Fragen werden in der Rechtswissenschaft seit jeher diskutiert. Die juristische Netzwerkforschung nähert sich ihnen formalisierend und quantifizierend – „welche Beziehungsstrukturen?“ heißt für sie: „Welche Graphen, welche Matrizen?“.

Die bisher vorgestellten Konzepte haben gemeinsam, dass sie *entweder* auf die Suche nach Gemeinschaften *oder* auf die Suche nach Positionen und Rollen zugeschnitten sind. Hat man nun aber etwa das Ziel vor Augen, Gerichtsentscheidungen anhand der strukturellen Eigenschaften des Zitiernetzwerks, von dem sie ein Teil sind, inhaltlich zu klassifizieren, so scheint es sinnvoll, beide Konzepte zu kombinieren. Das ergibt sich daraus, dass ein Zitiernetzwerk, das aus allen Entscheidungen eines Gerichts, einer Gerichtsbarkeit oder eines Gerichtssystems besteht, in der Regel sowohl Entscheidungen mit materiellrechtlichem Schwerpunkt als auch Entscheidungen mit prozessrechtlichem Schwerpunkt enthält. Die Intuition, dass Entscheidungen zusammengehören, die eng miteinander verbunden sind, ist nun vor allem für Entscheidungen zu *materiellrechtlichen Fragen* stimmig. Auch in diesen Entscheidungen treten allerdings regelmäßig *prozessrechtliche Fragen* auf, die mit knappem Verweis auf die einschlägige Rechtsprechung abgehandelt werden.<sup>301</sup> Entscheidungen, die primär prozessrechtliche Fragen betreffen, zitieren sich zwar ebenfalls gegenseitig. Will man sie aber von Entscheidungen mit materiellrechtlichem Schwerpunkt trennen, so hat ein dichtebasierter Ansatz, wie er zum Auffinden von Gemeinschaften üblich ist, mit erheblichem *Zitierrauschen* (*citation noise*) zu kämpfen: Prozessrechtliche Entscheidungen erhalten nicht nur von anderen prozessrechtlichen Entscheidungen, sondern auch von materiellrechtlichen Entscheidungen Zitate, die unter Umständen dazu führen können, dass ein dichtebasierter Algorithmus sie materiellrechtlichen Gemeinschaften zuschlägt, anstatt sie in separaten prozessrechtlichen Gemeinschaften zu erfassen.

Das geschilderte Problem lässt sich möglicherweise vermeiden, indem man der Suche nach Gemeinschaften eine symmetriebasierte Positionsanalyse vorschaltet: Prozedurale Entscheidungen, so die Hypothese, sind anders in das Zitiernetzwerk eingebettet als materiellrechtliche Entscheidungen. Sollte diese Hypothese zutreffen, so könnte man die prozessrechtlichen Entscheidungen anhand einer Positionsanalyse identifizieren und ihnen im Rahmen einer sich anschließenden Gemeinschaftserkennung eine Sonderbehandlung zukommen lassen. Eine

<sup>301</sup> Ein Beispiel ist der Verweis auf „spezifisches Verfassungsrecht“; dazu noch unten S. 270–273.

Möglichkeit wäre dann beispielsweise, die als prozessrechtlich erkannten Entscheidungen aus dem ursprünglichen Zitiernetzwerk zu entfernen und die Gemeinschaftserkennung auf zwei Netzwerken durchzuführen – einem aus prozessrechtlichen und einem aus materiellrechtlichen Entscheidungen. Zur Identifikation der prozessrechtlichen Entscheidungen könnte man auf die in ihren Zitierprofilen enthaltenen Informationen ebenso zurückgreifen wie auf Indikatoren, welche sich aus Maßen konstruieren lassen, die hier auf der Makroebene eingeführt werden;<sup>302</sup> je nach Datenlage könnten auch Informationen aus den Volltexten der Entscheidungen hinzugezogen werden.<sup>303</sup> Momentan ist allerdings weder bekannt, ob die Identifikation prozessrechtlicher Entscheidungen ein systematisches Problem der Gemeinschaftserkennung in Rechtsprechungszitiernetzwerken ist, noch ist gesichert, dass die Hypothese zur unterschiedlichen Einbettung prozessrechtlicher und materiellrechtlicher Entscheidungen in Rechtsprechungszitiernetzwerken tatsächlich zutrifft.<sup>304</sup>

Hat man ein Zitiernetzwerk mit Entscheidungen verschiedener Gerichte vorliegen, so kann sich noch eine weitere Möglichkeit zur Kombination unterschiedlicher Zusammengehörigkeitsvorstellungen auftun. Denn jedenfalls anekdotische Evidenz deutet darauf hin, dass Gerichte (außer sich selbst) insbesondere die Rechtsprechung anderer Gerichte zitieren, die ihnen übergeordnet sind (was auch aus prozessökonomischen Erwägungen sinnvoll ist). Auf der Ebene der einzelnen Entscheidung bedeutet dies, dass ihre ausgehenden Zitate nicht nur vom Inhalt, sondern auch vom Urheber der potenziell zitierten Entscheidungen abhängen. Ein Zitiernetzwerk mit Entscheidungen mehrerer Gerichte verschiedener Ebenen wird daher neben einer thematischen Organisation immer auch eine hierarchische Organisation aufweisen. Da Entscheidungen zudem mehrere verschiedene Themen ansprechen können, muss man dann sowohl überlappende als auch hierarchische Gemeinschaftsstrukturen erwarten. Soll hier eine inhaltsbasierte Entscheidungskategorisierung erreicht werden, so kann es sich anbieten, der Gemeinschaftserkennung eine Rollenanalyse auf dem Hierarchienetzwerk der Gerichte vorzuschalten. Danach ist es möglich, die inhaltliche Klassifizierung nur auf den Entscheidungen der höchsten Gerichtsebene durchzuführen und die dort gewonnenen Erkenntnisse für die Einordnung der Entscheidungen von Gerichten in anderen (normativen)

<sup>302</sup> Zu Zitierprofilen oben bei Abb. 3.19 (S. 133); zu den Maßen, die auf der Makroebene angesiedelt sind, unten S. 165–186.

<sup>303</sup> Generell kann nicht oft genug betont werden, dass die Analyse der Beziehungsstrukturen in einem juristischen Zitiernetzwerk ihr Potenzial vor allem in Kombination mit anderen Methoden entfaltet. Die Interpretation durch Experten oder die volltextbasierte Ähnlichkeitsanalyse mit den Verfahren des *Natural Language Processing* sind zwei Beispiele. Diese Methoden haben ebenso ihre Stärken und Schwächen wie die Analyse von Zitiernetzwerken. Hier soll lediglich gezeigt werden, wie weit man schon mit der Netzwerkanalyse kommen kann.

<sup>304</sup> Anstatt die Positionsanalyse der Gemeinschaftserkennung vorzuschalten, kann man auch versuchen, Positionen und Rollen anhand der Ergebnisse eines Verfahrens zur Gemeinschaftserkennung zu definieren; dazu *Guimerà und Amaral*, *Journal of Statistical Mechanics: Theory and Experiment* 2005, P02001-1–P02001-13.

Rollen zu berücksichtigen. Für die praktische Umsetzung eines solchen Vorgehens fehlt es aber – jedenfalls für Deutschland – an den geeigneten (insbesondere: in geeigneter Form verfügbaren) Rechtsprechungsdaten.

Die vorstehenden Erwägungen zeigen, wie rechtsspezifische Überlegungen in die Entwicklung der Methoden der juristischen Netzwerkforschung einfließen können. Ein solches Vorgehen scheint vor allem dann angemessen, wenn Netzwerke untersucht werden, die „von Juristen gemacht“ sind – also insbesondere Netzwerke aus juristischen Texten. Soweit sich dadurch Differenzierungen auf-tun, die auf nicht-juristische Netzwerke übertragbar sind – im hiesigen Kontext etwa, dass bei der Identifikation von Zusammengehörigem in Zitiernetzwerken unter Umständen Gemeinschaften, Positionen und Rollen gleichzeitig interessieren<sup>305</sup> – könnte die juristische Netzwerkforschung durch Erschließung der hier schlummernden Potenziale auch einen Beitrag zur Entwicklung der allgemeinen Netzwerkforschung leisten.<sup>306</sup>

#### IV. Makroebene

##### 1. Problemstellung: Was ist charakteristisch?

Wie im zweiten Kapitel erläutert, interessiert man sich auf der Makroebene insbesondere dafür, wie sich das betrachtete Netzwerk global charakterisieren lässt (Charakterisierungsfrage).<sup>307</sup> Ähnlich der Wichtigkeits- und der Gruppierungsfrage zerfällt auch die Charakterisierungsfrage prinzipiell in ein Definitions- und ein Messproblem: Was sind die Charakteristika eines Netzwerks? Wie lassen sie sich messen?

Praktisch erfolgt die Charakterisierung eines Netzwerks in der Regel dadurch, dass diverse Messungen angestellt und ihre Ergebnisse mit den Messwerten verglichen werden, die man für andere Netzwerke erhält. Diese Verknüpfung von Definition und Messung ist nicht unproblematisch, allerdings angesichts der kom-

<sup>305</sup> Es könnte sein, dass die Suche der naturwissenschaftlich geprägten allgemeinen Netzwerkforschung nach Möglichkeiten, hierarchische Gemeinschaftsstrukturen und Gemeinschaften in Netzwerken mit mehreren Beziehungsarten zu finden, zur (bewussten oder unbewussten) Wiederentdeckung und algorithmischen Weiterentwicklung der soziologischen Positions- und Rollenkonzepte führt.

<sup>306</sup> Ein konkretes Beispiel: Bei *Rosvall und Bergstrom*, PNAS 105 (2008), S. 1118 (1123) liest man: „We also exclude the only three major journals that span a broad range of scientific disciplines: *Science*, *Nature*, and *Proceedings of the National Academy of Sciences*; the broad scope of these journals otherwise creates an illusion of tighter connections among disciplines, when in fact few readers of the physics articles in *Science* also are close readers of the biomedical articles therein.“ (Hervorhebung im Original). Die Autoren scheinen hier gerade anzuerkennen, dass die genannten Zeitschriften eine atypische Position im wissenschaftlichen Diskurs haben, ohne diese Aussage allerdings zu belegen. Anstatt die Journals von vornherein aus der Analyse auszuschließen, könnte man versuchen, sie anhand einer Positionsanalyse tatsächlich als Zeitschriften mit Sonderstellung zu identifizieren. Möglicherweise würde ein solches Vorgehen noch weitere solcher Zeitschriften zutage fördern, welche die Forscher nicht im Blick hatten.

<sup>307</sup> Dazu oben S. 73.

parativen Komponente, die dem Begriff des Charakteristikums innewohnt, schwer vermeidbar. Die Maße, die bei den Messungen auf der Makroebene verwendet werden, basieren unmittelbar auf graphentheoretischen Grundbegriffen oder sind von Maßen auf der Mikro- und Mesoebene abgeleitet. Ziel ihres Einsatzes ist eine vielschichtige zusammenfassende Beschreibung des Netzwerks, die sowohl Überblicksmöglichkeiten als auch Vergleichsmöglichkeiten eröffnet.

Die Charakterisierung eines Netzwerks kann sich sowohl auf seine Struktur (seine Topologie) als auch auf seine Funktion (sein Verhalten) beziehen.<sup>308</sup> Im ersten Fall kann sie als *statisch*, im zweiten Fall als *dynamisch* bezeichnet werden. Der Unterschied zwischen den Charakterisierungsformen wird deutlich, wenn man sich die Charakterisierung eines Netzwerks als Parallele zur medizinischen Gesundheitskontrolle vorstellt: Das Blutbild liefert eine statische Charakterisierung, das Belastungs-Elektrokardiogramm (Belastungs-EKG) eine dynamische Charakterisierung des menschlichen Gesundheitszustands.<sup>309</sup> Die Parallele illustriert auch, dass für die dynamische Beschreibung entweder längere Beobachtungszeiträume oder aber Simulationen erforderlich sind. Vor diesem Hintergrund konzentrieren sich die nachfolgenden Ausführungen auf die statische Beschreibung von Netzwerken, während die dynamische Beschreibung im Rahmen der Ausführungen zur Zeit und zur Modellierung thematisiert wird.<sup>310</sup>

## 2. Lösungsansätze

Die Maße, die zur Charakterisierung eines Netzwerks auf der Makroebene eingesetzt werden, können die relevanten Informationen unterschiedlich stark komprimieren. Den höchsten Zusammenfassungsgrad erreichen statistische Kennzahlen (a)). Daneben spielen vor allem eindimensionale Verteilungen (b)) und zweidimensionale Verteilungen (c)) in der juristischen Netzwerkforschung eine Rolle.

---

<sup>308</sup> Bei Netzwerken, die durch dezentrale Prozesse entstehen (z.B. Zitiernetzwerken), ist manchmal nicht offensichtlich, welche Funktion oder welches Verhalten sie aufweisen könnten, das sich dynamisch charakterisieren ließe; dazu S. 211–223.

<sup>309</sup> Wer ein Beispiel mit juristischem Bezug präferiert, denke an Bankenbilanzen und Stresstests. Dieses Thema ist aus der Perspektive der Netzwerkforschung unmittelbar interessant (dazu bereits S. 22–28).

<sup>310</sup> Zur Rolle der Zeit in der juristischen Netzwerkforschung S. 187–193; zur Modellierung S. 193–223.

### a) Statistische Kennzahlen

Zu den statistischen Kennzahlen eines Netzwerks gehören zunächst die Angaben zur Größe des korrespondierenden Graphen, namentlich die Anzahl der Punkte (abgekürzt  $n$ ) und die Anzahl der Linien (abgekürzt  $m$ ). Auch die in Abschnitt A. (S. 93–99) erwähnten klassischen Statistiken wie arithmetisches Mittel, Varianz und Standardabweichung, die sich auf alle metrischen Merkmale von Objekten oder Beziehungen anwenden und zu globalen Netzwerkennzahlen aggregieren lassen, fallen hierunter.<sup>311</sup> Es existieren aber noch weitere Statistiken, die speziell auf die Zusammenfassung der Strukturinformationen eines Netzwerks zugeschnitten sind.

Zur Untersuchung der groben Struktur eines Netzwerks lassen sich Zahlen zu den Komponenten des zugehörigen Graphen erheben, etwa ihre Anzahl und Größe.<sup>312</sup> Das ist unter anderem deswegen interessant, weil man in vielen Graphen realer Netzwerke eine mit Abstand *größte Komponente* (*largest component*) findet, in der über 90 % der Punkte enthalten sind und die für die Funktionsfähigkeit des Netzwerks entscheidend ist.<sup>313</sup> Je nachdem, wie die Punkte innerhalb der größten Komponente miteinander verbunden sind, kann das Netzwerk bei zufälligem Ausfall einzelner Elemente oder infolge gezielter Attacken mehr oder weniger schnell auseinanderfallen.<sup>314</sup> Das ist vor allem für die Regulierung kritischer Infrastrukturen relevant, die unter anderem darauf abzielt, eine möglichst robuste Topologie (etwa des Internets, des Stromnetzes, des Straßennetzes oder auch des Interbankenmarktes) sicherzustellen.<sup>315</sup> Innerhalb einzelner Komponenten sind auch die Distanzen zwischen einzelnen Punkten definiert.<sup>316</sup> Man kann dann den Durchmesser der Komponente angeben und auch die durchschnittliche Länge der kürzesten Wege zwischen den einzelnen Punkten ermitteln.<sup>317</sup> Diese Werte sind

<sup>311</sup> Zu diesen Statistiken bereits oben S. 93–99.

<sup>312</sup> Zur Erinnerung: Komponenten sind maximale verbundene Subgraphen; dazu schon oben S. 150.

<sup>313</sup> Newman, Networks, 2010, S. 237–238.

<sup>314</sup> Beispiel: Ein Netzwerk aus Flughäfen und Flugstrecken, in dem ein Großteil des Verkehrs über wenige Drehkreuze (z.B. Frankfurt, Chicago, Dubai) abgewickelt wird, kann durch den Ausfall eines der Drehkreuze empfindlich gestört werden, während der Ausfall eines anderen Flughafens das globale Flugnetz nur geringfügig beeinträchtigen wird. Da es viele Nicht-Drehkreuze gibt, ist es wenig wahrscheinlich, dass zufällig gerade ein Drehkreuz ausfällt. Das Flugnetz ist also robust gegenüber zufälligem Ausfall, aber fragil, wenn einzelne Drehkreuze gezielt attackiert werden.

<sup>315</sup> Zur Robustheit von Netzwerken gegenüber Störungen und Angriffen noch S. 216–223.

<sup>316</sup> Zur graphentheoretischen Distanz bereits oben Fn. 163 (S. 126).

<sup>317</sup> Zur Erinnerung: Der Durchmesser ist die Länge des längsten kürzesten Wegs (d.h. die größte Distanz) zwischen zwei Punkten in einem Graphen. Der Durchmesser einer Komponente ist also die größte Distanz zwischen zwei Punkten in der Komponente. In Graphen mit mehreren Komponenten gibt es mehrere Möglichkeiten, die durchschnittliche Länge der kürzesten Wege zu definieren, beispielsweise kann man die durchschnittliche Länge der kürzesten Wege zunächst

auch in größeren Graphen (bzw. Komponenten) oft recht klein – eine Beobachtung, die man als *Kleine-Welt-Effekt* (*small-world effect*) bezeichnet.<sup>318</sup>

In gerichteten Graphen kann man zwischen stark verbundenen Komponenten und schwach verbundenen Komponenten unterscheiden.<sup>319</sup> Man findet dort mitunter eine große stark verbundene Komponente, der man eine Eingangskomponente (*in-component*) und eine Ausgangskomponente (*out-component*) zuordnen kann. Die Eingangskomponente enthält alle Punkte, von denen aus die stark verbundene Komponente erreicht werden kann; die Ausgangskomponente enthält alle Punkte, die aus der stark verbundenen Komponente heraus erreicht werden können – jeweils einschließlich der Punkte, die selbst der stark verbundenen Komponente angehören.<sup>320</sup> So ergibt sich eine Struktur, die optisch einer Fliege ähnelt (*bow-tie structure*), deren Bestandteile gemeinsam die große schwach verbundene Komponente bilden (Abb. 3.27 (a), S. 169). Die Fliegenstruktur wurde zuerst für den Graphen des *World Wide Web* als Netzwerk aus Webseiten und Hyperlinks festgestellt;<sup>321</sup> für den sie mittlerweile in Zweifel gezogen wird.<sup>322</sup> Der Ansatz lässt sich dennoch auf alle Netzwerke mit asymmetrischen Beziehungen – allen voran Zitiernetzwerke – übertragen, die durch gerichtete Graphen repräsentiert werden.<sup>323</sup> Das Web unterscheidet sich von anderen Zitiernetzwerken (insbesondere Rechtsprechungs zitiernetzwerken) allerdings dadurch, dass auch nach Veröffentlichung einer Webseite Links hinzugefügt werden können. Das führt dazu, dass der *Web Graph* nicht azyklisch ist<sup>324</sup> – anders als viele juristische und

für die einzelnen Komponenten ermitteln und diese dann nach ihrer jeweiligen Größe gewichten; dazu *Newman, Networks*, 2010, S. 184–185.

<sup>318</sup> Die Bezeichnung wurde populär durch *Milgram, Psychology Today* 1 (1967), S. 61–67. *Milgram* führte in den 1960er Jahren in den USA Experimente durch, deren Teilnehmer die Aufgabe hatten, eine Nachricht über möglichst wenige Zwischenstationen zu einem ihnen unbekanntem Empfänger zu befördern. Von 160 Nachrichten, die in einem der Experimente losgeschickt wurden, erreichten 44 (27,50 %) ihr Ziel. Die 44 erfolgreich überbrachten Nachrichten passierten im Median fünf Zwischenstationen, wurden also sechs Mal weitergeleitet. Dieses Ergebnis ist (trotz des ihm innewohnenden *sampling bias*) mit der Zeit als „six degrees of separation“ zu einem wissenschaftlichen und publizistischen Meme geworden; siehe etwa *Strogatz, Nature* 410 (2001), S. 268 (268); titelgebend für *Watts, Six Degrees*, 2003.

<sup>319</sup> Zur Erinnerung: Gerichtete Graphen (hier: Komponenten) sind stark verbunden, sofern jeder ihrer Punkte von jedem anderen Punkt aus erreichbar ist, wenn man sich bei der Bewegung über den Graphen an die Pfeilrichtungen hält. Sie sind schwach verbunden, sofern jeder ihrer Punkte von jedem anderen Punkt aus erreichbar ist, wenn man bei der Bewegung auf dem Graphen die Pfeilrichtungen ignoriert; dazu bereits oben Fn. 156 (S. 124).

<sup>320</sup> *Newman, Networks*, 2010, S. 239–240.

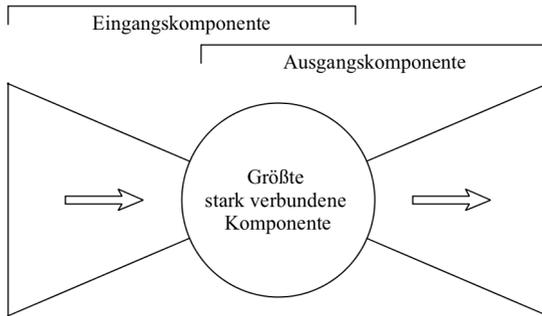
<sup>321</sup> *Broder u.a., Computer Networks* 33 (2000), S. 309–320.

<sup>322</sup> Es besteht der Verdacht, dass die festgestellte Fliegenstruktur ein Artefakt des Crawlers ist, der von den Forschern zur Exploration des *World Wide Web* verwendet wurde; dazu *Meusel u.a., The Journal of Web Science* 1 (2015), S. 33–47 m.w.N.

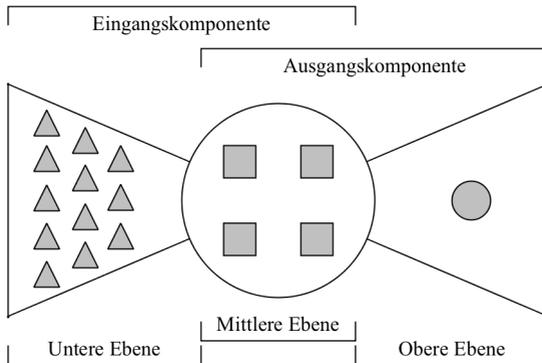
<sup>323</sup> Ebenso *Newman, Strogatz und Watts, Physical Review E* 2001, S. 026118-1 (10); *Cohen und Havlin, Complex Networks*, 2010, S. 123–125.

<sup>324</sup> Zur Erinnerung: Ein gerichteter Graph ist azyklisch, wenn es in ihm keine Folge von Pfeilen gibt, durch deren Nachzeichnen man am Ende wieder zum Ausgangspunkt gelangen könnte.

sonstige wissenschaftliche Zitiergraphen, sofern in ihnen einzelne Dokumente, etwa Gerichtsentscheidungen, Patente oder Aufsätze, als Punkte modelliert werden. In azyklischen Graphen kann es *per definitionem* keine stark verbundenen Komponenten geben; allerdings können die Graphen von abgeleiteten Netzwerken (etwa Zitiernetzwerken, deren Objekte einzelne Gerichte, Erfinder oder Zeitschriften sind) durchaus eine annähernde Fliegenstruktur aufweisen (Abb. 3.27 (b), S. 169).



(a) Schematische Darstellung



(b) Ziterverhalten in Flatland

Abbildung 3.27: Fliegenstruktur von Komponenten in gerichteten Graphen<sup>325</sup>

In Netzwerken mit gerichteten Beziehungen, die keine strikte temporale Struktur haben, kann es vorkommen, dass asymmetrische Beziehungen einander erwidern –

<sup>325</sup> Wer auf die Zitierzahlen in Abbildung 3.14 (S. 121) zurückblickt, wird feststellen, dass die Zuordnung in Abbildung 3.27 (b) (S. 169) nur approximativ zutrifft. Das ist beim Übereinanderlegen von Daten und Modell weniger eine Ausnahme als vielmehr die Regel.

eine Eigenschaft, die man *Reziprozität (reciprocity)* nennt.<sup>326</sup> Diese Eigenschaft kann auf der Makroebene erfasst werden, indem man den Anteil aller Pfeile im Graphen misst, für die es ein Gegenstück gibt, das in die entgegengesetzte Richtung zeigt.<sup>327</sup> Reziprozität spielt in sozialen Netzwerken eine herausgehobene Rolle,<sup>328</sup> kann für die juristische Netzwerkforschung aber auch in anderen Kontexten interessant sein – etwa wenn es darum geht abzuschätzen, wie klar die sachlichen Zuständigkeiten innerhalb einzelner Gerichte zwischen den Spruchkörpern tatsächlich voneinander abgegrenzt sind.<sup>329</sup>

In Tabelle 3.8 (S. 170) sind einige grundlegende Statistiken zu den für Flatland eingeführten Graphen (Zitiergraph, Unternehmensgraph und Personalgraph) zusammengestellt. Auffällig ist insbesondere, dass im Zitiergraphen fast 99 % der Punkte in der größten (schwach verbundenen) Komponente enthalten sind. Der Jurist, der in Flatland Fallrecht recherchiert, kann also 99 % der gesamten Rechtsprechung in maximal 9 Klicks erreichen (der Durchmesser der größten schwach verbundenen Komponente, behandelt als Komponente eines ungerichteten Graphen, ist 9) und muss durchschnittlich mindestens (aber eben auch nur) 4,11 Klicks aufwenden. Das hilft ihm zwar wenig, wenn seine Suchwerkzeuge ihn nicht in die Lage versetzen, bei der Recherche die kürzesten Wege zu den für ihn relevanten Entscheidungen zu finden. Hier können dann aber Maße der Mikroebene zum Einsatz kommen, die auf ein konkretes Erkenntnisinteresse zugeschnitten sind.<sup>330</sup>

<i>Graph</i>	<i>Typ</i>	<i>n</i>	<i>m</i>	<i>c</i>	<i>s<sub>LCC</sub></i>	<i>l<sub>LCC</sub></i>	<i>ø<sub>LCC</sub></i>
Zitiergraph (binär)	gerichtet	5.034	43.310	8,60	0,99	4,11	9
Unternehmensgraph	ungerichtet	100	537	10,74	0,97	2,33	5
Personalgraph	ungerichtet	1.000	705	1,41	0,20	3,00	6

Tabelle 3.8: Zusammenfassende Statistiken zu den für Flatland eingeführten Graphen<sup>331</sup>

<sup>326</sup> Newman, Networks, 2010, S. 204–205; Zweig, Network Analysis Literacy, 2016, S. 99–100.

<sup>327</sup> Erwiderte Beziehungen zeigen sich in gerichteten Graphen also als Zyklen der Länge 2; dazu Newman, Networks, 2010, S. 204–205.

<sup>328</sup> Siehe etwa Stegbauer, Reziprozität, in: Stegbauer und Häußling (Hrsg.), Handbuch Netzwerkforschung, 2010, S. 113–122.

<sup>329</sup> Wenn etwa zwei Senate des BGH permanent aufeinander Bezug nähmen, während alle anderen Senate sich vornehmlich selbst zitierten, könnte dies auf ein Problem der Kompetenzabgrenzung hindeuten. Für eine umfassende Untersuchung der Frage fehlt es bisher allerdings an geeigneten Zitierdaten.

<sup>330</sup> Dazu noch im vierten Kapitel anhand realer Daten zur Rechtsprechung des BVerfG (S. 278–281).

<sup>331</sup> Erfasst wurden die Anzahl der Punkte *n*, die Anzahl der Linien (Pfeile) *m*, der Durchschnittsgrad *c*, der Anteil der Punkte in der größten Komponente (*LCC*) *s*, die durchschnittliche Länge der kürzesten Wege *l* in der größten Komponente und der Durchmesser *ø* der größten Komponente. Bei Unternehmens- und Zitiergraphen wurde jeweils die einfache Projektion (ohne Mehrfachkanten) zugrunde gelegt. Für den Zitiergraphen bezieht sich *s* auf die größte

Ein wichtiger Ansatzpunkt für solche Maße, der auch für Charakterisierungs-bemühungen auf der Makroebene eine Rolle spielt, ist die vielfach beobachtete *Transitivität (transitivity)* von Beziehungen: „Der Freund meines Freundes ist auch mein Freund.“<sup>332</sup> Diese Aussage trifft nun nie auf alle Freunde zu, aber die Wahrscheinlichkeit, dass zwei Personen befreundet sind, ist tendenziell umso höher, je mehr gemeinsame Freunde sie haben. Diese Einsicht lässt sich auf juristische Netzwerke, insbesondere auf Netzwerke aus juristischen Texten, übertragen: Die Wahrscheinlichkeit, dass eine Gerichtsentscheidung oder ein rechtswissenschaftlicher Aufsatz mit einer anderen Entscheidung oder einem anderen Aufsatz durch ein Zitat verbunden ist, wird umso größer sein, je mehr gemeinsame Nachbarn die Entscheidungen oder Aufsätze aufweisen – so könnte man jedenfalls vermuten. Nun sind zahlreiche Gründe denkbar, warum es anders sein könnte. Ist es heute aus der Perspektive eines Gerichts nicht beispielsweise effizient, immer nur die jüngste einschlägige Entscheidung zu zitieren – „m.w.N.“? Kürzen die Autoren rechtswissenschaftlicher Aufsätze ihre Nachweis Ketten nicht ebenso ab? Andererseits: Sind die Kosten der Übernahme von Nachweisen aus anderen Quellen nicht verhältnismäßig gering?<sup>333</sup> Könnte sie in der Rechtsprechung nicht Kontinuität und Stabilität signalisieren? Wenn dies so wäre, ließen sich Transivitätsunterschiede in Ziternetzwerken dann als Indikatoren von Rechtsprechungswandel deuten?

Die Aufgabe der juristischen Netzwerkforschung besteht hier vor allem darin, die Faktengrundlage zu klären. Denn für jeden Graphen kann man abstrakt fragen, wie wahrscheinlich es ist, dass zwei zufällig herausgegriffene Nachbarn eines zufällig herausgegriffenen Punktes ihrerseits Nachbarn (adjazent) sind.<sup>334</sup> So gelangt man zum (globalen) *Clustering-Koeffizienten (clustering coefficient)*, der für ungerichtete Graphen unter anderem dadurch berechnet werden kann, dass man die Anzahl der *Dreiecke* (Zyklen mit drei Punkten, auch bezeichnet als transitive Tripel, *transitive triples*) im Graphen ermittelt und sie zur Anzahl der verbundenen Subgraphen mit drei Punkten (verbundene Tripel, *connected triples*) ins Verhältnis setzt.<sup>335</sup> Entsprechend lässt sich der *lokale Clustering-Koeffizient* eines einzelnen Punktes (ein Maß für die Mikroebene) definieren als Anteil aller

---

schwach verbundene Komponente, die zur Ermittlung von  $l$  und  $\emptyset$  als ungerichteter Graph behandelt wurde; der Durchschnittsgrad ist hier der durchschnittliche Innengrad, der stets dem durchschnittlichen Außengrad entspricht.

<sup>332</sup> Wasserman und Faust, *Social Network Analysis*, 1994, S. 150; Newman, *Networks*, 2010, S. 198–199.

<sup>333</sup> Für wissenschaftliche Publikationen deuten Daten zu Zitierfehlern darauf hin, dass ein Großteil von dem, was zitiert wird, aus den Zitaten anderer Publikationen übernommen wird; dazu Simkin und Roychowdhury, *Scientometrics* 62 (2005), S. 367–384.

<sup>334</sup> Übertragen auf Rechtsprechungsziternetzwerke: Was ist die Wahrscheinlichkeit, dass zwei Entscheidungen, die jeweils über Zitate mit einer dritten Entscheidung verbunden sind, untereinander ebenfalls durch ein Zitat miteinander verbunden sind?

<sup>335</sup> Eine Möglichkeit, den Clustering-Koeffizienten mathematisch zu fassen, ist:  $C = \frac{3 \cdot (\text{Anzahl transiver Tripel})}{\text{Anzahl verbundener Tripel}}$ . Die Multiplikation mit Drei im Zähler ist erforderlich, da für je-

Paare von Nachbarn eines Punktes, die ihrerseits miteinander verbunden sind.<sup>336</sup> Je geringer der lokale Clustering-Koeffizient eines Punktes ist, desto eher kommt ihm in seiner Nachbarschaft eine Brückenfunktion zu.<sup>337</sup> Umgekehrt bedeutet ein höherer lokaler Clustering-Koeffizient höhere *Redundanz* (*redundancy*), die erwünscht oder unerwünscht sein kann – je nachdem, welche Beziehungsart aus wessen Perspektive betrachtet wird.<sup>338</sup>

In gerichteten Graphen gibt es grundsätzlich sechs verschiedene Möglichkeiten für ein Tripel, transitiv zu sein. Führt man aber eine Reihenfolge ein und legt fest, dass Pfeile in dieser Reihenfolge nur „nach hinten“ zeigen dürfen (etwa, um zeitliche Abhängigkeiten zum Ausdruck zu bringen), so bleibt genau ein transitives Tripel, das mit der Reihenfolge kompatibel ist (Abb. 3.28, S. 173). Dies entspricht der Situation in vielen juristischen Zitiergraphen, die aufgrund ihrer zeitlichen Struktur (annähernd) azyklisch sind.<sup>339</sup> In solchen Fällen kann man daher entweder die Richtung der Pfeile ignorieren und den Clustering-Koeffizienten für ungerichtete Graphen als Transitivitätsmaß heranziehen oder aber die Anzahl der Dreiecke im Graphen durch die Anzahl der gerichteten Wege der Länge 2 teilen, um die Transitivität zu messen.<sup>340</sup>

Der Clustering-Koeffizient misst die Dichte von Dreiecken in einem Graphen; ebenso lassen sich auch die Dichten anderer kleiner Subgraphen untersuchen. Kleine Subgraphen mit mehr als zwei Punkten, die häufiger auftreten, als man es bei zufälliger Verteilung der Linien bzw. Pfeile erwarten würde, die also *überzufällig*

---

des Dreieck drei verbundene Tripel gezählt werden. Zu weiteren Formulierungsmöglichkeiten Newman, Networks, 2010, S. 199–200.

<sup>336</sup> Newman, Networks, 2010, S. 201. Mitunter wird der globale Clustering-Koeffizient auch als Durchschnitt aller lokalen Clustering-Koeffizienten definiert, so etwa bei Brinkmeier und Schank, Network Statistics, in: Brandes und Erlebach (Hrsg.), Network Analysis, 2005, S. 293 (303), die den Clustering-Koeffizienten im hiesigen Sinne als Transitivität bezeichnen. Der dadurch erlangte Wert kann aber insbesondere in Graphen, in denen einige wenige Punkte mit sehr vielen anderen Punkten und die meisten Punkte mit wenigen anderen Punkten verbunden sind, in die Irre führen; dazu Newman, Networks, 2010, S. 264 Fn. 8.

<sup>337</sup> In der Soziologie spricht man davon, dass *strukturelle Löcher* (*structural holes*) besetzt werden; dazu Burt, Structural Holes, 1995.

<sup>338</sup> Beispiel: Der Betreiber einer Fabrik wünscht sich Redundanz in seinem Anschluss ans Stromnetz, um beim Ausfall einer einzelnen Leitung nicht ohne Elektrizität zu sein; ein Job-suchender profitiert von wenig Redundanz in seinem Kontaktnetzwerk, da er so mit geringen Kosten (insbesondere zur Aufrechterhaltung von Beziehungen) auf viele verschiedene Informationen zugreifen kann (auch diskutiert unter dem Stichwort „strength of weak ties“, dazu Granovetter, American Journal of Sociology 78 (1973), S. 1360–1380).

<sup>339</sup> Ausnahmen von der Azyklizität juristischer Zitiergraphen können sich etwa daraus ergeben, dass ein Gericht zwei Fälle gleichzeitig entscheidet und die resultierenden Entscheidungen sich gegenseitig zitieren (bzw. aufeinander verweisen).

<sup>340</sup> Das erstgenannte Vorgehen scheint momentan weiter verbreitet zu sein, was auch daran liegen mag, dass der Clustering-Koeffizient in vielen Software-Lösungen nur für ungerichtete Graphen implementiert ist.

sind, werden *Motive (motifs)* genannt,<sup>341</sup> wobei insbesondere die Feststellung der Überzufälligkeit problematisch ist.<sup>342</sup> Sollte das einzige zeitkompatible transitive Tripel in Zitiernetzwerken der Rechtsprechung überzufällig häufig auftreten, so könnte man es möglicherweise als Ausdruck von Rechtsprechungskontinuität deuten und entsprechend als *Stabilitätsmotiv* bezeichnen.

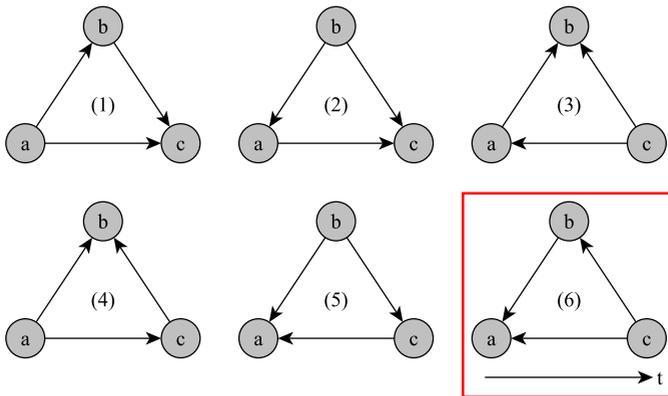


Abbildung 3.28: Transitive Tripel in gerichteten Graphen (zeitkompatibles Tripel hervorgehoben)

Wie ist der Wert einer strukturspezifischen zusammenfassenden Statistik, beispielsweise des Clustering-Koeffizienten, im konkreten Fall zu interpretieren? Diese Frage zielt unter anderem darauf ab, bedeutsame von zufälligen Strukturen zu unterscheiden: Ist die Statistik tatsächlich geeignet, das untersuchte Netzwerk zu charakterisieren? Ein erster Schritt in Richtung der Antwort kann darin bestehen,

<sup>341</sup> Für die Naturwissenschaften *Milo u.a.*, *Science* 298 (2002), S. 824–827; *Shen-Orr u.a.*, *Nature Genetics* 31 (2002), S. 64–68; *Itzkovitz und Alon*, *Physical Review E* 71 (2005), S. 026117-1–026117-9. In der Soziologie ist die Idee, kleine Subgraphen zu zählen und mit statistischen Nullmodellen zu vergleichen, um einiges älter. Gezählt werden insbesondere Subgraphen mit drei Punkten (Triadenzensus, *triad census*); dazu *Holland und Leinhardt*, *American Journal of Sociology* 76 (1970), S. 492–513 sowie *Burt*, *Annual Review of Sociology* 6 (1980), S. 79 (110–116, 125–126, mit Hinweis auf mathematische Grundlagenarbeiten, auf die sich heute auch Naturwissenschaftler beziehen, in Fn. 32). Zusammenfassend zum Triadenzensus *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 564–576 (die Autoren sprechen von Tripeln, wenn sie nur eine Menge von drei Punkten meinen, und von Triaden, wenn sie auch die Linien bzw. Pfeile zwischen diesen Punkten einbeziehen; a.a.O. S. 559). Die Ähnlichkeit der Ansätze von Motiven und Triadenzensus bemerkt auch *Watts*, *Annual Review of Sociology* 30 (2004), S. 243 (253, 263). In der Literatur zur Netzwerkforschung im Recht wird der Triadenzensus aufgegriffen von *Yang u.a.*, *Scientometrics* 105 (2015), S. 1319 (1325–1326).

<sup>342</sup> Dazu *Artzy-Randrup u.a.*, *Science* 305 (2004), S. 1107c, in Reaktion auf *Milo u.a.*, *Science* 298 (2002), S. 824–827, sowie die Replik von *Milo u.a.*, *Science* 305 (2004), S. 1107d; zusammenfassend *Zweig*, *Network Analysis Literacy*, 2016, S. 207–208.

den gemessenen Wert mit Werten zu vergleichen, die für andere Netzwerke (oder für dasselbe Netzwerk zu einem anderen Zeitpunkt) gemessen wurden. Damit wird die Vergleichsfunktion der betrachteten Kennzahl betont; es bleibt allerdings offen, inwieweit die beobachteten Ähnlichkeiten oder Differenzen nicht doch letztlich dem Zufall geschuldet sind. Man kann daher versuchen, den für ein konkretes Netzwerk gemessenen Wert mit Werten zu vergleichen, die sich für Modellnetzwerke messen lassen, deren Beziehungen zufällig verteilt sind, die dem untersuchten Netzwerk aber sonst in allen wesentlichen Punkten gleichen (*Modellwerte*). Problematisch (und sehr ergebnisrelevant) ist dabei insbesondere die Konkretisierung von „in allen wesentlichen Punkten gleichen“.<sup>343</sup> Hat man sich diesbezüglich festgelegt und Modellwerte gemessen, so kann man den Wertebereich bestimmen, in dem der Messwert bei zufälliger Verteilung der Beziehungen erwartungsgemäß liegen müsste, und analysieren, wie sich der konkret gemessene Wert zu diesem Wertebereich verhält. Wie genau der Bereich erwartbarer Werte zu ermitteln ist, hängt wiederum davon ab, wie die Werte verteilt sind, die für die Modellnetzwerke gemessen werden (dazu sogleich).

#### b) Eindimensionale Verteilungen

Inwieweit statistische Kennzahlen hinsichtlich einzelner Merkmale überhaupt aussagekräftig sind, zeigt sich oft erst, wenn man die Verteilungen dieser Merkmale betrachtet. Solche Verteilungen sind eindimensional (univariat), wenn nur ein Merkmal in den Blick genommen wird.<sup>344</sup> Soweit es dabei um empirische Beobachtungen geht, spricht man von *Häufigkeitsverteilungen* (*frequency distributions*);<sup>345</sup> geht es um Zufallsvariablen, spricht man von *Wahrscheinlichkeitsverteilungen* (*probability distributions*).<sup>346</sup> Häufigkeitsverteilungen kann man danach unterscheiden, ob die *Anzahl* der Beobachtungen mit identischer Ausprägung oder der *Anteil* dieser Beobachtungen an der Gesamtheit aller Beobachtungen wiedergegeben wird – im ersten Fall spricht man von *absoluten*, im zweiten Fall von *relativen* Häufigkeitsverteilungen.<sup>347</sup> Sofern das untersuchte Merkmal mindestens ordinal skaliert ist, kann man auch seine *kumulative Häufigkeitsverteilung* betrachten, die angibt, wie viele (bzw. welcher Anteil) der Beobachtungen kleiner oder gleich einem bestimmten Wert sind.<sup>348</sup>

<sup>343</sup> Dazu noch S. 195–211.

<sup>344</sup> Mittag, Statistik, 2014, S. 37; Fahrmeir u.a., Statistik, 2016, S. 29.

<sup>345</sup> Mittag, Statistik, 2014, S. 39; Fahrmeir u.a., Statistik, 2016, S. 30.

<sup>346</sup> Mittag, Statistik, 2014, S. 147. Zufallsvariablen sind Merkmale, deren Ausprägungen als Ergebnisse eines Zufallsvorgangs aufgefasst werden; sie können (wie empirische Merkmale) diskret oder stetig sein und haben den Vorzug, dass man sie abstrakt analysieren kann; Fahrmeir u.a., Statistik, 2016, S. 212.

<sup>347</sup> Fahrmeir u.a., Statistik, 2016, S. 30. Die relative Häufigkeitsverteilung wird aus der absoluten Häufigkeitsverteilung gewonnen, indem man die erfassten Anzahlen jeweils identischer Ausprägungen durch die Anzahl der Beobachtungen teilt.

<sup>348</sup> Fahrmeir u.a., Statistik, 2016, S. 45–46.

Grundsätzlich können Häufigkeitsverteilungen zu jedem geeignet skalierten Merkmal analysiert werden, das für die Objekte oder Beziehungen in einem Netzwerk erhoben wurde. Die Häufigkeitsverteilungen von strukturellen Merkmalen, also von Merkmalen, die aus der Beziehungsstruktur eines Netzwerks abgeleitet werden können, sind für dessen Charakterisierung auf der Makroebene allerdings besonders interessant, da sie sich für jedes Netzwerk anhand seines Graphen ermitteln und für die unterschiedlichsten Netzwerke miteinander vergleichen lassen. Bisher am intensivsten untersucht worden ist die Verteilung des Grads der einzelnen Punkte, die als *Gradverteilung* (*degree distribution*) bezeichnet wird; bei gerichteten Graphen unterscheidet man zwischen Außengradverteilung und Innengradverteilung (*outdegree distribution* und *indegree distribution*).<sup>349</sup> Speziell die Innengradverteilung ist im Abschnitt zur Visualisierung bereits aufgetaucht: Die graphische Darstellung der absoluten Häufigkeiten eingehender Zitate in Flatlands Rechtsprechung (Abb. 3.6, S. 107) zeigt nichts anderes als die Innengradverteilung für den Zitiergraphen in ihrer Darstellung als absolute Häufigkeitsverteilung in nicht-kumulierter und kumulierter Form. Ein erneuter Blick auf die Rechtsprechung in Flatland verdeutlicht nun, dass Innen- und Außengradverteilung sich sehr stark unterscheiden können, obwohl der durchschnittliche Innengrad und der durchschnittliche Außengrad mathematisch zwingend identisch sind (Abb. 3.29, S. 175).

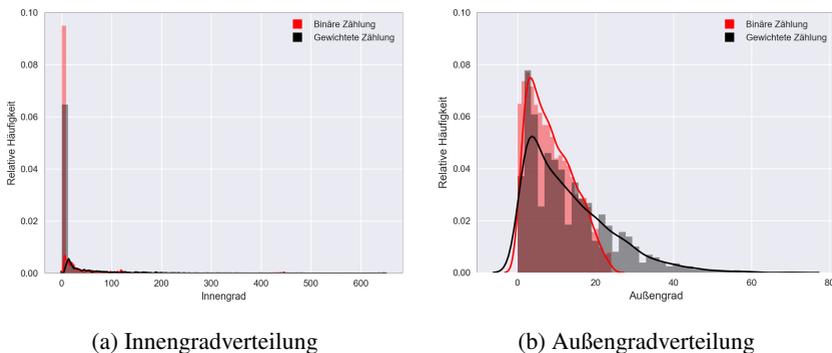


Abbildung 3.29: Innen- und Außengradverteilung für den Zitiergraphen der flatländischen Rechtsprechung<sup>350</sup>

<sup>349</sup> Newman, Networks, 2010, S. 244, 246.

<sup>350</sup> Die Darstellungen unterscheiden sich in mehrerer Hinsicht vom Histogramm in Abbildung 3.6 (S. 107). Erstens sind die Häufigkeiten bei binärer und bei gewichteter Zählung gemeinsam dargestellt, um einen direkten Vergleich zu ermöglichen und die Relevanz der Differenzierung zu verdeutlichen. Zweitens hat nicht mehr jede Gradzahl ihr eigenes Rechteck, sondern es wurden nach einfachen, in der Darstellungssoftware umgesetzten Faustregeln mehrere Klassen (*bins*) identischer Größe gebildet (das reicht offensichtlich nicht aus, um das Darstellungsproblem beim Innengrad zu beheben, dazu sogleich). Drittens ist zu jedem

Zur Charakterisierung des betrachteten Netzwerks interessiert man sich nun nicht primär für eine Gradverteilung als solche, sondern vielmehr für die Wahrscheinlichkeitsverteilung, mit der man sie am besten beschreiben kann – und für die Parameter dieser Verteilung.<sup>351</sup> Das liegt daran, dass Graphen, deren Gradverteilungen bestimmten Wahrscheinlichkeitsverteilungen folgen, abstrakt studiert werden können. Nun hat die allgemeine Netzwerkforschung bereits viel über die Eigenschaften und möglichen Entstehungsprozesse der Netzwerke, denen solche Graphen zugrunde liegen, in Erfahrung gebracht. Von diesem Wissen kann man bei der Untersuchung eines konkreten Netzwerks profitieren, sobald man festgestellt hat, durch welche Wahrscheinlichkeitsverteilung sich die Gradverteilung des zugrundeliegenden Graphen approximieren lässt. Das geschieht, indem man versucht, die Dichtekurven verschiedener Wahrscheinlichkeitsverteilungen an die Daten möglichst gut anzupassen (*curve fitting*).<sup>352</sup> Dann kann man für eine bestimmte Wahrscheinlichkeitsverteilung feststellen, ob diese als Verteilung für die Daten grundsätzlich in Betracht kommt, und unter den grundsätzlich plausiblen Verteilungen nach der relativ besten Anpassung suchen.<sup>353</sup>

Für die statistische Handhabung realer Daten ist es oft hilfreich, wenn sich die vorliegenden Daten gut durch eine Normalverteilung approximieren lassen. Diese stetige Verteilung wird durch zwei Parameter, Erwartungswert  $\mu$  und Varianz  $\sigma^2$  ( $\sigma$  ist die Standardabweichung), eindeutig beschrieben und ist insbesondere *symmetrisch*, wie auch in der Bezeichnung als „Gaußsche Glockenkurve“ zum Ausdruck kommt (Abb. 3.30 (b), S. 177).<sup>354</sup> Die Normalverteilung taucht außerhalb der Netzwerkforschung in den verschiedensten Kontexten auf – unter anderem, weil viele andere (auch diskrete) Verteilungen für eine große Anzahl von Beobachtungen durch sie approximiert werden können.<sup>355</sup> Für strukturelle Merkmale in realen Netzwerken ist allerdings oft charakteristisch, dass sie sehr

---

Histogramm jeweils eine dieser approximierende Dichtekurve eingezeichnet, die durch einen Kerndichteschätzer ermittelt wurde (*kernel density estimation*) und den Vergleich mit den im Folgenden auftretenden Darstellungen stetiger Wahrscheinlichkeitsverteilungen erleichtert. Zu Dichtekurven und Kerndichteschätzern einführend *Fahrmeir u.a.*, Statistik, 2016, S. 80–83, 91–94.

<sup>351</sup> Die Parameter einer Verteilung sind die Stellschrauben, an denen man ihre Form verändern kann; *Fahrmeir u.a.*, Statistik, 2016, S. 82–83. Durch Einsetzen konkreter Werte für die Parameter einer Verteilung wird diese vollständig bestimmt.

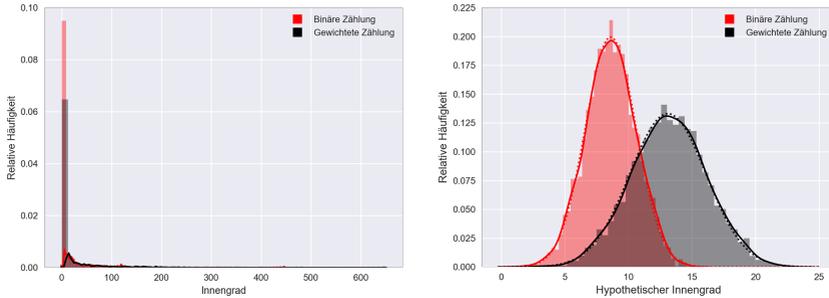
<sup>352</sup> Für schiefe Verteilungen *Clauaset, Shalizi und Newman*, SIAM Review 51 (2009), S. 661 (661–663); *Alstott, Bullmore und Plenz*, PLoS ONE 9 (2014), e85777-1 (1–2).

<sup>353</sup> Im Kontext schiefer Verteilungen *Clauaset, Shalizi und Newman*, SIAM Review 51 (2009), S. 661 (675–680); *Alstott, Bullmore und Plenz*, PLoS ONE 9 (2014), e85777-1 (5–7).

<sup>354</sup> *Fahrmeir u.a.*, Statistik, 2016, S. 271–272.

<sup>355</sup> *Fahrmeir u.a.*, Statistik, 2016, S. 271–272. Dass sich viele Verteilungen für eine große Anzahl von Beobachtungen durch eine Normalverteilung approximieren lassen, hängt damit zusammen, dass die Summe vieler unabhängiger, identisch verteilter Zufallsvariablen approximativ normalverteilt ist. Die theoretische Grundlage hierfür ist der zentrale Grenzwertsatz, dazu *Fahrmeir u.a.*, Statistik, 2016, S. 293–295.

*schief*, also deutlich asymmetrisch verteilt sind.<sup>356</sup> Das gilt besonders für den Grad (bzw. den Innengrad und den Außengrad) der Punkte in den entsprechenden Graphen und äußert sich visuell unter anderem darin, dass man in der klassischen Darstellung der Gradverteilung als Histogramm fast nichts erkennen kann (Abb. 3.30 (a), S. 177).



(a) Schiefe Innengradverteilung (tatsächlich) (b) Symmetrische Innengradverteilung (hypothetisch)

Abbildung 3.30: Tatsächliche und hypothetische Innengradverteilung des Zitiergraphen in Flatland<sup>357</sup>

Diese schiefen Verteilungen in realen Daten lassen sich nicht durch Normalverteilungen approximieren; Werte, die bei Normalverteilungen Ausreißer (*outliers*) wären, sind für sie charakteristisch, und Mittelwerte (Modus, Median und arithmetisches Mittel) in der Regel wenig geeignet, um das Wesen des beobachteten Phänomens zu erfassen.<sup>358</sup> Es steht aber eine ganze Familie schiefer Wahrscheinlichkeitsverteilungen zur Verfügung, die zur Anpassung an schief verteilte Daten

<sup>356</sup> Zur Terminologie *Fahrmeir u.a.*, Statistik, 2016, S. 44. Beispiele für schiefe Verteilungen in realen Netzwerken finden sich bei *Albert und Barabási*, Reviews of Modern Physics 74 (2002), S. 47 (49–54), *Newman*, SIAM Review 45 (2003), S. 167 (182), und *Boccaletti u.a.*, Physics Reports 424 (2006), S. 175 (251–275), jeweils m.w.N. In der Statistik unterscheidet man *rechtsschiefe* (*linkssteile*) und *linksschiefe* (*rechtssteile*) Verteilungen; *Fahrmeir u.a.*, Statistik, 2016, S. 44. Die Begriffe verwirren allerdings mehr als sie nützen, denn die Histogramme rechtsschiefer Verteilungen lehnen optisch nach links (s. Abb. 3.29, S. 175). In der Netzwerkforschung bedeutet „schief“ in den allermeisten Fällen „rechtsschief“.

<sup>357</sup> Abbildung 3.30 (a) (S. 177) zeigt noch einmal die Innengradverteilung des flatländischen Ziternetzwerks. Abbildung 3.30 (b) (S. 177) zeigt mögliche Innengradverteilungen für 5.034 Punkte mit durchschnittlichem Innengrad (d.h. arithmetischem Mittel  $\mu$ ) von 8,6035 (binär) bzw. 13,1013 (gewichtet) sowie Standardabweichung  $\sigma$  von 2 (binär) bzw. 3 (gewichtet) unter der Annahme, dass der Innengrad einer Normalverteilung folgt. Die durchgezogenen Linien entsprechen den Kerndichteschätzern, die gestrichelten Linien den Normalverteilungen mit  $\mu = 8,6035$  und  $\sigma = 2$  bzw.  $\mu = 13,1013$  und  $\sigma = 3$ . Dass Letztere so dicht bei Ersteren liegen, indiziert visuell, dass die Daten sich durch eine Normalverteilung annähernd beschreiben lassen (was wenig verwunderlich ist, da die Histogramme durch Ziehen aus einer Normalverteilung generiert wurden).

<sup>358</sup> *Clauset, Shalizi und Newman*, SIAM Review 51 (2009), S. 661 (661–662).

in Betracht kommen. In der allgemeinen wie in der juristischen Netzwerkforschung erfährt dabei die Gradverteilung, die ab einem bestimmten Minimalwert ( $x_{min}$ ) einem Potenzgesetz mit Parameter  $\alpha$  folgt (*power-law degree distribution*), besonders viel Aufmerksamkeit.<sup>359</sup> Diese ist dadurch gekennzeichnet, dass die Wahrscheinlichkeit  $p_k$ , einen Punkt mit Grad  $k$  zu finden, als negative Potenz des Grades abnimmt:  $p_k \sim k^{-\alpha}$ .<sup>360</sup>

Netzwerke mit Graphen, deren Gradverteilungen (jedenfalls ab einem bestimmten Grad  $x_{min}$  aufwärts) einem Potenzgesetz folgen, werden auch als *skalenfrei* (*scale-free*) bezeichnet.<sup>361</sup> Es ist üblich, für die visuelle Darstellung dieser Gradverteilungen logarithmische Achsen zu wählen, sodass sich eine an sie angepasste Wahrscheinlichkeitsfunktion (*probability density function*, PDF,  $p(X)$ ), die einem Potenzgesetz folgt, als annähernd gerade Linie zeigt.<sup>362</sup> Um visuell beurteilen zu können, wie gut eine Wahrscheinlichkeitsverteilung auf die Daten passt, hilft es manchmal, auch die komplementäre kumulative Verteilungsfunktion (*complementary cumulative distribution function*, CCDF,  $p(X \geq x)$ ) darzustellen (Abb. 3.31, S. 179).<sup>363</sup>

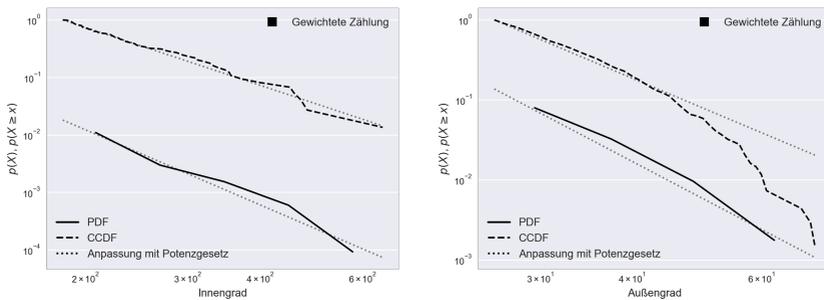
<sup>359</sup> Zur Identifikation von Potenzgesetzen allgemein *Clauset, Shalizi und Newman*, *SIAM Review* 51 (2009), S. 661–703. In der juristischen Netzwerkforschung werden Potenzgesetze beispielsweise diskutiert von *Strandburg u.a.*, *Berkeley Technology Law Journal* 21 (2006), S. 1293 (1302–1305); *Smith*, *San Diego Law Review* 44 (2007), S. 309 (317–322); *Cross, Smith und Tomarchio*, *Emory Law Journal* 57 (2008), S. 1227 (1237–1246); *Katz, Stafford und Provins*, *Georgia State University Law Review* 24 (2008), S. 977 (996–1000); *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457 (497–503).

<sup>360</sup> *Barabási und Albert*, *Science* 286 (1999), S. 509 (510); *Watts*, *Annual Review of Sociology* 30 (2004), S. 243 (250); *Newman*, *Networks*, 2010, S. 247–249. Allgemein zu Potenzgesetzen *Newman*, *Contemporary Physics* 46 (2005), S. 323–351.

<sup>361</sup> *Barabási*, *Network Science*, 2016, S. 117. Für die Bezeichnung der betroffenen Netzwerke als skalenfrei werden verschiedene Begründungen gegeben. Ein Erklärungsansatz stellt darauf ab, dass in einem Netzwerk mit  $n$  Objekten, dessen Gradverteilung einem Potenzgesetz folgt, für  $n \rightarrow \infty$  auch die Varianz unendlich groß wird, sodass der Durchschnittsgrad (arithmetisches Mittel) kein Maßstab (*scale*) für das betrachtete System ist; so zu finden bei *Barrat, Barthélemy und Vespignani*, *Dynamical Processes on Complex Networks*, 2008, S. 40–41. Eine andere Begründung stellt auf die mathematischen Eigenschaften der Potenzgesetze ab: Multipliziert man den Grad  $k$  mit einer Konstante, so verändert sich  $p_k$  (nur) proportional; so zu finden bei *Newman*, *SIAM Review* 45 (2003), S. 167 (186 Fn. 13); *Estrada*, *The Structure of Complex Networks*, 2012, S. 29. Praktisch bedeutet das, dass man für verschiedene Wertebereiche Beobachtungen machen kann, an deren Form sich der Wertebereich nicht erkennen lässt; dazu *Zweig*, *Network Analysis Literacy*, 2016, S. 166–168 (mit kritischem Hinweis zur insgesamt irreführenden Terminologie).

<sup>362</sup> Das liegt daran, dass multiplikatives Wachstum auf der logarithmischen Skala zu additivem Wachstum wird:  $\log_b(x \cdot y) = \log_b(x) + \log_b(y)$ .

<sup>363</sup> *Clauset, Shalizi und Newman*, *SIAM Review* 51 (2009), S. 661 (667) (dort wird die CCDF allerdings mit CDF abgekürzt); *Alstott, Bullmore und Plenz*, *PLoS ONE* 9 (2014), e85777-1 (2–3). Die Definition der CCDF als  $p(X \geq x)$  liegt dem zur Visualisierung verwendeten Programm zugrunde, ist aber unüblich. Die CCDF-Kurven in den nachfolgenden Abbildungen sind daher gegenüber der Standarddefinition ( $p(X > x)$ ) ein Stück versetzt.



(a) Innengradverteilung  
(Potenzgesetz potenziell passend)

(b) Außengradverteilung  
(Potenzgesetz sichtbar unpassend)

Abbildung 3.31: Anpassung von Potenzgesetzen an die Gradverteilungen des Zitiergraphen der Rechtsprechung in Flatland<sup>364</sup>

In Abbildung 3.31 (b) (S. 179) zeigt der visuelle Vergleich von komplementärer kumulativer Häufigkeitsverteilung und bester Anpassung durch ein Potenzgesetz deutlich, dass das Potenzgesetz die Außengradverteilung des Zitiergraphen schlecht beschreibt. Bei der Innengradverteilung in Abbildung 3.31 (a) (S. 179) kann man auf den ersten Blick optimistischer sein – und tatsächlich wurden auf Basis visueller Darstellungen in der Vergangenheit zahlreiche Potenzgesetze in Daten „gefunden“.<sup>365</sup> Bei diesem Vorgehen bleibt allerdings die Frage offen, ob sich nicht noch eine bessere Anpassung finden lässt, ohne das der Wahrscheinlichkeitsverteilung zugrunde liegende Modell zu sehr an die Daten anzupassen

<sup>364</sup> Abbildung 3.31 (a) (S. 179) zeigt die relative Häufigkeitsverteilung (PDF) und die komplementäre kumulative Häufigkeitsverteilung (CCDF) des Innengrads im Zitiernetzwerk der flatländischen Rechtsprechung bei gewichteter Zählung sowie die jeweils beste Anpassung durch ein Potenzgesetz. Abbildung 3.31 (b) (S. 179) zeigt die gleichen Verteilungen und Anpassungen für den Außengrad dieses Netzwerks. Wie an der Skalierung der x-Achse erkennbar ist, wird die Anpassung jeweils nur für einen kleinen Teil der Daten durchgeführt; für die Darstellung der PDF werden diese Daten in logarithmisch skalierte Klassen eingeteilt (*logarithmic binning*). Die Zahl der Datenpunkte, die in die Anpassung einbezogen werden, und die Anzahl der gebildeten Klassen ist dabei so gering, dass nicht zuverlässig auf ein Potenzgesetz geschlossen werden kann.

<sup>365</sup> In der juristischen Netzwerkforschung sind insbesondere Streudiagramme mit logarithmischen Achsen verbreitet, die viel Rauschen (*noise*) im rechten Teil der Verteilung zeigen. Potenzgesetze werden so „identifiziert“ von *Post und Eisen*, *The Journal of Legal Studies* 29 (2000), S. 545 (572–574) und *Smith*, *San Diego Law Review* 44 (2007), S. 309 (327–329, 331–335); auf Basis herkömmlicher Histogramme *Puig*, *The European Journal of International Law* 25 (2014), S. 387 (420); *Neale*, *Journal of Open Access to Law* 1 (2013), S. 1 (21–22). Vorsichtiger etwa *Katz und Stafford*, *Ohio State Law Journal* 71 (2010), S. 457 (497–500); *Katz u.a.*, *Journal of Legal Education* 61 (2011), S. 76 (94–95); sowie *Schaper*, *The Structure and Organization of EU Law in the Field of Direct Taxes*, 2013, S. 152–154.

(sog. *Überanpassung*, *overfitting*).<sup>366</sup> Bei schiefen Verteilungen kommen neben Potenzgesetzen insbesondere Exponentialverteilungen und Mischformen von Exponentialverteilungen und Potenzgesetzen (*crossover distributions*) in Betracht.<sup>367</sup> Meist reichen die verfügbaren Daten zwar nicht aus, um die einzelnen schiefen Verteilungen zuverlässig auseinanderzuhalten.<sup>368</sup> Der Vergleich bewahrt aber davor, ein Potenzgesetz zu proklamieren, wo auch andere Verteilungen einschlägig sein könnten. Das ist insbesondere dann relevant, wenn auf Basis von Annahmen über die Verteilung einzelner Merkmale Aussagen über das Verhalten des Netzwerks oder Prozesse seines Zustandekommens und seiner zukünftigen Entwicklung gemacht werden sollen.

Jedenfalls für die juristische Netzwerkforschung scheint es nach alledem sinnvoll, sich mit einer groben Einordnung der beobachteten Häufigkeitsverteilung auf einer Skala von symmetrisch bis sehr schief zu begnügen und den Fokus dann darauf zu lenken, was man aus diesem Befund für das Recht lernen kann. Hier könnten sich insbesondere für Rechtssoziologie, Rechtsvergleichung und Rechtstheorie interessante Erkenntnisse ergeben. Die Rechtssoziologie kann Gradverteilungen in den sozialen Netzwerken juristischer Akteure untersuchen und in Kooperation mit der Rechtsvergleichung über die Grenzen von Rechtssystemen hinweg miteinander in Beziehung setzen. Die Rechtsvergleichung selbst kann fragen, wie strukturelle Merkmale in den Zitiernetzwerken der Rechtsprechung unterschiedlicher Jurisdiktionen verteilt sind, und mit Rechtssoziologie und Rechtstheorie gemeinsam nach Gründen für die beobachteten Gemeinsamkeiten und Unterschiede suchen. Besonders für die Rechtstheorie interessant ist schließlich die in anderen Kontexten häufig beobachtete Selbstähnlichkeit (*self-similarity*) der analysierten Strukturen, die sich auch im Recht zeigen könnte.<sup>369</sup> Selbstähnlichkeit ist die Eigenschaft eines Objekts, in der Form (mindestens) einem seiner Teile ähnlich zu sein. Mathematische Objekte, die Selbstähnlichkeit aufweisen, heißen *Fraktale* (*fractals*);<sup>370</sup> aus dem Alltag bekannte Beispiele sind Kristalle, Bäume und Meeresküsten. Innen- und Außengradverteilung sind Aspekte der Form, die man etwa bei Zitiernetzwerken mit Entscheidungen verschiedener Gerichte sowohl für Entscheidungen aller Gerichte als auch für Entscheidungen ausgewählter Gerichte betrachten kann (Abb. 3.32, S. 181). Programmatischen Beiträgen, die das Recht als System mit fraktaler Struktur beschreiben oder all-

<sup>366</sup> Dazu speziell im Kontext von schiefen Verteilungen *Alstott, Bullmore und Plenz*, *PLoS ONE* 9 (2014), e85777-1 (5–6); im juristischen Kontext u.a. problematisiert von *Katz*, *Emory Law Journal* 62 (2013), S. 909 (942, insbes. Fn. 169); *Rich*, *University of Pennsylvania Law Review* 164 (2016), S. 871 (884).

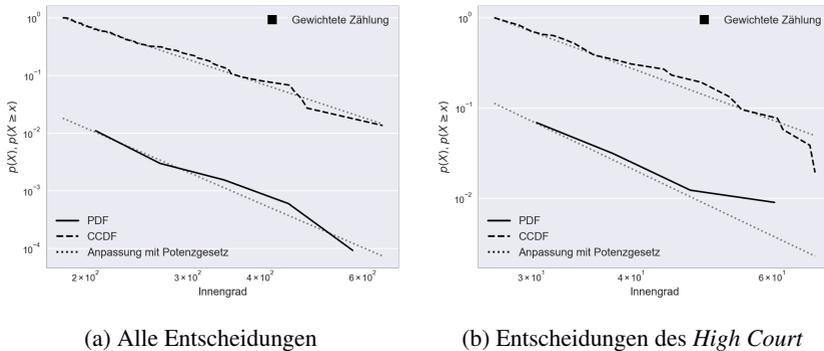
<sup>367</sup> *Barabási*, *Network Science*, 2016, S. 148–151.

<sup>368</sup> Das gilt auch für die Zitierdaten der flatländischen Rechtsprechung, wie sich im Online-Appendix nachvollziehen lässt.

<sup>369</sup> In diese Richtung mit quantitativen Methoden insbesondere *Post und Eisen*, *The Journal of Legal Studies* 29 (2000), S. 545–584.

<sup>370</sup> Grundlegend *Mandelbrot*, *Les Objets Fractals*, 1975, weiterentwickelt in *Mandelbrot*, *The Fractal Geometry of Nature*, 1982; zuvor bereits *Mandelbrot*, *Science* 156 (1967), S. 636–638.

gemeiner die Terminologie der Theorie komplexer adaptiver Systeme (*complex adaptive systems*, CAS) aufgreifen, könnte als Ergebnis *intrasystemischer Rechtsvergleichung* eine empirische Fundierung gegeben werden.<sup>371</sup>



(a) Alle Entscheidungen

(b) Entscheidungen des High Court

Abbildung 3.32: (Partielle) Selbstähnlichkeit der Innengradverteilungen im Zitiergraphen der Rechtsprechung in Flatland<sup>372</sup>

### c) Zweidimensionale Verteilungen

Betrachtet man statt der Verteilung *eines* Merkmals die gemeinsame Verteilung *zweier* Merkmale, so erhält man eine zweidimensionale (bivariate) Verteilung.<sup>373</sup> Diese lässt sich in tabellarischer Form darstellen (*Kontingenztafel*, *contingency table*), wobei die Felder im Inneren der Tabelle die *gemeinsame Verteilung* der Merkmale und jene an den Rändern deren jeweilige *Randhäufigkeiten* angeben.<sup>374</sup> Die gemeinsame Verteilung kann aber auch als Matrix aufgefasst werden.<sup>375</sup> In dieser Form ist sie bereits im Abschnitt zur Visualisierung aufgetaucht: Die in Abbildung 3.14 (S. 121) gezeigten Matrizen sind nichts anderes als die gemeinsame Verteilung der Merkmale „Zitierendes Gericht“ und „Zitiertes Gericht“, die sich den einzelnen Zitaten (!) anhand des Aktenzeichens der zitierenden bzw. zitierten Entscheidung zuschreiben lassen, mit unterschiedlichen Zeilen- und Spaltenanordnungen. Ähnlich kann man für die Unternehmen in Flatland erfassen, inwieweit sie gerade mit Unternehmen aus derselben Industrie personell verflochten sind, wobei die Symmetrie der Verflechtungsbeziehungen dazu führt, dass die Matrix

<sup>371</sup> Zahlreiche Beispiele für solche Beiträge finden sich in Kapitel 1 Fn. 20 (S. 9).

<sup>372</sup> Zur Darstellung in Abbildung 3.32 (a) (S. 181) schon Fn. 364 (S. 179). Abbildung 3.32 (b) (S. 181) zeigt die gleichen Verteilungen und Anpassungen für das auf Entscheidungen des *High Court* eingeschränkte Zitieretzwerk.

<sup>373</sup> Mittag, Statistik, 2014, S. 95; Fahrmeir u.a., Statistik, 2016, S. 105.

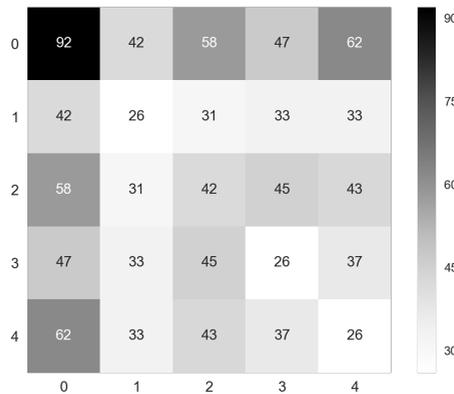
<sup>374</sup> Mittag, Statistik, 2014, S. 96–97; Fahrmeir u.a., Statistik, 2016, S. 105–110.

<sup>375</sup> Die Randverteilungen erhält man dann, indem man einen (geeignet dimensionierten) Vektor aus Einsen mit der Matrix multipliziert – je nachdem, welche Randverteilung man erhalten will, entweder als Zeilenvektor von links (dann untere Randverteilung) oder als Spaltenvektor von rechts (dann rechte Randverteilung).

symmetrisch ist und beide Randhäufigkeiten in der Kontingenztabelle identisch sind (Abb. 3.33, S. 182).

		Endpunkt 2					$\Sigma$
		0	1	2	3	4	
Endpunkt 1	0	92	42	58	47	62	301
	1	42	26	31	33	33	165
	2	58	31	42	45	43	219
	3	47	33	45	26	37	188
	4	62	33	43	37	26	201
$\Sigma$		301	165	219	188	201	1074

(a) Als Kontingenztabelle



(b) Als Matrix

Abbildung 3.33: Personalverflechtungen zwischen Unternehmen der einzelnen Industrien in Flatland<sup>376</sup>

Für die Netzwerkforschung sind vor allem die gemeinsamen Verteilungen der Merkmale interessant, die sich Beziehungen anhand ihrer Endpunkte zuschreiben lassen (z.B. die gemeinsame Verteilung der Spruchkörper an Start- und Endpunkt eines Rechtsprechungszitats) – hierin kommt ihre Konzentration auf das Relationale zum Ausdruck. In der Regel wird an beiden Endpunkten dasselbe Merkmal betrachtet, womit quadratische Matrizen analysiert und visualisiert werden können.<sup>377</sup> Man stellt sich dann vor allem die Frage, inwieweit die beiden

<sup>376</sup> Abbildung 3.33 (b) (S. 182) ist nichts anderes als eine kolorierte Version von Abbildung 3.33 (a) (S. 182) ohne die Randhäufigkeiten.

<sup>377</sup> Bei Netzwerken mit asymmetrischen Beziehungen (Graphen mit Pfeilen) kann es vorkommen, dass eine Merkmalsausprägung nicht an beiden Enden der Beziehungen auftritt. Man kann dann trotzdem quadratische Matrizen analysieren, indem man die fehlenden Ausprägungen als Zeile bzw. Spalte von Nullen ergänzt.

Endpunkte einer Beziehung dieselbe Merkmalsausprägung oder einander ähnliche Merkmalsausprägungen aufweisen. Netzwerke, in denen das verhältnismäßig häufig der Fall ist, zeigen in dieser Hinsicht *assortative Vermischung* (*assortative mixing*), die insbesondere im soziologischen Kontext auch als *Homophilie* (*homophily*) bezeichnet wird.<sup>378</sup> Ihre Essenz kommt in der Phrase „birds of a feather flock together“ zum Ausdruck.<sup>379</sup> Demgegenüber weisen Netzwerke, in denen die Endpunkte einzelner Beziehungen verhältnismäßig häufig unterschiedliche Merkmalsausprägungen haben, *disassortative Vermischung* (*disassortative mixing*) bzw. *Heterophilie* (*heterophily*) auf.<sup>380</sup>

Es stellt sich allerdings die Frage, was „verhältnismäßig häufig“ bedeutet. Die zweidimensionale Verteilung einzelner (von Objektmerkmalen abgeleiteter) Beziehungsmerkmale bildet hier lediglich den Ausgangspunkt der Antwort, die wiederum auf eine statistische Kennzahl hinausläuft. Wie eine solche Kennzahl aussehen kann, hängt davon ab, welche Skala die untersuchten Merkmale haben.<sup>381</sup> Betrachtet man nominale oder ordinale Merkmale, etwa das Gericht, das eine zitierende oder zitierte Entscheidung getroffen hat, so kann man zunächst die Anzahl der Beziehungen ermitteln, deren Endpunkte dieselbe Merkmalsausprägung aufweisen, und davon die Anzahl solcher Beziehungen abziehen, die man bei zufälliger Verteilung der Beziehungen zwischen allen Objekten erwarten würde. Teilt man anschließend durch die Anzahl der insgesamt vorhandenen Beziehungen, so erhält man ein Maß, dessen Werte immer im Intervall  $[-1, 1]$  liegen. Diese Werte sind genau dann positiv, wenn mehr Beziehungen zwischen Objekten desselben Typs verlaufen als man bei zufälliger Verteilung erwarten würde (dann nimmt man *Assortativität* an), und genau dann negativ, wenn weniger Beziehungen entsprechend verlaufen (dann nimmt man *Disassortativität* an).<sup>382</sup> So entspricht das zugrunde gelegte Maß der Modularität, die auf der Mesoebene bereits zur Beurteilung der Qualität von Zerlegungen eingeführt wurde (das dort betrachtete nominale Merkmal ist die Gemeinschaft, der die Endpunkte einer Beziehung in der Zerlegung zugeordnet werden).<sup>383</sup> Bevor man die Werte dieses Maßes für verschiedene Netzwerke miteinander vergleichen kann, muss man sie

<sup>378</sup> Newman, Networks, 2010, S. 222.

<sup>379</sup> Diese Phrase wird in der soziologischen Literatur zur Kennzeichnung des Phänomens häufig zitiert; dazu McPherson, Smith-Lovin und Cook, Annual Review of Sociology 27 (2001), S. 415 (416–417). Ähnliche Formulierungen finden sich bereits in der Antike; sie werden aufgegriffen von Burton, The Anatomy of Melancholy, 1638 [1621], S. 411 (bis zu dem das Sprichwort in der Netzwerkforschung üblicherweise zurückverfolgt wird), der allerdings formuliert: „... and birds of a feather will gather together.“

<sup>380</sup> Newman, Networks, 2010, S. 222.

<sup>381</sup> Newman, Networks, 2010, S. 222. Zu den verschiedenen Skalen von Merkmalen bereits S. 93–99.

<sup>382</sup> Newman, Networks, 2010, S. 224.

<sup>383</sup> Entsprechend ist wieder die Wahl des Nullmodells problematisch. Zur Modularität bereits oben S. 155–156.

allerdings durch den Modularitätswert für ein maximal assortatives Netzwerk teilen.<sup>384</sup>

Bei metrischen Merkmalen kann man die *Kovarianz (covariance)* der Merkmalsausprägungen an beiden Endpunkten einer Beziehung messen.<sup>385</sup> Diese ist positiv (negativ), wenn höhere Werte auf der einen Seite insgesamt mit höheren (niedrigeren) Werten auf der anderen Seite einhergehen. Sie verhält sich also wie die „klassische“ empirische Kovarianz;<sup>386</sup> allerdings werden das arithmetische Mittel des betrachteten Merkmals am Endpunkt einer Beziehung und die Kovarianz selbst mithilfe der Adjazenzmatrix ermittelt. Teilt man die Kovarianz der betrachteten Merkmalsausprägungen durch den Wert, den sie in einem Netzwerk mit maximaler Assortativität annimmt, so erhält man ein Maß, das Werte im Intervall  $[-1, 1]$  hat, *Korrelationskoeffizient* genannt und mit  $r$  abgekürzt wird.<sup>387</sup> Dieser Korrelationskoeffizient gibt an, inwiefern die für die Endpunkte einer Beziehung betrachteten Merkmale in einem Netzwerk linear miteinander zusammenhängen, ohne etwas über nichtlineare Zusammenhänge oder Kausalitäten auszusagen.<sup>388</sup> Die Werte, die der Korrelationskoeffizient für verschiedene Netzwerke annimmt, lassen sich miteinander vergleichen.

Welche Korrelationskoeffizienten auf die geschilderte Weise konstruiert werden können, hängt auch davon ab, welche metrischen Merkmale für die Objekte des betrachteten Netzwerks erhoben wurden. Immer lassen sich allerdings Korrelationskoeffizienten für Merkmale bilden, die auf strukturellen Maßen der Mikroebene basieren, beispielsweise auf Grad, Innengrad und Außengrad. Für die Rechtsprechung in Flatland kann man dann etwa fragen: Zitieren Entscheidungen, die viel zitiert werden, tendenziell eher Entscheidungen, die ihrerseits viel zitiert werden?<sup>389</sup> Da man bei Zitierbeziehungen Start- und Endpunkt unterscheiden kann, lassen sich hier sogar vier Korrelationen für das Paar Startpunkt/Endpunkt untersuchen: (1) Innengrad/Innengrad (so die vorangegangene Formulierung), (2) Innengrad/Außengrad, (3) Außengrad/Innengrad und (4) Außengrad/Außengrad (Abb. 3.34 (a), S. 185).<sup>390</sup>

<sup>384</sup> Dazu Newman, Networks, 2010, S. 224–225.

<sup>385</sup> Allgemein ist die *empirische Kovarianz* zwischen zwei Merkmalen  $x_i$  und  $y_i$  für eine Menge von  $n$  Beobachtungen gegeben als  $\frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$ ; siehe etwa Fahrmeir u.a., Statistik, 2016, S. 128. Dabei steht  $\bar{x}$  bzw.  $\bar{y}$  („ $x$  quer“ bzw. „ $y$  quer“) für das arithmetische Mittel von  $x$  bzw.  $y$ , also für  $\frac{1}{n} \cdot \sum_{i=1}^n x_i$  bzw.  $\frac{1}{n} \cdot \sum_{i=1}^n y_i$ . Das arithmetische Mittel ist für nominale und ordinale Merkmale nicht definiert, daher kann man für sie auch keine Kovarianz berechnen.

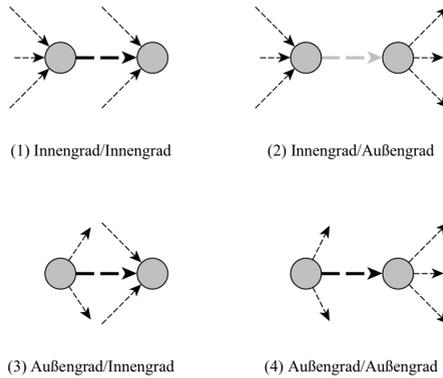
<sup>386</sup> Dazu bereits Fn. 385 (S. 184).

<sup>387</sup> Hierbei handelt es sich um eine auf Netzwerke zugeschnittene Variante des Bravais-Pearson-Korrelationskoeffizienten (daher auch die Abkürzung als  $r$ ); kleinschrittige Herleitung bei Newman, Networks, 2010, S. 228–229.

<sup>388</sup> Zu dieser Limitation von Korrelationskoeffizienten allgemein Fahrmeir u.a., Statistik, 2016, S. 129–131.

<sup>389</sup> Die Frage ähnelt der Konstellation, die manche Rückkopplungszentralitäten auf der Mikroebene als Wichtigkeit einzelner Entscheidungen erfassen wollen (oben S. 135–139).

<sup>390</sup> Foster u.a., PNAS 107 (2010), S. 10815 (10816). Diese Korrelationskoeffizienten sind nicht zu verwechseln mit der (ebenso konstruierbaren) Korrelation von Innen- und Außengrad



(a) Mögliche Maße

binär		Zitiert	
		Innengrad	Außengrad
Zitierend	Innengrad	0,044	-0,038
	Außengrad	-0,144	0,020

gewichtet		Zitiert	
		Innengrad	Außengrad
Zitierend	Innengrad	0,056	-0,035
	Außengrad	-0,085	0,020

(b) Werte des Ziternetzwerks in Flatland

Abbildung 3.34: Gradassortativitäten in Netzwerken mit asymmetrischen Beziehungen<sup>391</sup>

Dabei fällt auf, dass die Beziehung, deren Startpunkt und Endpunkt betrachtet werden, nur in Maß (2) *nicht* unmittelbar in die Ermittlung des Korrelationskoeffizienten einfließt – vorausgesetzt, das Netzwerk ist azyklisch. Da der Spruchkörper

*einzelner Objekte*: Dort wird zur Ermittlung der Korrelation immer nur ein Objekt betrachtet, hier werden zwei Objekte betrachtet. Nur deshalb sind auch vier verschiedene Maße möglich.

<sup>391</sup> Eine Abbildung 3.34 (a) (S. 185) ähnliche Darstellung findet sich auch bei *Foster u.a.*, PNAS 107 (2010), S. 10815 (10816). Die Werte in Abbildung 3.34 (b) (S. 185) liegen alle nahe Null, was bedeutet, dass keine starke lineare Korrelation zwischen den verschiedenen Graden der einzelnen Punkte im Zitiergraphen der flatländischen Rechtsprechung besteht. Das ist nicht verwunderlich, da in das Modell auch keine solche Korrelation eingebaut wurde (Einzelheiten im Online-Appendix). Zum Vergleich die in *Foster u.a.*, PNAS 107 (2010), S. 10815 (Supporting Information, S. 3) kommunizierten Werte der gerichteten Graphen für einige andere Informationsnetzwerke, die allerdings nicht zwingend azyklisch sind (Reihenfolge wie in Abb. 3.34 (a), S. 185): *World Wide Web* (-0,023; 0,257; -0,062; -0,014); *Wikipedia* (-0,014; 0,017; -0,070; -0,032); *Political Blogs* (-0,094; -0,023; -0,230; -0,0515).

der zitierenden Entscheidung zwar ihren Außengrad, nicht aber ihren Innengrad unter Kontrolle hat, kann bei der Interpretation von Messwerten teils auf Faktoren bei der zitierenden Entscheidung, teils auf Faktoren im Gerichtssystem abgestellt werden. Zuvor stellt sich allerdings wieder die Frage, ob die gemessenen Werte dem Zufall geschuldet sind. Man kann versuchen, das herauszufinden, indem man den gemessenen Wert mit den Werten von Modellnetzwerken vergleicht, die ähnliche Eigenschaften haben, wobei insbesondere die Konstruktion geeigneter Modellnetzwerke problematisch ist.<sup>392</sup>

Für die Unternehmen und Führungsfiguren in Flatland kann man beispielsweise fragen: Sind diejenigen Unternehmen, mit denen intensiv verflochtene Unternehmen personell verflochten sind, tendenziell mit Unternehmen verflochten, die ihrerseits intensiv verflochten sind? Sitzen Führungsfiguren, die mit vielen anderen Führungsfiguren in Gremien sitzen, tendenziell mit Führungsfiguren in Gremien, die ihrerseits in vielen verschiedenen Gremien sitzen? Diese Fragen sind unter anderem deshalb relevant, weil mögliche Antworten sehr unterschiedliche Netzwerktopologien implizieren. Netzwerke mit hoher Gradassortativität (*degree assortativity*) weisen oft eine zentralisierte Struktur auf, bei der wenige auch untereinander stark vernetzte Objekte in der Mitte mit vielen weniger stark vernetzten Objekten am Rand verbunden sind (Zentrum-Peripherie-Struktur, *core/periphery structure*).<sup>393</sup> In solchen Netzwerken breiten sich Informationen (oder andere Diffusionsgegenstände) ganz anders aus als in Netzwerken mit einer tendenziell dezentralisierten Struktur.<sup>394</sup>

---

<sup>392</sup> Zur Konstruktion von Modellnetzwerken allgemein S. 195–211. *Foster u.a.*, PNAS 107 (2010), S. 10815 (10816) testen die Signifikanz der (Dis-)Assortativität eines Netzwerks, indem sie die Differenz des gemessenen Werts zum Durchschnittswert eines randomisierten Graphenensembles bilden und durch die Standardabweichung der Werte dieses Ensembles teilen (zu Graphenensembles bei Fn. 429 [S. 197]); die Graphen im Ensemble weisen dieselbe Innen- und Außengradsequenz auf wie der Graph des untersuchten Netzwerks (eine *Gradsequenz* ist eine Folge von Zahlen, die den Grad für eine Folge von Punkten in einem Graphen angibt). Unter der Annahme, dass die Werte von  $r$  im randomisierten Graphenensemble normalverteilt sind, erhalten sie so eine Messstatistik, die standardnormalverteilt ist und angibt, wie viele Standardabweichungen der beobachtete Wert über oder unter dem Durchschnittswert des Graphenensembles liegt. Die verwendete Statistik wird *Z-score* genannt und entspricht dem, was in der allgemeinen Statistik unter diesem Ausdruck bekannt ist (siehe etwa *Fahrmeir u.a.*, Statistik, 2016, S. 85–86) – der Graph des beobachteten Netzwerks wird als Stichprobe aus einer Population von Graphen betrachtet, die durch das Graphenensemble repräsentiert wird. Problematisch ist dabei neben der Wahl des Nullmodells (Graph mit identischer Innen- und Außengradsequenz) auch die Normalverteilungsannahme (deren Plausibilität sich unter anderem durch explorative Visualisierung der für einzelne Modellinstanzen erzielten Werte abschätzen lässt).

<sup>393</sup> *Newman*, Networks, 2010, S. 230–231.

<sup>394</sup> Zu Diffusionsprozessen auf Netzwerken noch unten S. 216–223.

### V. Zusatzfaktor: Zeit

Die in den vorigen Abschnitten eingeführten Methoden sind darauf zugeschnitten, Netzwerke in ihrem Zustand zu einem bestimmten Zeitpunkt zu untersuchen. Die Analyse ist *statisch*; das betrachtete Netzwerk ist eine *Momentaufnahme* (*snapshot*). Viele juristische Netzwerke verändern sich aber mit der Zeit: Ein Rechtsprechungsziernetzwerk wächst mit jeder neuen Gerichtsentscheidung, ein Verflechtungsnetzwerk wandelt sich mit jeder personellen Veränderung in den betrachteten Gremien.<sup>395</sup> Daher können sich auch die zu verschiedenen Zeitpunkten entstandenen Momentaufnahmen eines juristischen Netzwerks unterscheiden – und daraus, wie und worin sie dies tun, lässt sich wiederum viel über das Netzwerk lernen. Erforscht man die Entwicklung eines Netzwerks über die Zeit (*Netzwerk-evolution, network evolution*), so ist die Analyse *dynamisch*. Dazu kann man das Netzwerk permanent beobachten und alle Messungen laufend aktualisieren – dann ist die Zeit eine stetige Variable (Zeit als Kontinuum). Oder aber man macht mehrere, zeitlich versetzte Momentaufnahmen vom Netzwerk und vergleicht diese miteinander – dann wird die Zeit als diskrete Variable behandelt (Zeit als Folge von Zeitpunkten) und die Momentaufnahmen bilden eine *Zeitreihe* (*time series*).<sup>396</sup>

Da die Frequenz von Momentaufnahmen (im Rahmen des technisch Möglichen) beliebig erhöht werden kann, ist der Übergang fließend. Die Behandlung von Zeit als Kontinuum ist allerdings nur dann sinnvoll, wenn auch kontinuierlich Veränderungen auftreten können. Das ist etwa bei einem Netzwerk von Banktransaktionen (Konten als Objekte und Zahlungsströme als Beziehungen) oder bei einem Infrastrukturnetzwerk der Fall, auf dem ein Prozess abläuft, der sich in Echtzeit beobachten lässt (z.B. bei einem Straßennetz mit Kraftfahrzeugverkehr oder einem Schienennetz mit Bahnverkehr). Bei einem Zitiernetzwerk aus Gerichtsentscheidungen, dessen Objekte (Entscheidungen) nur mit dem Tag der Entscheidung gekennzeichnet sind, reicht es hingegen aus, für jeden Tag eine Momentaufnahme zu machen – zu welcher Uhrzeit eine Entscheidung öffentlich wird, hängt von der Kommunikationspraxis der Gerichte ab und ist kein Umstand, der sich sinnvoll zur Einordnung einzelner Entscheidungen auf einem Zeitstrahl heranziehen lässt. Ebenso verhält es sich mit einem Netzwerk aus Personalverflechtungen zwischen Unternehmen: Unabhängig davon, ob man für die Registrierung personeller Veränderungen auf Bekanntgabe oder Inkrafttreten der Neuerung abstellt, lässt sich auf Basis von Momentaufnahmen im Tagesrhythmus ein (für die Untersuchungszwecke) vollständiges Bild von der Netzwerkevolution zeichnen. Auch zur dynamischen Analyse anderer juristischer Netzwerke genügt meist die Untersuchung von Zeitreihen aus ihren Momentaufnahmen. Um längerfristige Trends in

<sup>395</sup> Das setzt natürlich eine entsprechend flexible, zeitlich „relative“ Netzwerkdefinition voraus, etwa „alle Entscheidungen des höchsten Gerichts, die bis zum letzten Tag des vergangenen Monats getroffen wurden (Objekte), und ihre Zitate untereinander (Beziehungen)“.

<sup>396</sup> Einführend zur Analyse von Zeitreihen *Fahrmeir u.a.*, Statistik, 2016, S. 503–524.

der Entwicklung juristischer Netzwerke zu erkennen, können unter Umständen sogar Momentaufnahmen ausreichen, die Monate oder Jahre auseinanderliegen. Sind allerdings Momentaufnahmen in kürzeren Zeitabständen verfügbar, so bietet es sich an, auf die statistischen Methoden der Zeitreihenanalyse zurückzugreifen und für einzelne Merkmale des Netzwerks beispielsweise *gleitende Durchschnitte* (*moving averages*) zu bilden, um möglichst viele Informationen in die Trenderkennung einzubeziehen.<sup>397</sup> Dies ist insbesondere für Rechtsprechungsziernetzwerke relevant, deren tägliche Momentaufnahmen sich anhand der Entscheidungsdaten rekonstruieren lassen.

Doch was heißt es eigentlich, die Evolution eines juristischen Netzwerks zu beobachten, und welche Trends sind dabei von Interesse? Grundsätzlich kann jedes Konzept, das zur statischen Analyse eines Netzwerks geeignet ist (z.B. Gradzentralitäten auf der Mikroebene, Gemeinschaften auf der Mesoebene oder Clustering-Koeffizienten auf der Makroebene), auch zu seiner dynamischen Analyse eingesetzt werden. Sind die Momentaufnahmen eines Netzwerks als Zeitreihe gegeben, so lassen sich daraus Zeitreihen von Messwerten für alle Maße ableiten, die in den vorigen Abschnitten eingeführt wurden, indem für jede Momentaufnahme des Netzwerks eine entsprechende statische Analyse durchgeführt wird. Welche Maße für die Beschreibung der Netzwerkevolution relevant sind, hängt insbesondere davon ab, in welcher Hinsicht sich das betrachtete Netzwerk mit der Zeit überhaupt verändern kann. Prinzipiell können Objekte und Beziehungen als Elemente des Netzwerks hinzukommen, wegfallen oder ausgetauscht werden.<sup>398</sup> So ergeben sich sechs Veränderungsprozesse auf der Makroebene, die in manchen Netzwerken auch (oder sogar nur) kombiniert auftreten können (Tab. 3.9, S. 188).

	<i>Objekte</i>	<i>Beziehungen</i>
<i>mehr</i>	Wachstum	Verdichtung
<i>weniger</i>	Schrumpfung	Ausdünnung
<i>andere</i>	Fluktuation	Umstrukturierung

Tabelle 3.9: Veränderung von Netzwerken

<sup>397</sup> Der gleitende Durchschnitt eines Merkmals  $Y$  zum Zeitpunkt  $t$  bezieht neben dem Wert von  $Y$  zu diesem Zeitpunkt ( $y_t$ ) auch die Werte zu jeweils  $k$  Zeitpunkten vor und nach  $t$  mit ein. Er kann auf verschiedene Weisen gebildet werden; die einfachste Möglichkeit besteht darin, das arithmetische Mittel der Werte innerhalb der  $2k + 1$  Zeitpunkte zu bilden, das als *lokales arithmetisches Mittel* bezeichnet wird:  $y_t^* = \frac{1}{2k+1} \cdot \sum_{i=-k}^k y_{t+i}$ . Der Durchschnitt wird gleitend genannt, weil sich durch die Berücksichtigung von  $k$  Zeitpunkten vor und nach  $t$  ein Fenster der Weite  $2k + 1$  ergibt, das sich mit  $t$  über die Zeitachse bewegt (also „gleitet“). Zum Ganzen *Toutenburg und Heumann, Deskriptive Statistik*, 2009, S. 2013–214; *Fahrmeir u.a., Statistik*, 2016, S. 516–517.

<sup>398</sup> Der Austausch eines Netzwerkelements impliziert, dass ein Element hinzukommt und ein Element wegfällt. Es lohnt sich dennoch, den Austausch als separate Kategorie zu betrachten, da hier Hinzukommen und Wegfall untrennbar miteinander verknüpft sind.

Im Verflechtungsnetzwerk der flatländischen Unternehmenslandschaft können alle sechs Veränderungsprozesse unabhängig voneinander auftreten: Neue Unternehmen können gegründet, bestehende liquidiert werden (Wachstum und Schrumpfung), einzelne Führungsfiguren können andere ersetzen (Fluktuation), Führungsfiguren können von einem Kontrollgremium in ein anderes wechseln (Umstrukturierung), mehr oder weniger Kontrolleure in einzelnen Gremien eingesetzt werden (Verdichtung und Ausdünnung). Demgegenüber scheiden Schrumpfung, Fluktuation und Umstrukturierung als Veränderungsprozesse für das Ziternetzwerk der flatländischen Rechtsprechung von vornherein aus, während Wachstum ein beinahe zwingendes Resultat des Gerichtssystems ist. Dabei kommt es fast zwangsläufig zu Ausdünnung anstatt zu Verdichtung, selbst wenn die hinzukommenden Entscheidungen mehr andere Entscheidungen zitieren als der bisherige Durchschnitt. Denn die Anzahl möglicher Zitate wächst mit der Anzahl der Punkte im Netzwerk quadratisch – und ohne Wachstum ist eine Veränderung der Beziehungsdichte für ein Rechtsprechungsziternetzwerk nicht möglich. Die dargestellten Einschränkungen der Veränderungsmöglichkeiten gelten für Rechtsprechungsziternetzwerke im Allgemeinen ebenso wie für Ziternetzwerke aus anderen Dokumenten, die nach ihrer Veröffentlichung nicht mehr modifiziert werden können. Anders verhält es sich insbesondere beim *World Wide Web*, in dem wiederum alle sechs Veränderungsprozesse auftreten können. Diese verschiedenen Möglichkeiten der Netzwerkevolution sind bei der Übertragung von Methoden, die für den *Web Graph* entwickelt wurden, auf juristische Ziternetzwerke gegebenenfalls zu berücksichtigen.

Ähnlich wie das Netzwerk als Ganzes können sich auch seine Strukturen auf der Mesoebene mit der Zeit verändern. Aus rechtswissenschaftlicher Perspektive von besonderem Interesse ist dabei, wie sich die Gemeinschaftsstrukturen juristischer Ziternetzwerke, die von bestimmten Algorithmen identifiziert werden, über längere Zeiträume hinweg entwickeln.<sup>399</sup> Im Rahmen der Evolution eines Netzwerks können grundsätzlich neue Gemeinschaften entstehen (Geburt) oder bestehende Gemeinschaften verschwinden (Tod), bestehende Gemeinschaften können Mitglieder gewinnen (Vergrößerung) oder verlieren (Verkleinerung), mehrere Gemeinschaften können zu einer Gemeinschaft zusammenwachsen (Vereinigung), oder eine Gemeinschaft kann in mehrere Gemeinschaften zerfallen (Teilung).<sup>400</sup> Geht man davon aus, dass sich die in Rechtsprechungsziternetzwerken entdeckten Gemeinschaften primär an den Entscheidungsinhalten orientieren (was aufgrund der vielfältigen anderen Einflussfaktoren, allen voran der Person des Richters, nicht unzweifelhaft ist), so können die beschriebenen Veränderungen als bestimmte Formen der Rechtsentwicklung interpretiert werden: Die Geburt oder die

<sup>399</sup> Dass es sich auch lohnen kann, die Veränderung der Gemeinschaftsstrukturen in den Netzwerken juristischer Akteure zu untersuchen, zeigen *Mucha u.a.*, *Science* 328 (2010), S. 876 (877–878).

<sup>400</sup> Dazu im Kontext sozialer Netzwerke *Palla, Barabási und Vicsek*, *Nature Letters* 446 (2007), S. 664 (664); zusammenfassend *Barabási*, *Network Science*, 2016, S. 359–360.

Vergößerung einer Gemeinschaft (soweit sie eine kritische Größe überschreitet) spricht für die Entstehung eines juristischen Problembereichs in dem Sinne, dass offenbar neue Inhalte justiziabel werden. Die Verkleinerung einer Gemeinschaft (soweit sie eine kritische Größe unterschreitet) deutet hingegen auf das Aussterben eines juristischen Problembereichs hin, denn die bisher gemeinsam diskutierten Inhalte scheinen vermehrt in unterschiedlichen Kontexten behandelt zu werden.<sup>401</sup> Teilung von Gemeinschaften könnte als Ausdifferenzierung von Rechtsprechung, Vereinigung von Gemeinschaften als ihre Entdifferenzierung aufgefasst werden.

Im Rahmen der dynamischen Analyse auf der Mesoebene kann man auch nach Bedeutungsverschiebungen in der Rechtsprechung suchen. Dazu analysiert man, welche Entscheidungen zu welchen Zeitpunkten in derselben Gemeinschaft zusammengefasst werden und wie sich diese Zusammenfassungen mit der Zeit verändern. Werden Entscheidungen, die bisher verschiedenen Gemeinschaften angehörten, erstmals zu einer Gemeinschaft zusammengefasst, so steht dahinter wahrscheinlich eine zunehmende Co-Zitation (die Entscheidungen werden von neuen Entscheidungen öfter gemeinsam zitiert), die dadurch bedingt sein kann, dass bisher getrennt voneinander diskutierte Konzepte nunmehr gemeinsam zur Lösung von Rechtsfragen herangezogen werden. Bisher ist allerdings noch nicht untersucht worden, inwieweit reale Zitierdaten der Rechtsprechung solche Deutungen ermöglichen.<sup>402</sup>

Sind potenzielle Bedeutungsverschiebungen identifiziert, so lässt sich die Co-Zitation beteiligter Entscheidungen (als Beziehung zweiten Grades) auch direkt auf der Mikroebene dynamisch untersuchen. Je nachdem, welchen Mustern die Gerichte bei der Zitierung folgen und wie die Zitierdaten extrahiert wurden, kann dabei zwischen Co-Zitation auf der Ebene einzelner Dokumente und Co-Zitation auf der Ebene einzelner Zitierblöcke unterschieden werden.<sup>403</sup> Die Inhalte von Entscheidungen, die vermehrt in einzelnen Zitierblöcken auftauchen, werden von den zitierenden Entscheidungen tendenziell stärker miteinander verknüpft als die Inhalte von Entscheidungen, die nur verstärkt jeweils irgendwo in einzelnen Entscheidungen zitiert werden (Abb. 3.35, S. 191).<sup>404</sup>

<sup>401</sup> Der Tod einer Gemeinschaft ist in einem Rechtsprechungszitienetzwerk unmöglich, da dieses nicht schrumpfen kann – das „Verschwinden“ einer Gemeinschaft von einer Momentaufnahme zur anderen muss daher Folge von Verkleinerung, Vereinigung oder Teilung sein.

<sup>402</sup> Mit statischen Gemeinschaftsstrukturen in Rechtsprechungszitienetzwerken befassen sich etwa *Derlén u.a.*, *Europarättslig Tidskrift* 2013, S. 517–535, sowie *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327–349 (341–344); eine ausführliche statische Analyse unternimmt *Schaper*, *The Structure and Organization of EU Law in the Field of Direct Taxes*, 2013, S. 155–226.

<sup>403</sup> Zur Erinnerung: Zitierblöcke sind Sequenzen von Zitaten, die in einem Dokument unmittelbar hintereinander stehen (in deutschen Gerichtsentscheidungen typischerweise in einem Klammerzusatz); dazu schon oben S. 76–79.

<sup>404</sup> Bei der dynamischen Analyse von Zitierblöcken geht es im Grunde darum, die Entwicklung „struktureller Kollokationen“ nachzuverfolgen – ähnlich, wie in der Linguistik die Entwicklung von Konzepten auf Basis von *n-grams* (Sequenzen von Textfragmenten) untersucht wird, indem

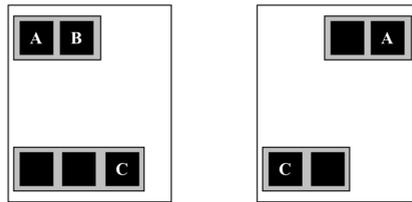


Abbildung 3.35: Co-Zitation auf der Ebene einzelner Zitierblöcke und auf der Ebene einzelner Dokumente<sup>405</sup>

Eine weitere Möglichkeit, ein Netzwerk auf der Mikroebene dynamisch zu analysieren, besteht darin, die nach den verschiedenen Zentralitäten zu unterschiedlichen Zeitpunkten jeweils „wichtigsten“ Entscheidungen miteinander zu vergleichen und den Weg einzelner Entscheidungen hin zu oder weg von ihrer Spitzenposition nachzuzeichnen. Speziell bei zeitlich eingebetteten juristischen Ziternetzwerken, in denen sich Zitate nur (oder fast nur) in die Vergangenheit richten, kann man auch direkt Gradzentralitäten definieren, welche die Zeitabhängigkeit berücksichtigen. Innengrad und Außengrad in ihrer klassischen Formulierung haben für solche Netzwerke den Nachteil, dass ältere Entscheidungen sehr viel häufiger zitiert werden können als jüngere Entscheidungen und jüngere Entscheidungen mehr Entscheidungen zitieren können als ältere Entscheidungen. Damit dominieren beim Innengrad in der Regel Entscheidungen, die zum Betrachtungszeitpunkt relativ alt sind, und beim Außengrad dominieren Entscheidungen, die

---

man erkundet, welche Wörter aus statistischer Sicht besonders häufig gemeinsam (d.h. in einem bestimmten Abstand voneinander) in den einzelnen Dokumenten eines Korpus auftauchen. Die Analyse von *n-grams* wird im Rahmen der Untersuchung deutscher Rechtsprechung zum Arbeitnehmerbegriff eingesetzt von *Vogel, Pötters und Christensen*, Richterrecht der Arbeit – empirisch untersucht, 2015, S. 100–126; zur Untersuchung des Konzepts der Effektivität (*effectiveness*) wird sie mit netzwerkanalytischen Methoden kombiniert von *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327 (335, 339–341).

<sup>405</sup> Betrachtet man die Co-Zitation der Dokumente A, B und C (schwarze Kästen mit Buchstaben) auf der Ebene einzelner Dokumente (weiße Kästen), so sind die Dokumente A und C hier stärker miteinander verknüpft als die Dokumente A und B; außerdem sind die Dokumente B und C ebenso stark miteinander verknüpft wie die Dokumente B und A. Betrachtet man die Co-Zitation hingegen auf der Ebene von Zitierblöcken (graue Kästen), so scheinen die Dokumente A und B stärker miteinander verknüpft als die Dokumente A und C, die nicht enger verbunden scheinen als die Dokumente B und C. Auf welche Art der Co-Zitation in welchem Maße abzustellen ist (beide Nuancen lassen sich auch kombinieren), hängt vom konkreten Erkenntnisinteresse ab. Denkbar ist auch, bei beiden Maßen danach zu differenzieren, wie viele andere Dokumente (neben den konkret betrachteten) in demselben Dokument bzw. in demselben Zitierblock wie häufig (und in wie vielen anderen Dokumenten) zitiert werden. Eine Einbeziehung der Co-Zitation auf der Ebene von Zitierblöcken setzt allerdings voraus, dass die Daten entsprechend differenziert erhoben werden.

zum Betrachtungszeitpunkt relativ jung sind.<sup>406</sup> Um die ungleichen Startvoraussetzungen bei Innengrad und Außengrad zu beheben, kann man definieren, dass eine Entscheidung im Zitiernetzwerk zu einem bestimmten Zeitpunkt umso wichtiger ist, je größer der auf sie entfallende Anteil der nach ihrer Veröffentlichung eingehenden Zitate (*temporaler Innengrad*, *temporal indegree*) oder der auf sie entfallende Anteil der bis zu ihrer Veröffentlichung ausgehenden Zitate (*temporaler Außengrad*, *temporal outdegree*) ist. Temporaler Innengrad und temporaler Außengrad einer Gerichtsentscheidung zu einem bestimmten Zeitpunkt lassen sich berechnen, indem man den Innen- bzw. Außengrad des ihr zugeordneten Punktes durch die jeweils einschlägige Anzahl von (eingehenden bzw. ausgehenden) Pfeilen teilt.<sup>407</sup> Damit sind die beschriebenen Maße für Graphen mit Mehrfachkanten (gewichtete Zählung der Zitate) ebenso geeignet wie für Graphen ohne Mehrfachkanten (binäre Zählung der Zitate).<sup>408</sup>

Der temporale Außengrad einer Entscheidung steht (wie der herkömmliche Außengrad) mit ihrer Veröffentlichung fest; die Berücksichtigung der Zeit stellt hier lediglich sicher, dass die Werte von Entscheidungen, die zu verschiedenen Zeitpunkten getroffen wurden, miteinander vergleichbar sind. Demgegenüber kann sich der temporale Innengrad von Gerichtsentscheidungen (wie der herkömmliche Innengrad) mit der Zeit verändern. Beobachtet man diese Veränderungen über einen längeren Zeitraum, so erhält man juristische *Zitierkurven* (*citation curves*), die den Rezeptionszyklus einzelner Entscheidungen beschreiben.<sup>409</sup> Dann

<sup>406</sup> Zur Dominanz älterer Entscheidungen bei der Beurteilung ihrer Wichtigkeit anhand des Innengrads siehe bereits Abb. 3.7 (S. 108) und den begleitenden Text.

<sup>407</sup> Problematisch ist die Behandlung der Fälle, in denen nach der Definition der Nenner Null sein müsste, denn die Division durch Null ist nicht definiert. Dies betrifft für den Außengrad die ältesten Entscheidungen, für den Innengrad die zum Betrachtungszeitpunkt jüngsten Entscheidungen im Netzwerk. Setzt man den temporalen Außen- bzw. Innengrad in diesen Fällen auf Null, so schließt man die betroffenen Entscheidungen effektiv von der Analyse der Zentralitäten aus. Da es sich um wenige Entscheidungen handelt, die man ohne Weiteres identifizieren kann und deren Wichtigkeit im jeweiligen Sinne sich aufgrund ihrer Sonderposition ohnehin nicht sinnvoll beurteilen lässt, scheint dies vertretbar.

<sup>408</sup> Ein dem temporalen Innengrad ähnliches Maß wird von *Tarissan und Nollez-Goldbach*, Temporal Properties of Legal Decision Networks: A Case Study from the International Criminal Court, in: *Rotolo* (Hrsg.), Legal Knowledge and Information Systems, 2015, S. 111–120 und *Tarissan, Panagis und Sadl*, IEEE/ACM ASONAM 2016, S. 661 (662) vorgeschlagen und als *relativer Innengrad* (*relative indegree*, dort *relative in-degree*) bezeichnet. Die von den Autoren gewählte Definition hat allerdings den Nachteil, dass sie sich in Graphen mit Mehrfachkanten nicht sinnvoll anwenden lässt (der Innengrad eines Punktes  $v$  wird durch die Anzahl der nach  $v$  zum Graphen hinzu gekommenen Punkte geteilt). Die Bezeichnung als relativer Innengrad ist außerdem unglücklich, da dieser Ausdruck oft für das Maß verwendet wird, das man erhält, wenn man den Innengrad durch die Anzahl *aller* Punkte im Netzwerk teilt (was unter anderem dann nützlich ist, wenn man die Innengrade einzelner Punkte in Netzwerken vergleichen will, die unterschiedlich groß sind).

<sup>409</sup> Auf realen Daten basierende Beispiele für juristische Zitierkurven finden sich in Abbildung 4.20 (S. 276).

kann man auch die Zitierkurven verschiedener Gerichtsentscheidungen miteinander vergleichen und versuchen, charakteristische Muster zu identifizieren. Sind Entscheidungen zu prozessrechtlichen Fragen vielleicht generell langlebiger als solche zu materiellrechtlichen Fragen? Oder gibt es bestimmte Themen, deren Konjunktur sich an den Zitierkurven einschlägiger Entscheidungen ablesen lässt? Diese und ähnliche Fragen sind noch weitgehend ungeklärt.

## D. Modellierung

Neben Visualisierung und Quantifizierung ist Modellierung ein wesentliches Element der Methodik der Netzwerkforschung. Modellierung bezeichnet das indirekte Studium realer Systeme durch die Konstruktion und Analyse von Modellen.<sup>410</sup> Modelle sind *interpretierte Strukturen* (*interpreted structures*), d.h. Verbindungen einer Struktur mit einer Deutung dieser Struktur.<sup>411</sup> Die bedeutungstragenden Strukturen können verschiedene Formen annehmen, insbesondere können sie als physische Objekte (*physische Modelle*), mathematische Abstraktionen (*mathematische Modelle*) oder Mengen von Algorithmen (*algorithmische Modelle*) auftreten.<sup>412</sup> Dabei kann eine Struktur mit verschiedenen Interpretationen und eine Interpretation mit verschiedenen Strukturen verbunden werden, die zu jeweils unterschiedlichen Modellen führen – das Verhältnis von Strukturen zu Interpretationen ist also *viele-zu-viele* (*n:m, many-to-many*).<sup>413</sup>

Physische Modelle sind zum Beispiel die aus der Architektur bekannten Miniaturen, die eingesetzt werden, um die Eigenschaften geplanter Neubauten zu studieren; sie spielen in der Netzwerkforschung kaum eine Rolle und werden hier nicht weiter behandelt. Mathematische Modelle enthalten mathematische Strukturen (z.B. Mengen und Funktionen), welche die Zustände von Systemen und ihre Beziehungen zueinander (insbesondere die Übergänge zwischen einzelnen Zuständen) darstellen können.<sup>414</sup> Ermöglicht ein mathematisches Modell die Beschreibung von Zustandsübergängen, so ist es *dynamisch*, beschränkt sich das Modell auf die Zustandsbeschreibung, so ist es *statisch*.<sup>415</sup> Die Graphen, die in dieser Arbeit zur Untersuchung von Netzwerken zu einem bestimmten Zeitpunkt

<sup>410</sup> Weisberg, *Simulation and Similarity*, 2013, S. 4. In der Literatur finden sich verschiedene Definitionen der Begriffe „Modellierung“ und „Modell“.

<sup>411</sup> Weisberg, *Simulation and Similarity*, 2013, S. 15; für die Netzwerkforschung aufgegriffen von Zweig, *Network Analysis Literacy*, 2016, S. 366.

<sup>412</sup> Die Differenzierung orientiert sich an Weisberg, *Simulation and Similarity*, 2013, S. 7, der von *concrete models*, *mathematical models* und *computational models* spricht.

<sup>413</sup> Weisberg, *Simulation and Similarity*, 2013, S. 38.

<sup>414</sup> Weisberg, *Simulation and Similarity*, 2013, S. 29.

<sup>415</sup> Ähnlich Weisberg, *Simulation and Similarity*, 2013, S. 29, der allerdings von einem „non-dynamical model“ spricht, das zur Untersuchung von „static, structural features of a system“ eingesetzt wird.

eingesetzt werden, sind statische mathematische Modelle.<sup>416</sup> Dynamische mathematische Modelle werden häufig in Form von Differenzialgleichungen angegeben, welche die Veränderungsrate einer interessierenden Größe (z.B. die Veränderung des Bruttoinlandsprodukts) als Funktion ihres aktuellen Werts (z.B. des aktuellen Bruttoinlandsprodukts) beschreiben.<sup>417</sup>

Dynamische mathematische Modelle haben den Nachteil, dass sie Kausalität, Zufall und parallel ablaufende Prozesse nicht oder nur schwer abbilden können.<sup>418</sup> In der Netzwerkforschung spielen daher neben statischen mathematischen Modellen vor allem algorithmische Modelle eine Rolle, in denen nicht (wie bei dynamischen mathematischen Modellen) die Eingaben und Ausgaben von Prozeduren, sondern die Prozeduren selbst als interpretierte Strukturen im Vordergrund stehen.<sup>419</sup> Bekannte Beispiele für algorithmische Modelle sind agentenbasierte Modelle (*agent-based models*) für Segregation (*segregation*) und Schwarmverhalten (*flocking*).<sup>420</sup> In diesen Modellen ergibt sich der globale Systemzustand zu einem bestimmten Zeitpunkt  $t_n$  aus dem Ausgangszustand  $t_0$  und einer Menge nachfolgender, zwischen  $t_0$  und  $t_n$  (oft parallel) getroffener Entscheidungen einzelner Agenten (*agents*), z.B. Menschen. Die Entscheidungen basieren auf Regeln, die insbesondere Informationen aus der Umgebung der Agenten einbeziehen. Für die Netzwerkforschung sind diese Modelle interessant, wenn die Agenten als Objekte in einem Netzwerk dargestellt werden und ihre Entscheidungen von den Beziehungen abhängen, an denen sie beteiligt sind.<sup>421</sup>

Bisher sind in dieser Arbeit vor allem Graphen als statische mathematische Modelle von Netzwerken aufgetaucht, wobei die Netzwerke als Abstraktionen von verschiedenen juristisch interessanten Phänomenen in der Realität fungierten. Die Erläuterungen haben sich an der Durchführung konkreter netzwerkanalytischer Studien im Recht orientiert, weshalb konkrete (wenngleich fiktive) Netzwerkdaten

<sup>416</sup> Siehe auch Weisberg, *Simulation and Similarity*, 2013, S. 29.

<sup>417</sup> Weisberg, *Simulation and Similarity*, 2013, S. 26.

<sup>418</sup> Weisberg, *Simulation and Similarity*, 2013, S. 30–31.

<sup>419</sup> Weisberg, *Simulation and Similarity*, 2013, S. 31. Da jeder Algorithmus auf mathematische Operationen zurückgeführt werden kann und Zustandsübergänge beschreibt, sind algorithmische Modelle stets mathematisch und dynamisch, sodass sie sich auch als Teilmenge der dynamischen mathematischen Modelle auffassen lassen. Die Schwerpunkte mathematischer und algorithmischer Modelle sind aber so unterschiedlich, dass eine getrennte Behandlung sinnvoll ist, zumal algorithmische Modelle oft nicht übersichtlich mathematisch dargestellt werden können; ebenso Weisberg, *Simulation and Similarity*, 2013, S. 20, 30.

<sup>420</sup> Allgemein zu agentenbasierten Modellen Epstein, *Complexity* 4 (1999), S. 41–60; Bona-beau, *PNAS* 99 (suppl. 3) (2002), S. 7280–7287; Helbing (Hrsg.), *Social Self-Organization*, 2012, S. 25–70. Zu agentenbasierten Segregationsmodellen früh beispielsweise Schelling, *The American Economic Review* 59 (1969), S. 488–493; zu Modellen für Schwarmverhalten Vicsek u.a., *Physical Review Letters* 75 (1995), S. 1226–1229.

<sup>421</sup> Siehe etwa Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 85–91, 195–197; Jackson, *Social and Economic Networks*, 2008, S. 406–407; Kruschke und Röpcke, *Graphen und Netzwerktheorie*, 2015, S. 233–235.

als Graphen modelliert wurden.<sup>422</sup> Modellierung als methodischer Baustein der Netzwerkforschung hat zwar auch Interpretationsszenarien vor Augen, setzt aber abstrakter an und greift dabei auf mathematische und algorithmische Modelle zurück. Nach ihren primären Gegenständen lassen sich dabei zwei Klassen von Modellen unterscheiden: Modelle für Strukturen von Netzwerken (I.) und Modelle für Prozesse in und auf Netzwerken (II.).

### I. Strukturen

Viele Quantifizierungsmethoden der juristischen Netzwerkforschung ermöglichen Aussagen über die Struktur der untersuchten Netzwerke. So kann etwa festgestellt werden, dass ein Netzwerk Gemeinschaftsstrukturen aufweist (S. 157), dass die Distanzen zwischen je zwei Objekten im Durchschnitt sehr kurz sind (S. 168) oder dass die Gradverteilung sehr schief ist (S. 177). Dabei bleibt allerdings offen, wie die einzelnen Charakteristika von Netzwerken miteinander zusammenhängen und wie Netzwerke mit bestimmten Charakteristika zustande kommen. Die Modellierung von Netzwerkstrukturen ermöglicht eine Annäherung an diese Fragen.<sup>423</sup>

Im Rahmen der Modellierung von Netzwerkstrukturen werden Netzwerke mit bestimmten, vorgegebenen Eigenschaften in ihrer Darstellung als Graphen konstruiert und analysiert. Zur Erstellung der Graphen müssen Punkte und Verbindungen (Linien oder Pfeile) definiert werden. Dabei sind zwei Herangehensweisen denkbar, die auch kombiniert werden können: Erstens kann man mit einer nichtleeren Menge von Punkten beginnen und diese durch Linien oder Pfeile untereinander verdrahten, ohne dabei Punkte hinzuzufügen (*Verdrahtungsmodelle*, 1.). Zweitens kann man mit einer (möglicherweise leeren) Punktmenge beginnen und den Graphen schrittweise wachsen lassen, indem man einzelne Punkte gemeinsam mit von ihnen ausgehenden Verbindungen hinzufügt (*Wachstumsmodelle*, 2.).

Nachfolgend werden einige Verdrahtungs- und Wachstumsmodelle vorgestellt, die in der Netzwerkforschung bisher intensiv untersucht worden sind. Ihnen ist in ihren Grundformen gemein, dass sie eine geringe Anzahl von Parametern haben – es gibt also nur wenige Stellschrauben, an denen man drehen kann, um die Modelle einzustellen. Das macht es leichter, diese Modelle mathematisch zu analysieren und den Ablauf von Prozessen auf ihnen zu simulieren, führt aber dazu, dass sie realen Netzwerken (bzw. deren Graphen) oft nur hinsichtlich ausgewählter Eigenschaften entsprechen können. Dies wird im Folgenden laufend anhand

---

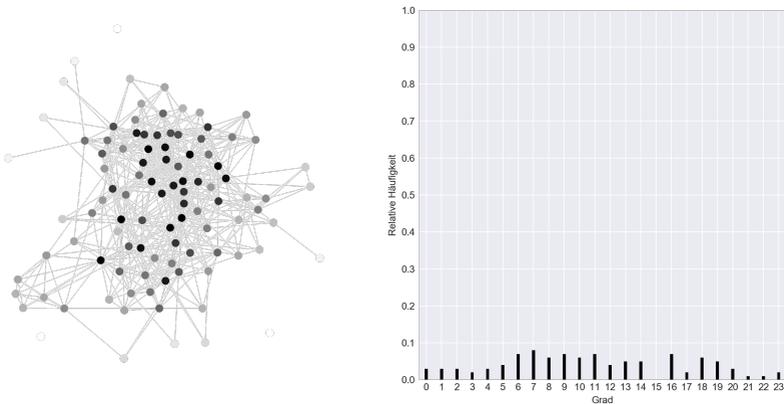
<sup>422</sup> Inwieweit die in Abschnitt A.II. (S. 88–93) eingeführten Graphen auch Modelle im Sinne von Weisberg, *Simulation and Similarity*, 2013, sind, ist nicht ganz klar, da dieser zwischen einem Modell für ein Zielsystem (einer abstrakten Darstellung des interessierenden Phänomens) und der mathematischen Repräsentation eines Zielsystems unterscheidet (a.a.O. S. 95–96). Diese (in anderen Kontexten unter Umständen sinnvolle) Differenzierung wird hier nicht aufgegriffen, da sie die Darstellung für die hiesigen Zwecke komplizierter machen würde, ohne wesentlich zum Erkenntnisgewinn beizutragen.

<sup>423</sup> Allgemein zur Modellierung von Netzwerkstrukturen Newman, *SIAM Review* 45 (2003), S. 167 (196–224).

von Vergleichen mit dem Zitiergraphen (Gerichtsentscheidungen als Punkte und Zitate als Pfeile) und dem Unternehmensgraphen (Unternehmen als Punkte und gemeinsame Führungsfiguren als Linien) aus Flatland illustriert.<sup>424</sup>

### 1. Verdrahtungsmodelle

Um die Unterschiede zwischen den verschiedenen Verdrahtungsmodellen zu verstehen, ist es nützlich, die Eigenschaften der Graphen, die sie produzieren, mit den Eigenschaften der aus den vorangegangenen Abschnitten bereits bekannten Graphen zu vergleichen, die auf fiktiven Daten basieren. Zunächst sei dazu der Unternehmensgraph noch einmal betrachtet, den man erhält, wenn die einzelnen Unternehmen als Punkte dargestellt und die Personalverflechtungen zwischen ihnen als Linien repräsentiert werden. Lässt man keine Mehrfachkanten zu (diese entstehen, wenn zwei Unternehmen mehrere Führungsfiguren gemeinsam haben), so hat der Unternehmensgraph 100 Punkte und 537 Linien; der Durchschnittsgrad beträgt 10,74. Der Clustering-Koeffizient liegt mit 0,41 relativ hoch, die durchschnittliche Länge der kürzesten Wege zwischen je zwei Punkten ist mit 2,33 verhältnismäßig gering. Weitere Strukturinformationen zum Graphen lassen sich seiner Visualisierung im kräftebasierten Layout und der Histogrammdarstellung seiner Gradverteilung entnehmen (Abb. 3.36, S. 196).



(a) Unternehmensgraph  
(ohne Mehrfachkanten)

(b) Gradverteilung

Abbildung 3.36: Unternehmensgraph als Vergleichsgrundlage für Verdrahtungsmodelle<sup>425</sup>

<sup>424</sup> Zur Datengrundlage oben S. 88–93.

<sup>425</sup> Der Layout-Algorithmus für Abbildung 3.36 (a) (S. 196) ist *Fruchterman-Reingold* in der Implementierung von *networkx-1.11* mit *seed 1234*; die Farbe der Punkte ist nach ihrem Grad von 0 (weiß) bis 20+ (schwarz) skaliert. Diese Einstellungen gelten auch für alle weiteren Graphdiagramme des Unternehmensgraphen, die in diesem Abschnitt gezeigt werden.

Kennt man nun lediglich die Anzahl der Unternehmen und die Anzahl der Verflechtungen zwischen ihnen, ohne irgendetwas über die Eigenschaften des Unternehmensnetzwerks oder über die Mechanismen zu wissen, die zu den Verflechtungsbeziehungen führen, so lässt sich für die Zwecke der Modellierung annehmen, dass die Verflechtungsbeziehungen zwischen den Unternehmen zufällig verteilt sind. Dadurch gelangt man zum schlichtesten Modell eines Netzwerks, das lediglich die Anzahl  $n$  der Punkte und die Anzahl  $m$  der Linien im Graphen als Parameter kennt: dem *Zufallsgraphen* (*random graph*),  $G(n, m)$ , der auch *Erdős-Rényi-Graph* genannt wird.<sup>426</sup> Die mathematische Analyse von Zufallsgraphen hat eine lange Tradition, sodass viel über ihre Eigenschaften bekannt ist.<sup>427</sup> Außerdem bilden Zufallsgraphen oft die Vergleichsgrundlage zur Bewertung erkannter Gemeinschaften anhand der Modularität, da man davon ausgeht, dass sie *per definitionem* keine Gemeinschaftsstrukturen enthalten.<sup>428</sup>

Durch  $G(n, m)$  wird nun nicht nur *ein* Graph, sondern ein *statistisches Ensemble* (*statistical ensemble*) von Graphen angegeben (*Graphenensemble*).<sup>429</sup>  $G(n, m)$  steht für alle Graphen, die  $n$  Punkte und  $m$  Linien haben und außerdem schlicht sind (also keine Schleifen oder Mehrfachkanten aufweisen); das sind  $\binom{N}{m}$  Graphen, wobei  $N = \binom{n}{2}$ .<sup>430</sup> Abbildung 3.37 (a) (S. 198) zeigt eine konkrete Realisierung (*Instanz*) des Ensembles  $G(100, 537)$ , zu dem auch der Unternehmensgraph gehört.

<sup>426</sup> Namensgebend ist *Erdős und Rényi*, *Publicationes Mathematicae* 6 (1959), S. 290–297. Dort wird  $G(n, m)$  notiert als  $\Gamma_{n,N}$ ; andernorts finden sich auch Varianten wie  $G_{N,E}$ ,  $G_{n,m}$  oder  $G_{n,M}$ .

<sup>427</sup> Eine Zusammenfassung bietet das Lehrbuch von *Bollobás*, *Random Graphs*, 2001. Zu den zahlreichen Erweiterungen, die unter Stichwörtern wie *p\* models* oder *exponential random graph models* (ERGM) diskutiert werden, etwa *Wasserman und Pattison*, *Psychometrika* 61 (1996), S. 401–425; *Pattison und Wasserman*, *British Journal of Mathematical and Statistical Psychology* 52 (1999), S. 169–193; *Anderson, Wasserman und Crouch*, *Social Networks* 21 (1999), S. 37–66; *Robins u.a.*, *Social Networks* 29 (2007), S. 173–191; *Robins u.a.*, *Social Networks* 29 (2007), S. 192–215; sowie die Beiträge in *Lusher* (Hrsg.), *Exponential Random Graph Models for Social Networks*, 2012.

<sup>428</sup> *Newman und Girvan*, *Physical Review E* 69 (2004), S. 026113-1 (7). Zur Modularität bereits S. 155–156.

<sup>429</sup> *Dorogovtsev und Mendes*, *Advances in Physics* 51 (2002), S. 1079 (1082); *Boccaletti u.a.*, *Physics Reports* 424 (2006), S. 175 (191).

<sup>430</sup> Dazu *Bollobás*, *Random Graphs*, 2001, S. 34. Herleitung: Für einen Graphen mit  $n$  Punkten gibt es  $N = \frac{n!}{2!(n-2)!} = \binom{n}{2}$  (lies: „ $n$  über 2“; im Englischen: „ $n$  choose 2“) Möglichkeiten, zwei verschiedene Punkte als Endpunkte einer Linie auszuwählen. Dann gibt es in einem Graphen  $\binom{N}{m}$  Möglichkeiten,  $m$  Paare von Endpunkten zur Definition von Linien auszuwählen.

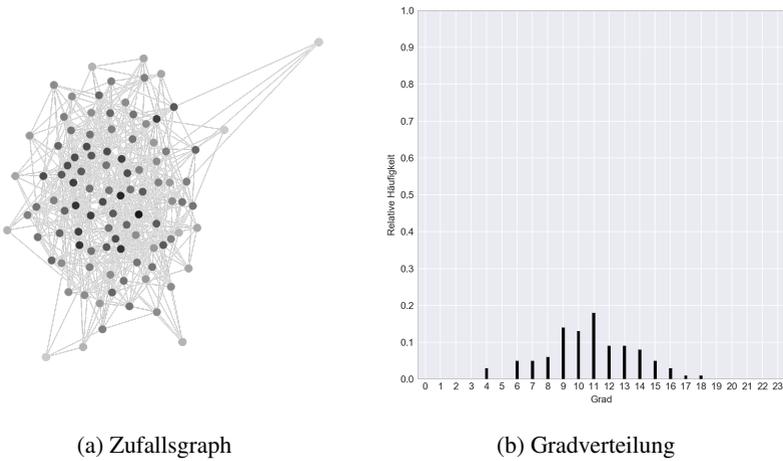


Abbildung 3.37: Annäherung des Unternehmensgraphen durch einen Zufallsgraphen<sup>431</sup>

Alle Mitglieder dieses Ensembles haben (wie der Unternehmensgraph) einen Durchschnittsgrad von 10,74 (abstrakt  $\frac{2m}{n}$ ). Die Gradverteilung des abgebildeten Zufallsgraphen (Abb. 3.37 (b), S. 198) weicht allerdings in ihrer Form deutlich von der Gradverteilung des Unternehmensgraphen (Abb. 3.36 (b), S. 196) ab. Die durchschnittliche Länge der kürzesten Wege liegt mit 2,17 relativ nah an der des Unternehmensgraphen (2,33); allerdings liegt der Clustering-Koeffizient mit 0,11 deutlich niedriger (im Unternehmensgraphen liegt er bei 0,41).

Der Vergleich zwischen dem Unternehmensgraphen und *einer* Instanz von  $G(100, 537)$  sagt allerdings wenig darüber aus, wie gut man ihn mit dem Modell des Zufallsgraphen beschreiben kann. Entscheidend ist, inwieweit der *durchschnittliche* Zufallsgraph – hier also die durchschnittliche Instanz von  $G(100, 537)$  – ein adäquates Modell für den Unternehmensgraphen ist.<sup>432</sup> Dass dies nicht der Fall ist, lässt sich abstrakt mit einer geringfügig abgewandelten Definition des Zufallsgraphen zeigen, die  $G(n, p)$  genannt wird. Hier steht  $n$  wieder für die Anzahl der Linien;  $p$  repräsentiert die Wahrscheinlichkeit, dass eine einzelne, mögliche Linie im Graphen tatsächlich vorhanden ist.<sup>433</sup> Die Gradverteilung von  $G(n, p)$  folgt einer Binomialverteilung  $B(n-1, p)$ ,<sup>434</sup> die sich für große  $(n-1)$  und kleine  $p$  durch eine Poissonverteilung mit Parameter  $\lambda = (n-1) \cdot p$ , d.h.  $Po((n-1) \cdot p)$

<sup>431</sup> Zu Layout und Färbung von Abbildung 3.37 (a) (S. 198) bereits Fn. 425 (S. 196).

<sup>432</sup> Wenn von den Eigenschaften eines Zufallsgraphen gesprochen wird, sind in der Regel die durchschnittlichen Eigenschaften des Ensembles gemeint (womit wiederum bestimmte Verteilungsannahmen verbunden sind); dazu auch *Newman*, *Networks*, 2010, S. 399.

<sup>433</sup> *Bollobás*, *Random Graphs*, 2001, S. 34–35. Frühe Formulierungen von  $G(n, p)$  finden sich bei *Solomonoff und Rapoport*, *Bulletin of Mathematical Biophysics* 13 (1951), S. 107 (107) sowie bei *Gilbert*, *The Annals of Mathematical Statistics* 30 (1959), S. 1141 (1141).

<sup>434</sup> *Bollobás*, *Discrete Mathematics* 33 (1981), S. 1 (5); *Newman*, *SIAM Review* 45 (2003), S. 167 (180); *Brinkmeier und Schank*, *Network Statistics*, in: *Brandes und Erlebach* (Hrsg.),

approximieren lässt.<sup>435</sup> Beide Verteilungen fallen (ähnlich wie die Normalverteilung) um ihren wahrscheinlichsten Wert herum symmetrisch ab (womit der Durchschnittsgrad eine gute Schätzung für den Grad eines beliebigen Punktes ist), während die Gradverteilung für den Unternehmensgraphen eher einer diskreten Gleichverteilung auf der Trägermenge  $\{0, 1, \dots, 23\}$  der natürlichen Zahlen nahe kommt. Außerdem ergibt sich der Clustering-Koeffizient für den durchschnittlichen Zufallsgraphen aus seinem Durchschnittsgrad, geteilt durch  $n - 1$ , im hiesigen Fall ist das Ergebnis 0,1084, gerundet auf zwei Nachkommastellen wiederum 0,11.

Allgemein hat man festgestellt, dass Zufallsgraphen reale Netzwerke zwar hinsichtlich ihrer Distanzen relativ gut beschreiben – der Durchmesser von Zufallsgraphen (die Länge des längsten kürzesten Weges zwischen zwei Punkten im Graphen) wächst nur logarithmisch mit der Anzahl der Punkte, sodass sie den Kleine-Welt-Effekt aufweisen.<sup>436</sup> Doch gerade beim Clustering-Koeffizienten zeigen Zufallsgraphen unrealistisch geringe Werte (bei gleichbleibendem Durchschnittsgrad geht der Clustering-Koeffizient mit wachsendem  $n$  gegen Null).<sup>437</sup> Hohe Clustering-Koeffizienten sind allerdings typisch für Graphen mit bestimmten, sehr regelmäßigen Strukturen, unter anderem für ringförmige Gittergraphen (*circular lattices* bzw. *ring lattices*).<sup>438</sup> Hier werden  $n$  Punkte im Kreis angeordnet und jeder Punkt wird mit seinen  $k$  (bzw.  $k - 1$ , falls  $k$  ungerade ist) nächsten Nachbarn (links und rechts von ihm) verbunden. Abbildung 3.38 (a) (S. 200) zeigt einen ringförmigen Gittergraphen mit  $n = 100$  und  $k = 11$ , jeder Punkt ist hier also mit seinen  $11 - 1 = 10$  nächsten Nachbarn verbunden. Für jedes Paar  $(n, k)$  gibt es genau einen ringförmigen Gittergraphen, dessen Durchschnittsgrad  $k$  bzw. (bei ungeradem  $k$ )  $k - 1$  entspricht. Der Durchschnittsgrad des in Abbildung 3.38 (a) (S. 200) gezeigten Graphen ist also 10,00; alle Punkte haben den Grad 10 und der Graph hat 500 Linien (etwas weniger als der Unternehmensgraph).

---

Network Analysis, 2005, S. 293 (294). Mithilfe der Binomialverteilung lässt sich unter anderem die Anzahl der Erfolge bei  $n$  voneinander unabhängigen Versuchen modellieren, von denen jeder die Erfolgswahrscheinlichkeit  $p$  hat; sie folgt einer Binomialverteilung mit den Parametern  $n$  und  $p$  ( $B(n, p)$ ). Ein Beispiel die Anzahl der Erfolge bei  $n$  Versuchen, aus einer Urne mit einem Anteil von  $p$  schwarzen Kugeln eine schwarze Kugel zu ziehen, wobei nach jedem Versuch die gezogene Kugel zurückgelegt wird (Urnenmodell: Ziehen mit Zurücklegen).

<sup>435</sup> *Fahrmeir u.a.*, Statistik, 2016, S. 243. Die Poissonverteilung wird außerhalb der Netzwerkforschung unter anderem zur Modellierung der Anzahl seltener Ereignisse innerhalb eines bestimmten Zeitintervalls eingesetzt; ihr Parameter  $\lambda$  gibt die Intensitätsrate an, mit der die seltenen Ereignisse auftreten; *Fahrmeir u.a.*, Statistik, 2016, S. 242. Als Faustregel für die Möglichkeit, eine Binomialverteilung  $B(n, p)$  durch eine Poissonverteilung  $Po(n \cdot p)$  zu approximieren, gilt in der Statistikkliteratur  $n > 30$  und  $p \leq 0,05$ ; *Fahrmeir u.a.*, Statistik, 2016, S. 297; in der Physikkliteratur wird meist nur von „large  $n$ “ gesprochen.

<sup>436</sup> *Albert und Barabási*, Reviews of Modern Physics 74 (2002), S. 47 (58); *Newman*, SIAM Review 45 (2003), S. 167 (199). Zum Kleine-Welt-Effekt bereits S. 168.

<sup>437</sup> *Newman*, SIAM Review 45 (2003), S. 167 (199).

<sup>438</sup> *Watts und Strogatz*, Nature 393 (1998), S. 440 (440).

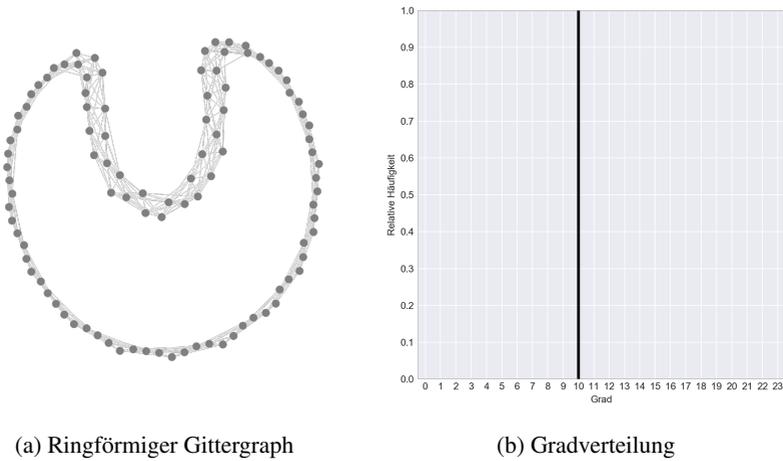


Abbildung 3.38: Annäherung des Unternehmensgraphen durch einen ringförmigen Gittergraphen<sup>439</sup>

Für ringförmige Gittergraphen ist neben ihrem hohen Clustering-Koeffizienten (für  $n = 100$  und  $k = 11$  beträgt er  $\frac{2}{3}$ ) auch charakteristisch, dass die kürzesten Wege zwischen den einzelnen Punkten im Durchschnitt verhältnismäßig lang sind (hier sind sie mit 5,45 knapp doppelt so lang wie im Unternehmensgraphen) – ringförmige Gittergraphen zeigen keinen Kleine-Welt-Effekt, was sie wiederum zu schlechten Modellen für die meisten realen Netzwerke macht.<sup>440</sup> Anders sieht es allerdings aus, wenn zu der regulären Struktur ringförmiger Gittergraphen eine Prise Zufall hinzugegeben wird (womit durch die Modellparameter wieder ein Graphenensemble definiert wird). Dies kann geschehen, indem man, ausgehend von einem ringförmigen Gittergraphen mit Parametern  $n$  und  $k$ , mit einer Wahrscheinlichkeit  $p$  eine der bestehenden Linien neu verdrahtet.<sup>441</sup> Dann erhält man für einige Werte von  $p$  (in etwa zwischen 0,001 und 0,1) Graphen, die sowohl niedrige Durchschnittsdistanzen als auch einen hohen Clustering-Koeffizienten aufweisen.<sup>442</sup> Dieses Modell wird oft als *Kleine-Welt-Modell* (*small-world model*)

<sup>439</sup> Zu Layout und Färbung von Abbildung 3.38 (a) (S. 200) bereits Fn. 425 (S. 196).

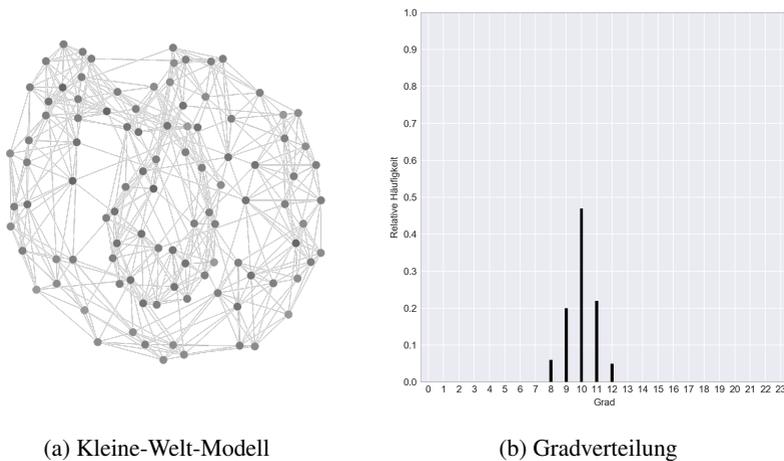
<sup>440</sup> Watts und Strogatz, Nature 393 (1998), S. 440 (440).

<sup>441</sup> Watts und Strogatz, Nature 393 (1998), S. 440–442; kritisch dazu Newman, SIAM Review 45 (2003), S. 167 (209). In einer später vorgeschlagenen Abwandlung des Vorgehens fügt man mit Wahrscheinlichkeit  $p$  neue Linien hinzu, anstatt bestehende Linien neu zu verdrahten; siehe z.B. Monasson, The European Physical Journal B 12 (1999), S. 555 (556); Newman und Watts, Physics Letters A 263 (1999), S. 341 (342); Newman und Watts, Physical Review E 60 (1999), S. 7332 (7333).

<sup>442</sup> Watts und Strogatz, Nature 393 (1998), S. 440 (441).

bezeichnet, teilweise nach den Forschern, auf die es zurückgeht, auch *Watts-Strogatz-Modell* genannt.<sup>443</sup>

Abbildung 3.39 (a) (S. 201) zeigt einen mit dem Kleine-Welt-Modell und den Parametern  $n = 100$ ,  $k = 11$  und  $p = 0,1$  erzeugten Graphen, der wiederum aus 100 Punkten und 500 Linien zusammengesetzt ist (Durchschnittsgrad: 10,00). Dieser Graph hat einen Clustering-Koeffizienten von 0,49 und eine durchschnittliche Länge der kürzesten Wege von 2,70, womit er relativ nah an den Werten des Unternehmensgraphen (0,41 und 2,33) liegt. Seine Gradverteilung ist aber weiterhin einer Poissonverteilung ähnlich und damit von der Gradverteilung des Unternehmensgraphen sehr verschieden.



(a) Kleine-Welt-Modell

(b) Gradverteilung

Abbildung 3.39: Annäherung des Unternehmensgraphen durch Umstecken (*rewiring*) von Linien eines ringförmigen Gittergraphen<sup>444</sup>

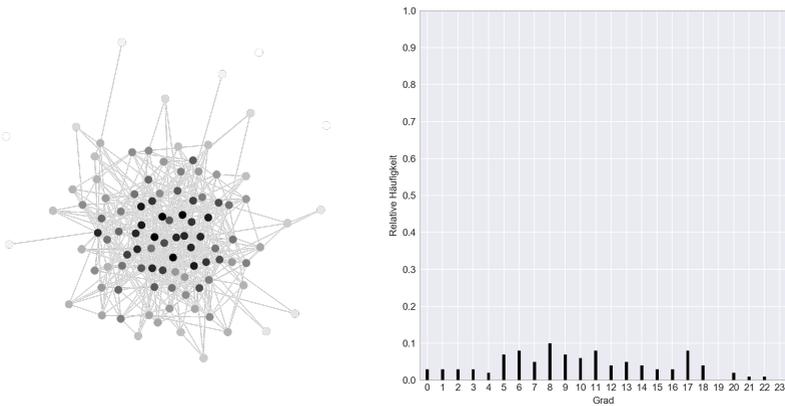
Will man Graphen mit einer bestimmten Gradverteilung erzeugen, so bietet das *Konfigurationsmodell* (*configuration model*) eine relativ simple Lösung. Dieses Modell erhält als Eingabe eine Gradverteilung (aus der auch die Anzahl der Punkte im Graphen hervorgeht).<sup>445</sup> Daraus wird ein Graph erzeugt, indem man jedem Punkt so viele *Stümpfe* (*stubs*), d.h. Endpunkte von Linien zuordnet, wie es seinem

<sup>443</sup> Namensgebend sind *Watts und Strogatz*, *Nature* 393 (1998), S. 440–442. Die Bezeichnung als Kleine-Welt-Modell ist allerdings unglücklich, da auch Zufallsgraphen mit niedrigen Clustering-Koeffizienten den Kleine-Welt-Effekt zeigen.

<sup>444</sup> Zu Layout und Färbung von Abbildung 3.39 (a) (S. 201) bereits Fn. 425 (S. 196).

<sup>445</sup> Das Modell wird z.B. studiert von *Molloy und Reed*, *Random Structures and Algorithms* 6 (1995), S. 161–179; *Molloy und Reed*, *Combinatorics, Probability and Computing* 7 (1998), S. 295–305; *Newman, Strogatz und Watts*, *Physical Review E* 2001, S. 026118-1–026118-17; zusammenfassend *Newman, Networks*, 2010, S. 434–445.

Grad entspricht, und die Stümpfe dann zufällig miteinander verbindet.<sup>446</sup> Dabei können grundsätzlich Mehrfachkanten und Schleifen entstehen, und aufgrund der Zufallskomponente ist durch die Gradverteilung als Parameter wieder ein Graphenensemble gegeben.<sup>447</sup> Abbildung 3.40 (a) (S. 202) zeigt eine Instanz des Ensembles, das durch die Gradverteilung des Unternehmensgraphen definiert ist. Mehrfachkanten wurden entfernt, sodass 100 Punkte und 494 Linien bei einem Durchschnittsgrad von 9,88 übriggeblieben sind und die Gradverteilung derjenigen des Unternehmensgraphen (nur) näherungsweise entspricht (Abb. 3.40 (b), S. 202). Die durchschnittliche Länge der kürzesten Wege liegt mit 2,26 nah an der des Unternehmensgraphen; hingegen ist der Clustering-Koeffizient mit 0,15 wieder unpassend niedrig.



(a) Konfigurationsmodell für ungerichtete Graphen

(b) Gradverteilung

Abbildung 3.40: Annäherung des Unternehmensgraphen anhand der Gradverteilung<sup>448</sup>

Die vorgestellten Verdrahtungsmodelle werden hauptsächlich abstrakt studiert oder als Grundlage der Simulation von Prozessen auf Netzwerken genutzt.<sup>449</sup> Besonders interessiert ist man an Situationen, in denen mit minimalen Veränderungen von Modellparametern große Veränderungen von Netzwerkcharakteristika einhergehen. Diese abrupten Veränderungen werden *Phasenübergänge* (*phase transi-*

<sup>446</sup> Newman, Networks, 2010, S. 435. Eine andere Möglichkeit, aus einer gegebenen Gradverteilung einen Graphen zu konstruieren, besteht darin, für jeden Punkt so viele Kopien in einer Urne vorzusehen, wie es seinem gewünschten Grad entspricht, und die Linien als Paare von Kugeln aus der Urne zu ziehen. In diesem Sinne (wenngleich in der Terminologie der Mengen formuliert) Molloy und Reed, Random Structures and Algorithms 6 (1995), S. 161 (165).

<sup>447</sup> Molloy und Reed, Random Structures and Algorithms 6 (1995), S. 161 (165); Newman, Strogatz und Watts, Physical Review E 2001, S. 026118-1 (2).

<sup>448</sup> Zu Layout und Färbung von Abbildung 3.40 (a) (S. 202) bereits Fn. 425 (S. 196).

<sup>449</sup> Zu Prozessen auf Netzwerken unten S. 211–223.

tions) genannt und spielen eine zentrale Rolle, wenn es darum geht, das dynamische Verhalten von komplexen Systemen zu verstehen.<sup>450</sup> Beispielsweise besteht ein Zufallsgraph mit Durchschnittsgrad kleiner oder gleich 1 (d.h. für  $G(n, m)$ :  $\frac{2m}{n} \leq 1$ ) in der Regel aus vielen kleinen Komponenten, während es für Zufallsgraphen mit Durchschnittsgrad größer als 1 eine Zusammenhangskomponente gibt, die proportional mit  $n$  wächst (sog. *Riesenkompone*te, *giant component*).<sup>451</sup>

Wer sich für die näherungsweise Beschreibung eines konkreten Netzwerks, beispielsweise des Unternehmensnetzwerks, durch ein Verdrahtungsmodell interessiert, wird auf unterschiedliche Modelle zurückgreifen oder diese kombinieren müssen, je nachdem, welche Charakteristika des Netzwerks ihm besonders wichtig sind. Der Verflechtungsgraph, aus dem der Unternehmensgraph abgeleitet ist, wurde mit einem Modell konstruiert, das Teile des Konfigurationsmodells mit Teilen des Zufallsgraphen kombiniert und dabei die gewünschte bipartite Struktur garantiert.<sup>452</sup> Jedes Unternehmen wurde mit so vielen Führungskräften zufällig verbunden, wie es Sitze in seinem Führungsgremium gibt, wobei die Wahrscheinlichkeit einer Führungsfigur, Mitglied eines Gremiums zu werden, zusätzlich von ihrer Form abhängig gemacht wurde.<sup>453</sup>

Alle skizzierten Verdrahtungsmodelle gehen in ihrer Grundform von ungerichteten Graphen aus, lassen sich aber auf gerichtete Graphen übertragen. So kann man etwa mit einem leicht modifizierten Konfigurationsmodell einen gerichteten Graphen erzeugen, der annähernd dieselbe Innen- und Außengradverteilung hat wie der Zitiergraph der flatländischen Rechtsprechung (Abb. 3.41, S. 204). Problematisch ist dieses Vorgehen allerdings insbesondere dann, wenn es um gerichtete Graphen geht, die – wie die meisten Zitiergraphen – annähernd azyklisch sind.<sup>454</sup> Denn die vorgestellten Modelle bieten bei ihrer Anpassung an gerichtete Graphen keinerlei Gewähr dafür, dass keine Zyklen auftreten.<sup>455</sup> Eine Möglichkeit, die Eigenschaft von Zitiergraphen als gerichtete, azyklische Graphen zu erhalten, besteht darin, ihre Struktur nicht mittels Verdrahtung, sondern durch Nachempfindung

<sup>450</sup> Ein speziell auf die Untersuchung von Netzwerken zugeschnittener Überblick aus der Perspektive der statistischen Physik findet sich bei *Barrat, Barthélemy und Vespignani*, *Dynamical Processes on Complex Networks*, 2008, S. 92–115; knapp und anschaulich außerdem *Barabási*, *Network Science*, 2016, S. 109–110.

<sup>451</sup> *Cohen und Havlin*, *Complex Networks*, 2010, S. 10; *Newman*, *Networks*, 2010, S. 404; in der Formulierung  $G(n, m)$  liegt der Phasenübergang für die Existenz und Größe einer Riesenkompone

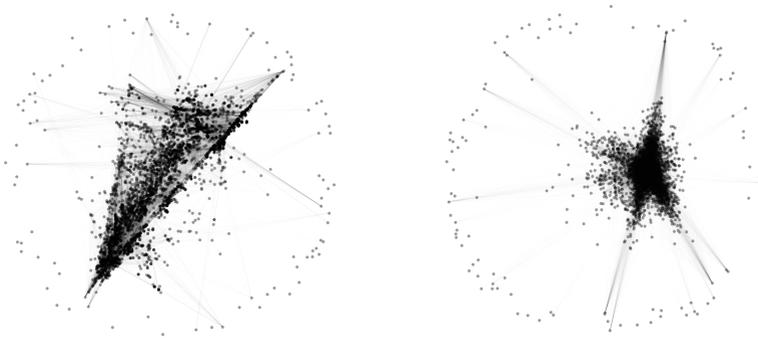
<sup>452</sup> Zur Erinnerung: Ein Graph ist bipartit, wenn man seine Punktmenge so in zwei disjunkte Teilmengen einteilen kann, dass alle Linien zwischen Punkten in unterschiedlichen Mengen verlaufen.

<sup>453</sup> Der Code, mit dem der Verflechtungsgraph erstellt wurde, ist im Online-Appendix einsehbar.

<sup>454</sup> Zu Konfigurationsmodellen für gerichtete azyklische Graphen *Karrer und Newman*, *Physical Review E* 80 (2009), S. 046110-1–046110-12.

<sup>455</sup> Beispielsweise ist der in Abbildung 3.41 (b) (S. 204) gezeigte Graph im Gegensatz zum Zitiergraphen selbst (Abb. 3.41 (a), S. 204) nicht azyklisch.

ihres Entstehungsprozesses zu definieren – auf Basis eines Wachstumsmodells (dazu sogleich).



(a) Zitiergraph ohne Mehrfachkanten

(b) Konfigurationsmodell für gerichtete Graphen

Abbildung 3.41: Annäherung des Zitiergraphen anhand von Innen- und Außengradverteilung<sup>456</sup>

## 2. Wachstumsmodelle

Während Verdrahtungsmodelle in der Regel davon ausgehen, dass die Anzahl der Punkte im Graphen sich bei der Definition der Verbindungsstruktur nicht verändert, beginnen Wachstumsmodelle typischerweise mit einigen wenigen oder gar keinen Punkten. Der Graph wird dann konstruiert, indem man nach und nach einzelne Punkte und Verbindungen dieser Punkte zu anderen, bereits im Graphen befindlichen Punkten hinzufügt – die Verbindungsstruktur wird durch Wachstum erzeugt.<sup>457</sup>

<sup>456</sup> Abbildung 3.41 (a) (S. 204) zeigt den Zitiergraphen der flatländischen Rechtsprechung (Version ohne Mehrfachkanten) im kräftebasierten Layout (Fruchterman-Reingold-Algorithmus), generiert mit *networkx-1.11* und *seed 1234*. Dieser Graph hat 5.034 Punkte, 43.310 Pfeile und einen durchschnittlichen Innen- bzw. Außengrad von 8,60. Abbildung 3.41 (b) (S. 204) zeigt einen Graphen mit ebenfalls 5.034 Punkten, 42.772 Pfeilen und einem durchschnittlichen Innen- bzw. Außengrad von 8,50, der mithilfe eines Konfigurationsmodells für gerichtete Graphen erzeugt wurde, ohne Mehrfachkanten zuzulassen, sodass Innengradverteilung und Außengradverteilung den entsprechenden Verteilungen des Zitiergraphen (nur) näherungsweise entsprechen.

<sup>457</sup> Der Gedanke hinter diesem Vorgehen ist, dass die Prozesse, die im Modell dazu führen, eine in der Realität beobachtete Struktur hervorzubringen, auch Ursache der beobachteten Struktur sein könnten; man spricht daher auch von generativen Netzwerkmodellen (*generative network models*); siehe etwa Newman, *Networks*, 2010, S. 486.

Diese Vorgehensweise kommt der Entstehung von Rechtsprechungszitiernetzwerken recht nahe: Unabhängig davon, wie genau man das Netzwerk definiert, gibt es immer eine oder mehrere „älteste“ Entscheidungen, die zeitlich zuerst ins Netzwerk eintreten und von denen daher regelmäßig keine Zitate zu anderen Entscheidungen im Netzwerk ausgehen können. Die Beziehungsstruktur eines Rechtsprechungszitiernetzwerks ergibt sich allein daraus, dass juristische Akteure einzelner Spruchkörper (in der Regel Richter oder ihre wissenschaftlichen Mitarbeiter) in den Entscheidungstexten auf andere Entscheidungen hinweisen;<sup>458</sup> je älter eine Entscheidung, desto häufiger kann sie prinzipiell zitiert werden.

Will man nicht nur ein konkretes Rechtsprechungszitiernetzwerk beschreiben, sondern Erkenntnisse über Rechtsprechungszitiernetzwerke im Allgemeinen gewinnen, so bietet sich die Untersuchung von Wachstumsmodellen an. Da Zitiernetzwerke der Hauptanwendungsfall für zeitgebundene, reine Wachstumsmodelle sind und Zitate in Graphen durch Pfeile dargestellt werden, haben die folgenden Ausführungen vor allem Wachstumsmodelle für gerichtete Graphen im Blick. Sie lassen sich aber ohne Weiteres auf ungerichtete Graphen übertragen. Alle diskutierten Modelle beinhalten stochastische Elemente (Wahrscheinlichkeiten bzw. Zufall), sodass sie jeweils Ensembles von Graphen definieren.

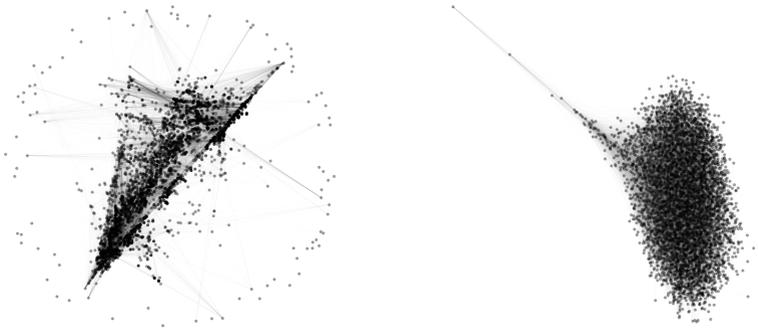
Das schlichteste Wachstumsmodell hat als Parameter lediglich die Anzahl der Punkte  $n$ , die der Graph nach Abschluss des Wachstums haben soll. Das Wachstum des Graphen erfolgt dann so, dass jeder Punkt (mit Ausnahme des ersten) zufällig mit genau einem bereits im Graphen befindlichen Punkt verbunden wird. Bei gerichteten Graphen erfolgt die Verbindung durch einen Pfeil vom jüngeren zum älteren Punkt; der resultierende Graph ist dann ein Baum mit  $n$  Punkten,  $n - 1$  Pfeilen und einem durchschnittlichen Innen- bzw. Außengrad von  $\frac{n-1}{n}$ .<sup>459</sup> Das entspricht einem Modell, in dem jede Gerichtsentscheidung (mit Ausnahme der ältesten) zufällig genau eine andere Entscheidung zitiert, und ist nicht sehr realistisch. Man kann das Modell verallgemeinern, indem man mehr als einen Pfeil pro neu hinzu kommendem Punkt zulässt – beispielsweise durch einen Parameter  $k$ , der die Anzahl der gewünschten Pfeile vorgibt. So kann man bereits einen Graphen konstruieren, der hinsichtlich der Anzahl von Punkten und Pfeilen annähernd einem Vergleichsgraphen (etwa dem Zitiergraphen ohne Mehrfachkanten) entspricht, indem man für  $k$  einen Wert nahe am Durchschnittsgrad des Vergleichsgraphen wählt.<sup>460</sup>

<sup>458</sup> Zur temporalen Struktur von Rechtsprechungszitiernetzwerken bereits S. 187–193.

<sup>459</sup> *Krapivsky und Redner*, Physical Review E 63 (2001), S. 066123-1 (1–2).

<sup>460</sup> *Krapivsky und Redner*, Physical Review E 63 (2001), S. 066123-1 (13). Für ein ähnliches Modell, das allerdings keine Azyklichkeit garantiert, siehe außerdem *Peterson*, Advances in Applied Probability 47 (2015), S. 858–879.

Wie Abbildung 3.42 (S. 206) verdeutlicht, wird das der Struktur des flatländischen Zitiergraphen allerdings nicht gerecht. Das liegt unter anderem daran, dass nicht jede Entscheidung gleich viele andere Entscheidungen zitiert. Es liegt aber vor allem daran, dass die Zitate nicht zufällig sind – nicht jedes mögliche Zitat wird mit gleicher Wahrscheinlichkeit realisiert. Die Innengradverteilung des in Abbildung 3.42 (b) (S. 206) gezeigten Modellgraphen ist zwar schief, aber längst nicht so schief wie in den meisten Zitiernetzwerken. Gesucht wird also ein Wachstumsmodell, das sehr schiefe Innengradverteilungen produziert.



(a) Zitiergraph ohne Mehrfachkanten

(b) Einfaches Wachstumsmodell

Abbildung 3.42: Annäherung des Zitiergraphen durch ein einfaches Wachstumsmodell<sup>461</sup>

Eines der ältesten Wachstumsmodelle für gerichtete Graphen macht die Wahrscheinlichkeit eines Dokuments, zu einem konkreten Zeitpunkt zitiert zu werden, davon abhängig, wie viele Zitate es zuvor erhalten hat: je häufiger ein Dokument bereits zitiert wurde, desto höher die Wahrscheinlichkeit, dass es erneut zitiert wird.<sup>462</sup> Dies führt zur Akkumulation von Vorteilen (*kumulativer Vorteil, cumulative advantage*),<sup>463</sup> weshalb das Modell hier als *Kumulationsmodell* bezeichnet wird. Das Kumulationsmodell erfasst ein Phänomen, welches auch *Matthäus-*

<sup>461</sup> Abbildung 3.42 (a) (S. 206) zeigt den Zitiergraphen wie in Abbildung 3.41 (a) (S. 204). Abbildung 3.42 (b) (S. 206) zeigt im identischen Layout einen gerichteten Graphen mit 5.034 Punkten, 40.236 Pfeilen und einem durchschnittlichen Innen- bzw. Außengrad von 7,99, der durch das im Haupttext beschriebene simple Wachstumsmodell mit  $k = 8$  generiert wurde.

<sup>462</sup> Price, Journal of the American Society for Information Science 27 (1976), S. 292–306.

<sup>463</sup> Merton, Science 159 (1968), S. 56–63; Price, Journal of the American Society for Information Science 27 (1976), S. 292–306. Die Studie von Price wird im juristischen Diskurs rezipiert unter anderem von Wagner-Döbler, Scientometrics 29 (1994), S. 15 (18).

Effekt (*Matthew effect*)<sup>464</sup> oder ‚*the rich get richer*‘-phenomenon<sup>465</sup> genannt wird. Dieses Phänomen bringt Graphen mit sehr schiefer Innengradverteilung hervor – wie schief die Innengradverteilung ist, hängt davon ab, wie man die Abhängigkeit von der Zitierhäufigkeit im Einzelnen modelliert. In der jüngeren Literatur wird die Bevorzugung von Punkten mit vielen Verbindungen bei der Gewinnung neuer Verbindungen unter dem Stichwort *Präferenzenanbindung* (*preferential attachment*) diskutiert.<sup>466</sup> Die Bezeichnung geht auf einen Artikel zurück, in dem das Kumulationsmodell für ungerichtete Graphen wiederentdeckt wurde, und der so einflussreich geworden ist, dass das Modell mittlerweile unter dem Namen der Autoren als *Barabási-Albert-Modell* (*Barabási-Albert model*) firmiert.<sup>467</sup>

Bedenkt man, wie Rechtsprechungsziernetzwerke wachsen, so lässt sich erwägen, die Zitierwahrscheinlichkeit einer Entscheidung nicht nur von ihrem Innengrad, sondern auch von ihrem Außengrad abhängig zu machen: Juristische Datenbanken führen sowohl die zitierten als auch die zitierenden Entscheidungen in den Annotationen zu einer Entscheidung an, sodass der Zitiergraph sich bei der Recherche als ungerichteter Graph darstellt.<sup>468</sup> Damit ist die Wahrscheinlichkeit, dass jemand, der Fallrecht recherchiert, bei einer zufälligen Bewegung über den Zitiergraphen auf eine bestimmte Entscheidung stößt (und damit überhaupt erst

<sup>464</sup> Merton, *Science* 159 (1968), S. 56–63, in Anlehnung an Mt 25, 29: „Denn wer da hat, dem wird gegeben werden, und er wird die Fülle haben; wer aber nicht hat, dem wird auch, was er hat, genommen werden.“ Dass die Bezeichnung des beschriebenen Phänomens angesichts des zweiten Halbsatzes nicht ganz passend ist, bemerkt Price, *Journal of the American Society for Information Science* 27 (1976), S. 292 (293); aufgegriffen von Newman, *SIAM Review* 45 (2003), S. 167 (213 Fn. 29).

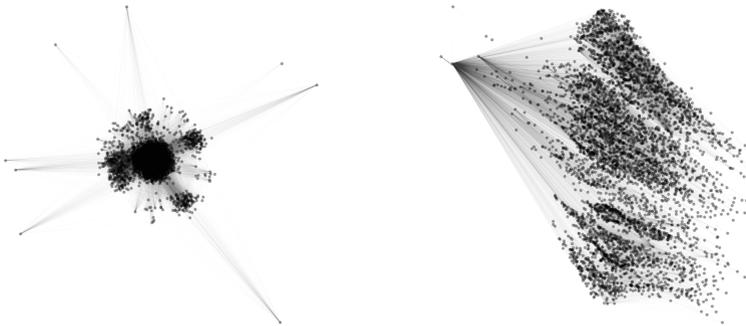
<sup>465</sup> Baumann und Stiller, *Network Models*, in: Brandes und Erlebach (Hrsg.), *Network Analysis*, 2005, S. 341 (348).

<sup>466</sup> Das Grundmodell ist zu finden bei Barabási und Albert, *Science* 286 (1999), S. 509–512, deren Terminologie sich anscheinend durchgesetzt hat; siehe etwa Dorogovtsev, Mendes und Samukhin, *Physical Review Letters* 85 (2000), S. 4633 (4633); Watts, *Annual Review of Sociology* 30 (2004), S. 243 (251); Boccaletti u.a., *Physics Reports* 424 (2006), S. 175 (95). In der juristischen Netzwerkforschung wird *preferential attachment* als Konzept aufgegriffen z.B. von Smith, *San Diego Law Review* 44 (2007), S. 309 (320); Cross, Smith und Tomarchio, *Emory Law Journal* 57 (2008), S. 1227–1257 (1241); Strandburg u.a., *North Carolina Law Review* 87 (2009), S. 1657 (1667); Boulet, Mazzeza und Bourcier, *Network Analysis of the French Environmental Code*, in: Casanovas u.a. (Hrsg.), *AI Approaches to the Complexity of Legal Systems*, 2010, S. 39–53 (45); Katz und Stafford, *Ohio State Law Journal* 71 (2010), S. 457 (501).

<sup>467</sup> Barabási und Albert, *Science* 286 (1999), S. 509–512. Es gibt in erster Linie drei Modelle für Netzwerkstrukturen, die (auch) unter den Namen von Autorenpaaren bekannt sind: Das Erdős-Rényi-Modell, das Watts-Strogatz-Modell und das Barabási-Albert-Modell. Die Popularität dieser Bezeichnungen dürfte sich auch daraus ergeben, dass man sie leicht abkürzen kann: ER-Modell, WS-Modell und BA-Modell. Hier werden die autorengelunden Bezeichner vermieden, um die hinter den Modellen stehenden Ideen in den Vordergrund zu rücken und der Tatsache Rechnung zu tragen, dass die Genannten nur einen kleinen Teil der Forscher repräsentieren, welche die Modellierung von Netzwerkstrukturen vorangetrieben haben.

<sup>468</sup> Dazu außerdem oben bei Abb. 3.18 (S. 128) und unten bei Abb. 3.44 (S. 213).

in die Position gelangt, diese zu zitieren), abhängig nicht allein vom Innengrad des entsprechenden Punktes, sondern von der Summe von Innengrad und Außengrad (d.h. vom Grad des Punktes, wenn man den Zitiergraphen als ungerichteten Graphen betrachtet). Dann kann man zur Graphenkonstruktion sogar das Kumulationsmodell für ungerichtete Graphen benutzen, indem man den Graphen zunächst mit Linien wachsen lässt und die Linien dann entsprechend der Punktreihenfolge zeitkompatibel in Pfeile überführt (Abb. 3.43 (a), S. 208).



(a) Kumulation anhand der Summe von Innen- und Außengrad (b) Kopie von 100 % der ausgehenden Pfeile eines Vorbilds

Abbildung 3.43: Annäherung des Zitiergraphen durch Kumulationsmodell und Kopiermodell<sup>469</sup>

Unbefriedigend ist allerdings, dass das Kumulationsmodell die wechselseitige Unabhängigkeit der einzelnen, von einem Dokument (z.B. einer Gerichtsentscheidung) ausgehenden Zitate impliziert. Das entspricht der Situation, in der ein Jurist für jedes Zitat einen neuen Rechercheanlauf unternimmt, bei dem er sich zufällig über den Zitiergraphen bewegt – und ist entsprechend unrealistisch. Praktisch ist die Wahrscheinlichkeit, dass der Autor einer Entscheidung A in dieser Entscheidung eine Entscheidung C zitiert, aber erhöht, wenn er bereits die Entscheidung B zitiert hat und Entscheidung C in den Annotationen zu B enthalten ist.<sup>470</sup> Mit

<sup>469</sup> Abbildung 3.43 (a) (S. 208) zeigt im Layout von Abbildung 3.41 (a) (S. 204) einen gerichteten Graphen mit 5.034 Punkten, 40.208 Pfeilen und einem durchschnittlichen Innen- bzw. Außengrad von 7,99, der mithilfe eines ungerichteten Kumulationsmodells erzeugt und dann in einen gerichteten Graphen überführt wurde. Abbildung 3.43 (b) (S. 208) zeigt in demselben Layout einen gerichteten Graphen mit 5.034 Punkten, 34.615 Pfeilen und einem durchschnittlichen Innen- bzw. Außengrad von 6,88, der mit einem extremen Kopiermodell generiert wurde, in dem ein Punkt 100 % der Pfeile seines Vorbilds kopiert.

<sup>470</sup> Hierin liegt ein Grund für das Auftreten von bibliographischer Kopplung und Co-Zitation als Beziehungen zweiten Grades, dazu bereits S. 38–39.

dieser Motivation kann man ein Wachstumsmodell definieren, in dem jeder hinzukommende Punkt sich unter den bereits im Graphen befindlichen Punkten ein Vorbild sucht und einen Teil der Verbindungen dieses Vorbilds kopiert.<sup>471</sup> Dieses Modell wird hier als *Kopiermodell* bezeichnet.<sup>472</sup>

So kann man beispielsweise für ein Rechtsprechungsziernetzwerk festlegen, dass eine Entscheidung sich eine „Leitentscheidung“ aussucht und alle Entscheidungen zitiert, die in dieser Entscheidung zitiert werden. Das entspricht einem extremen, weil 100 % der ausgehenden Verbindungen des Vorbilds reproduzierenden Kopiermodell für gerichtete Graphen (Abb. 3.43 (b), S. 208).<sup>473</sup> Eine mögliche Modifikation besteht darin, einen Punkt nicht alle, sondern nur  $m$  zufällig ausgewählte Pfeile kopieren zu lassen und einen Anteil  $p$  dieser Pfeile, ähnlich wie im Kleine-Welt-Modell, zufällig neu zu verdrahten, um eine realistischere Zitiernetzstruktur zu erzeugen.<sup>474</sup> Man kann zeigen, dass die Wachstumsdynamik, die sich aus diesem Kopiermodell ergibt, für große Graphen zur Wachstumsdynamik des Kumulationsmodells äquivalent ist.<sup>475</sup> Das ist unter anderem deshalb bemerkenswert, weil das Kumulationsmodell nur mit globalem Wissen funktioniert (für jeden Punkt muss dessen [Innen-]Grad bekannt sein), während das Kopiermodell mit lokalem Wissen auskommt (es müssen nur ein Punkt als Vorbild und dessen Nachbarschaft im Graphen bekannt sein).<sup>476</sup>

Für reale Zitiernetzwerke ist ein reines Kopiermodell ebenso unpassend wie ein reines Kumulationsmodell. Beide Wachstumsmodelle lassen sich zwar (wie oben skizziert) durch die Eigenheiten eines Recherchevorgangs motivieren, sodass sie auf abstrakte Weise interessante Aspekte von konkreten Zitierprozessen erfassen und – im Gegensatz zu vielen Verdrahtungsmodellen – nicht nur eine Struktur, sondern (zumindest teilweise) auch Erklärungen für ihr Zustandekommen anbieten. Tatsächlich fließen aber noch viele weitere Faktoren in das Phänomen „Zitation“ ein, die ein Wachstumsmodell berücksichtigen kann. Modelle, die sol-

---

<sup>471</sup> Man kann sich das Prinzip auch am Beitritt zu einem Online-Netzwerk verdeutlichen, zu dem man eingeladen werden muss: Wer eine Einladung akzeptiert, geht oft zuerst die Kontaktliste des Einladenden durch, um weitere Menschen im Netzwerk zu finden, die er bereits kennt. Dass das nicht unwahrscheinlich ist, hängt wiederum mit der hohen Transitivität von Beziehungen in sozialen Netzwerken zusammen, dazu bereits oben S. 170.

<sup>472</sup> Im Englischen spricht man vom (*vertex copying model*; siehe etwa Newman, *SIAM Review* 45 (2003), S. 167 (223–224). In der juristischen Netzwerkforschung werden Kopiermodelle erwähnt von Katz und Stafford, *Ohio State Law Journal* 71 (2010), S. 457 (502 Fn. 53).

<sup>473</sup> So zu finden bei Krapivsky und Redner, *Physical Review E* 71 (2005), S. 036118-1–036118-7, und implementiert in *networkx-1.11*.

<sup>474</sup> Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 68–70.

<sup>475</sup> Herleitung bei Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 69–70.

<sup>476</sup> Dies wird auch hervorgehoben von Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 70.

che Faktoren als Eigenschaften einzelner Punkte (hier: Gerichtsentscheidungen) berücksichtigen, werden *Fitnessmodelle* (*fitness models*) genannt.<sup>477</sup>

Das einfachste Fitnessmodell für ein Rechtsprechungsziiternetzwerk beginnt mit einem Kumulationsmodell und führt lediglich einen zusätzlichen Parameter für jede Entscheidung ein, der ihre Wahrscheinlichkeit beeinflusst, zitiert zu werden.<sup>478</sup> In diesem Parameter sind alle Faktoren zusammengefasst, die neben dem (Innen-)Grad zur Zitierwahrscheinlichkeit einer Entscheidung beitragen. Diese Lösung ist mathematisch elegant, für Rechtsprechungsziiternetzwerke aber möglicherweise nicht differenziert genug. Denn die Faktoren, welche die Zitierwahrscheinlichkeit einer einzelnen Entscheidung beeinflussen, können konstant oder variabel sein (Beispiele: das entscheidende Gericht [konstant] und der Innengrad [variabel]); ihr Einfluss kann als absolut (für alle potenziell zitierenden Entscheidungen gleich) oder relativ (abhängig von den Eigenschaften einer potenziell zitierenden Entscheidung) modelliert werden.<sup>479</sup> Intuitiv kann man beispielsweise davon ausgehen, dass die Zitierwahrscheinlichkeit für Entscheidungen von Gerichten auf der mittleren Ebene bei drei Instanzen auch davon abhängt, auf welcher Ebene das potenziell zitierende Gericht in der Gerichtshierarchie positioniert ist (relativer Einfluss eines konstanten Faktors). Wie sich das Entscheidungsdatum und – davon abgeleitet – das Alter einer Entscheidung auf die Zitierwahrscheinlichkeit auswirken, hängt vielleicht nicht nur vom entscheidenden Gericht, sondern auch vom Thema der Entscheidung ab. Ferner dauert es eine Weile, bis neue Entscheidungen in der Rechtsprechung verarbeitet werden, sodass man nicht davon ausgehen kann, dass die Zitierwahrscheinlichkeit einer Entscheidung nach ihrer Veröffentlichung monoton wächst oder fällt. Der Zitiergraph der flatländischen Rechtsprechung, der für die Zwecke dieser Arbeit konstruiert wurde, macht die Zitierwahrscheinlichkeit einer Entscheidung vor allem von der Position des Gerichts in der Gerichtshierarchie, vom (jeweils zum potenziellen Zitierzeitpunkt betrachteten) Entscheidungsalter sowie davon abhängig, dass die Themen von zitierender und zitierter Entscheidung einander „nahestehen“.<sup>480</sup>

<sup>477</sup> Albert und Barabási, *Reviews of Modern Physics* 74 (2002), S. 47 (81–82); Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 67; Cohen und Havlin, *Complex Networks*, 2010, S. 59.

<sup>478</sup> So die Formulierung bei Bianconi und Barabási, *Europhysics Letters* 54 (2001), S. 436 (437), für ungerichtete Graphen. Eine andere Formulierung kommt ohne das Kumulationsmodell als Grundlage aus und macht die Verbindungswahrscheinlichkeit zwischen zwei Punkten ausschließlich von deren Fitness abhängig; dazu Caldarelli u.a., *Physical Review Letters* 2002, S. 258702-1–258702-4.

<sup>479</sup> Gerade variable Faktoren wie der Innengrad können dabei auch von externen, nur begrenzt beobachtbaren Einflüssen wie etwa der Tatsache getrieben sein, dass manche Themen besonders häufig vor Gericht getragen werden. Nimmt man an, dass Zitate in Gerichtsentscheidungen in erster Linie aufgrund inhaltlicher Relevanz erfolgen, so werden Entscheidungen, deren Themen häufiger vor Gericht getragen werden, insgesamt mit höherer Wahrscheinlichkeit zitiert.

<sup>480</sup> Der Code, mit dem der Zitiergraph erstellt wurde, ist im Online-Appendix einsehbar; die inhaltliche Nähe wird dort anhand von Inhaltsziffern simuliert. Das Modell ist nicht darauf

Ob und wie sich die Faktoren, die das Zitierverhalten beeinflussen, sinnvoll in ein Fitnessmodell integrieren lassen, ohne die Anzahl seiner Parameter deutlich zu erhöhen (was seine analytische Attraktivität mindern würde), ist noch weitgehend ungeklärt. Möglicherweise sind agentenbasierte Modelle, die das Zitierverhalten von Spruchkörpern als Agenten bei der Produktion von Entscheidungen modellieren, besser geeignet, um das Wachstum realer Rechtsprechungsziernetzwerke zu beschreiben.<sup>481</sup> Aus rechtsvergleichender und rechtstheoretischer Sicht besonders interessant wäre ein agentenbasiertes Modell, welches das Wachstum von Rechtsprechungsziernetzwerken in *Common Law Jurisdictions* ebenso gut beschreibt wie jenes von Rechtsprechungsziernetzwerken in *Civil Law Jurisdictions*: Die Existenz eines solchen Modells spräche dafür, dass die Unterschiede zwischen beiden Systemen in der Praxis geringer sind, als oft angenommen wird. Hier begibt man sich auf ein Terrain, das in der juristischen Netzwerkforschung bisher weitgehend unerschlossen ist.

## II. Prozesse

Bei der Modellierung von Netzwerkstrukturen bildet die Netzwerktopologie selbst das Zentrum der Analyse. Demgegenüber steht bei der Modellierung von Prozessen in und auf Netzwerken die Frage im Mittelpunkt, wie Ablauf und Ergebnisse bestimmter Vorgänge mit einzelnen Netzwerktopologien zusammenhängen. Für die juristische Netzwerkforschung sind dabei insbesondere zwei Arten von Prozessen relevant, die unterschiedlichen Bewegungen auf Graphen entsprechen: Suchprozesse (1.) und Diffusionsprozesse (2.).<sup>482</sup>

---

ausgelegt, möglichst ästhetisch oder realistisch zu sein. Vielmehr soll das Modell interessante und nicht vollkommen unrealistische Strukturen produzieren, anhand derer die Methoden der juristischen Netzwerkforschung illustriert werden können.

<sup>481</sup> Das entspräche dem generativen Paradigma der agentenbasierten Modellierung; dazu *Epstein*, *Complexity* 4 (1999), S. 41–60.

<sup>482</sup> Im Folgenden nicht behandelt werden damit insbesondere Transportprozesse, bei denen es darum geht, Transportgegenstände, die nicht an mehreren Orten gleichzeitig sein können oder sollen, von bestimmten Startpunkten zu vorgegebenen Endpunkten zu befördern. Transportprozesse laufen beispielsweise auf Straßennetzen ab (Objekte: je nach Betrachtungsebene Orte oder Kreuzungen; Beziehungen: Verkehrswege; Ladung: Verkehrsteilnehmer), ebenso im Internet (Objekte: Clients, Server und Router; Beziehungen: Netzwerkverbindungen [z.B. WLAN oder Glasfaserkabel]; Ladung: Datenpakete). Ihre Analyse ist vor allem für die Regulierung kritischer Infrastrukturen relevant, die unter anderem darauf abzielt, Versorgungssicherheit zu gewährleisten und die Existenz effektiver und effizienter Verfahren für den Umgang mit Störungen zu garantieren. Die Modellierung von Transportprozessen auf Netzwerken kann bei der Abschätzung helfen, welche regulatorischen Maßnahmen zur Erfüllung regulatorischer Intentionen möglicherweise geeignet sind.

### 1. Modelle für Suchprozesse

Im Rahmen von Suchprozessen versuchen Suchende (z.B. Juristen), Suchgegenstände (z.B. Informationen zu einer Rechtsfrage) zu finden. Werden Suchprozesse auf Netzwerken modelliert, so befinden sich die Suchgegenstände bei den einzelnen Objekten und die Suchenden, denen die Position der Suchgegenstände unbekannt ist, bewegen sich über das Netzwerk, um sie zu finden. Für die juristische Netzwerkforschung ist insbesondere die Suche nach Informationen interessant.

Die Suche nach Informationen ist in dieser Arbeit bereits mehrfach als Bewegung über einen juristischen Zitiergraphen aufgetaucht.<sup>483</sup> Doch Suchprozesse auf Netzwerken sind auch an anderen Stellen im juristischen Rechercheprozess relevant: Wer eine klassische Suchmaschine (z.B. Google) benutzt, bekommt Informationen präsentiert, die von *Crawlern* gesammelt worden sind.<sup>484</sup> Diese Crawler bewegen sich im *World Wide Web*, einem Informationsnetzwerk aus Webseiten<sup>485</sup> (Objekte) und Hyperlinks (asymmetrische Beziehungen),<sup>486</sup> um möglichst viele Ressourcen zu finden, die für die Nutzer der Suchmaschine relevant sein könnten. Dabei lesen sie aus neu entdeckten Seiten systematisch die Hyperlinks aus, um diese später nachzuverfolgen. Da nicht jede Webseite von anderen Webseiten aus erreicht werden kann, finden die Crawler immer nur einen Teil der Webseiten, die über das *World Wide Web* abrufbar sind.<sup>487</sup> Es ist die Position der einzelnen Webseiten in dem Netzwerk, das sich aus allen *gefundenen* Webseiten zusammensetzt, welche anschließend (z.B. über den PageRank) in das Ranking der Suchergebnisse einfließt, die auch ein Jurist präsentiert bekommt, wenn er eine Web-Suchmaschine nutzt.

Es wurde bereits darauf hingewiesen, dass sich der PageRank auch als Bewegung eines Zufallssurfers über einen gerichteten Graphen (bzw. ein Netzwerk

---

<sup>483</sup> Beispiel: Abb. 3.18 (S. 128).

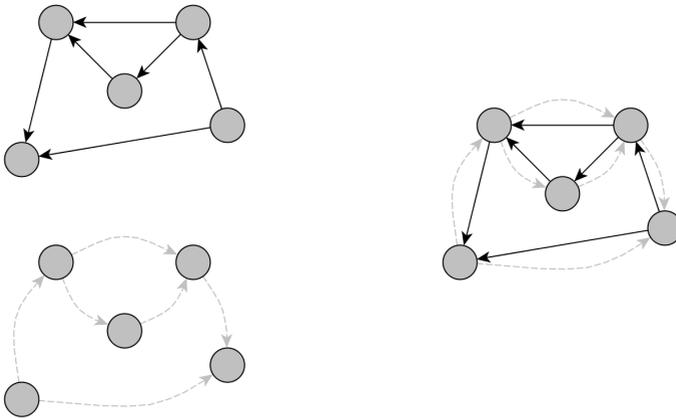
<sup>484</sup> Für einen systematischen Literaturüberblick zu den hierbei eingesetzten Technologien siehe Kumar, Bhatia und Rattan, WIREs Data Mining and Knowledge Discovery 7 (2017), S. 1–45; speziell zu Google Brin und Page, Computer Networks and ISDN Systems 30 (1998), S. 107–117. Wie Googles Web Crawler heute arbeiten, hat das Unternehmen in allgemeinverständlicher Sprache aufbereitet unter [www.google.com/search/howsearchworks/crawling-indexing/](http://www.google.com/search/howsearchworks/crawling-indexing/).

<sup>485</sup> Zur Erinnerung: Eine Webseite (*web page*) ist ein einzelnes Dokument, das über das *World Wide Web* abrufbar und anhand seiner URL eindeutig identifizierbar ist (z.B. [www.bundesgerichtshof.de/DE/DasGericht/dasGericht\\_node.html](http://www.bundesgerichtshof.de/DE/DasGericht/dasGericht_node.html)). Eine Website (*website*) ist eine Menge von Webseiten, die unter einer Domäne (z.B. [www.bundesgerichtshof.de](http://www.bundesgerichtshof.de)) abrufbar sind.

<sup>486</sup> Zur Struktur des *World Wide Web* siehe etwa Kleinberg u.a., The Web as a Graph: Measurements, Models, and Methods, in: Asano u.a. (Hrsg.), Computing and Combinatorics, 1999, S. 1–17; Broder u.a., Computer Networks 33 (2000), S. 309–320; Meusel u.a., The Journal of Web Science 1 (2015), S. 33–47.

<sup>487</sup> Zum Vorstehenden insgesamt Kumar, Bhatia und Rattan, WIREs Data Mining and Knowledge Discovery 7 (2017), S. 1 (1–2).

mit asymmetrischen Beziehungen) interpretieren lässt.<sup>488</sup> Nutzt ein Jurist zur Recherche von Fallrecht eine juristische Datenbank, so kann er sich aufgrund der vom Anbieter hinzugefügten Annotationen in der Regel in beide Richtungen über Zitate bewegen – so, als ob der Zitiergraph ungerichtet wäre. Ein diesen Recherchemöglichkeiten entsprechendes Modell eines juristischen Surfers muss den Zufallssurfer daher über einen Graphen surfen lassen, bei dem für jeden Pfeil in die eine Richtung (jedes Zitat) ein Pfeil in die andere Richtung (Annotation) hinzugefügt wurde.<sup>489</sup> Denn der juristische Surfer bewegt sich bei der Suche nach Rechtsprechung über eine Kombination aus Zitiergraphen und Annotationsgraphen, die hier als *Recherchegraph* bezeichnet wird (Abb. 3.44, S. 213). Für ein möglichst realistisches Netzwerkmodell der juristischen Datenbankrecherche wären außerdem alle dort über Annotationen oder Links im Text verfügbaren Quellen, also neben Gerichtsentscheidungen insbesondere Gesetzestexte, Verwaltungsvorschriften und Literaturbeiträge zu berücksichtigen.



(a) Zitiergraph und Annotationsgraph

(b) Recherchegraph

Abbildung 3.44: Konstruktion des Recherchegraphen aus Zitaten und Annotationen<sup>490</sup>

Eine weitgehend ungeklärte Frage ist, inwieweit das Rechercheverhalten von Juristen in juristischen Datenbanken systematisch vom Modell des Zufallssurfers abweicht. Denn wo es regelmäßig auf eine möglichst umfassende Kenntnis der Rechtslage ankommt, da muss auch die Quellenarbeit umfassend sein – so könnte

<sup>488</sup> Dazu oben S. 135–139.

<sup>489</sup> Man könnte also den PageRank für Entscheidungen in einem Rechtsprechungszitiernetzwerk auf einem um entsprechende Pfeile ergänzten Zitiergraphen berechnen.

<sup>490</sup> Die Abbildung geht davon aus, dass die Annotationen fehlerfrei sind, dass also zu jeder Entscheidung vollständig und ausschließlich diejenigen Entscheidungen unter „Diese Entscheidung wird zitiert“ (o.Ä.) vermerkt werden, die diese Entscheidung auch tatsächlich zitieren. Wahrscheinlich weicht die Realität von dieser Annahme ab.

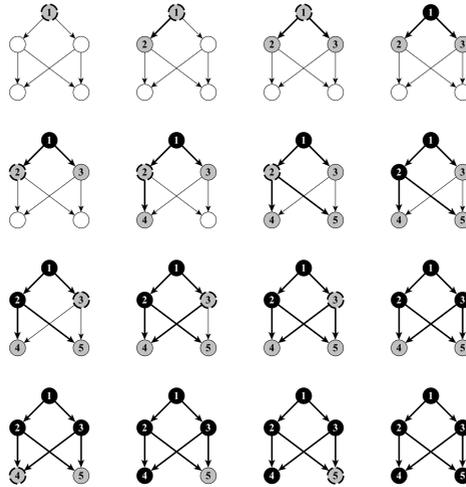
man zumindest vermuten. „Umfassende Quellenarbeit“ in Netzwerkstrukturen lässt sich algorithmisch als systematische Erkundung von Graphen formulieren. Diese Erkundung kann zuerst in die Breite oder zuerst in die Tiefe gehen. In beiden Fällen führt man Buch über die bereits gefundenen Quellen und fügt nach und nach neu entdeckte Quellen hinzu. Bei der *Breitensuche* (*breadth first search*, BFS) werden die gefundenen Quellen in der Reihenfolge auf weitere Quellen gesichtet, in der sie gefunden wurden; bei der *Tiefensuche* (*depth first search*, DFS) verhält es sich genau umgekehrt (Abb. 3.45, S. 215).<sup>491</sup> Für die Breitensuche bilden die noch nicht ausgewerteten Quellen eine klassische Warteschlange (*queue*), die nach dem „First In, First Out“-Prinzip (FIFO) abgearbeitet wird (Alltagsbeispiel: Supermarktkasse). Bei der Tiefensuche hingegen bilden die noch nicht behandelten Quellen einen Stapel (*stack*), der nach dem „Last In, First Out“-Prinzip (LIFO) abgearbeitet wird (Alltagsbeispiel: Tellerstapel).

Welcher Algorithmus ist näher am Suchverhalten von Juristen? Öffnen sie regelmäßig mehrere Fenster oder Tabs im Browser, um mehrere Links aus einem Dokument im Stile einer Breitensuche nachzuverfolgen? Oder bewegen sie sich erst durch eine Reihe von Quellen, um danach im Browser die Schaltfläche für „Zurück“ zu klicken und durch eine Reihe weiterer Klicks neue Dokumente zu erschließen? Weist das Rechercheverhalten allgemein so wenig Muster auf, dass es doch wie das Verhalten eines Websurfers modelliert werden sollte – also als Zufallslauf? Oder gibt es vielleicht Recherchemuster, die einem Modell eher entsprechen als den anderen und denen sich bestimmte Nutzertypen oder Suchanfragen zuordnen lassen? Und schließlich: Nutzen Juristen überhaupt die Netzwerkstruktur der Dokumentmenge, in der sie recherchieren? Solche Fragen wird die juristische Netzwerkforschung nur in Kooperation mit den Anbietern juristischer Datenbanken beantworten können.

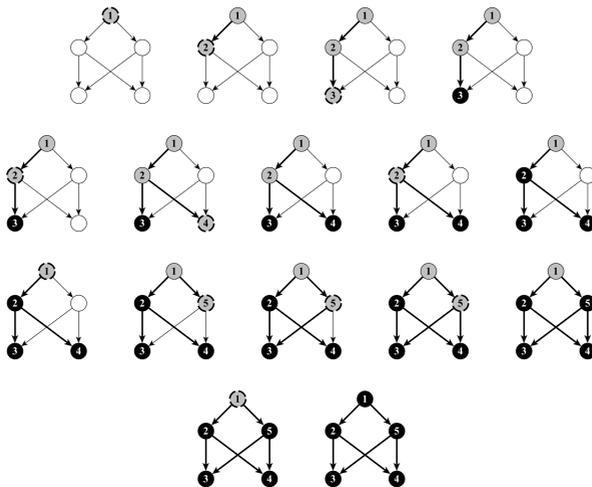
Zwar ist nicht zu erwarten, dass ein Jurist auch nur einen Bruchteil des gesamten Recherchegraphen vollständig erkundet. Denn nicht alle Dokumente, die von einem Dokument aus verlinkt sind, das vom Recherchierenden in einem konkreten Fall als interessant eingestuft wird, sind für dessen Suchanfrage ebenso interessant (etwa der Link zu § 242 BGB). Sollte man aber feststellen, dass Juristen ihre Recherche im Wesentlichen über das Suchfeld und die Liste der Suchergebnisse durchführen, so könnte dies auf grundsätzliche Probleme deuten, deren Ursache nicht nur in der Sphäre der Datenbanknutzer, sondern auch in jener der Datenbankanbieter liegen könnte: Stellt beispielsweise ein Nutzer regelmäßig viele neue, jeweils leicht modifizierte Anfragen hintereinander oder durchforstet er viele Seiten von Suchergebnissen, ohne dass er einen der angezeigten Treffer sichtet und sich von dort aus zu anderen, als potenziell relevant eingestuften Quellen weiterbewegt, so fällt es ihm offenbar schwer, durch die Interaktion mit der Datenbank Zugang zu denjenigen Dokumenten zu bekommen, die für ihn wichtig sind. Könnte man dem Nutzer helfen, indem man ihm visuelle Repräsentationen

<sup>491</sup> Allgemein zur Breitensuche *Cormen u.a.*, Introduction to Algorithms, 2009, S. 594–601; zur Tiefensuche a.a.O. S. 603–610.

von Teilen des Recherchegraphen zur Verfügung stellt, damit es ihm leichter fällt, darin zu navigieren, und er den Kontext eines Dokuments schneller erfassen kann? Inwiefern würde sich sein Suchverhalten durch Bereitstellung solcher Recherchehilfen verändern?



(a) Breitensuche



(b) Tiefensuche

Abbildung 3.45: Möglichkeiten zur systematischen Erkundung eines Graphen<sup>492</sup>

Wie Juristen die ihnen zur Verfügung stehenden Informationsnetzwerke praktisch nutzen, ist aus rechtswissenschaftlicher Perspektive auch deshalb relevant, weil sehr wahrscheinlich ist, dass der Rechercheprozess der Juristen erheblichen Einfluss auf den juristischen Diskurs und damit auch auf die Rechtsentwicklung hat. Was in den Suchergebnissen der gängigen Datenbanken weiter oben oder auf einer der vorderen Seiten in der Trefferdarstellung angezeigt wird, wird eher wahrgenommen. Was eher wahrgenommen wird, wird eher geklickt, eher zitiert und dadurch in späteren Suchen noch leichter auffindbar, sei es als Ergebnis konkreter Suchen oder durch Exploration des Recherchegraphen. Das Informationsnetzwerk, auf dem sich Juristen bewegen, und die Art und Weise, wie sie sich darauf bewegen, beeinflussen sich gegenseitig (*Co-Evolution*); beide gemeinsam wirken darauf ein, was diskutiert und vor allem auch judiziert wird. Macht dies juristische Datenbanken nicht zu einem naheliegenden Gegenstand für rechtswissenschaftliche Grundlagenforschung zu den normativen Wirkungen von Technologie? Müssten Juristen nicht hier zuerst über Regulierungsbedarf nachdenken?

## 2. Modelle für Diffusionsprozesse

Diffusionsprozesse sind Vorgänge, bei denen sich ein Diffusionsgegenstand mit der Zeit ausbreitet. Worin die Ausbreitung erfolgt (Diffusionsmedium) und wie sie vonstatten geht (Diffusionsmodus), hängt vom konkret betrachteten Diffusionsprozess ab. Für die juristische Netzwerkforschung sind unter anderem Prozesse der Ausbreitung von Krankheitserregern und Prozesse der Ausbreitung von Informationen interessant. Diese können ähnlich modelliert werden, repräsentieren aus regulatorischer Perspektive aber zwei Klassen von Diffusionsgegenständen: Krankheitserreger stehen für Diffusionsgegenstände, deren Ausbreitung in der Regel unerwünscht ist – hierunter fallen Pathogene für viele biologische Organismen (insbesondere Menschen) ebenso wie solche für technische Geräte (z.B.

---

<sup>492</sup> Die Punkte im Graphen haben drei Zustände: unentdeckt (weiß); entdeckt, aber nicht abgearbeitet (grau); und abgearbeitet (schwarz). Die Pfeile haben zwei Zustände: abgearbeitet (dick) und nicht abgearbeitet (dünn). Die Zahlen in den Punkten geben die Reihenfolge an, in der die Punkte entdeckt worden sind. Der Punkt, der gerade bearbeitet wird, ist mit einer gestrichelten Linie umrandet. Bei der Breitensuche wird von den entdeckten, aber nicht abgearbeiteten Punkten immer der Punkt mit der niedrigsten Zahl bearbeitet; bei der Tiefensuche wird immer der Punkt mit der höchsten Zahl bearbeitet. In jedem Bearbeitungsschritt wird ein ausgehender Pfeil erkundet, der noch nicht als abgearbeitet markiert ist (in welcher Reihenfolge die Pfeile erkundet werden, hängt von der Implementierung ab). Falls der Endpunkt eines solchen Pfeils noch nicht als entdeckt markiert ist, wird er mit einer entsprechenden Markierung (entdeckt, aber nicht abgearbeitet) versehen und zur Datenstruktur hinzugefügt, in der die „wartenden“ Punkte festgehalten werden. Ein Punkt wird als abgearbeitet markiert, wenn er zur Bearbeitung an der Reihe ist und alle von ihm ausgehenden Pfeile als abgearbeitet markiert sind. Breitensuche und Tiefensuche sind hier auf gerichteten Graphen dargestellt, um die Parallele zum Rechercheprozess zu verdeutlichen. Sie funktionieren ebenso auf ungerichteten Graphen – mit dem Unterschied, dass eine Linie von beiden angrenzenden Punkten aus erkundet werden kann und nicht nur von einem Startpunkt.

Computerviren oder -würmer). Informationen hingegen stehen für Diffusionsgegenstände, deren Ausbreitung je nach Kontext erwünscht oder unerwünscht sein kann – hierzu zählen auch Ideen und Innovationen, bei denen weder abstrakt noch objektiv beurteilt werden kann, ob ihre Diffusion erstrebenswert ist.<sup>493</sup>

Bei der Modellierung von Diffusionsprozessen auf Netzwerken versucht man zu beschreiben, wie sich ein Diffusionsgegenstand (z.B. ein Krankheitserreger oder eine Information) mit der Zeit über die einzelnen Beziehungen (z.B. soziale Interaktionen) ausbreitet und dabei nach und nach einzelne Objekte (z.B. Individuen in einer Population) erreicht. Ein Objekt kann sich dann in einem von mehreren Zuständen befinden, die als *Kompartments (compartments)* bezeichnet werden.<sup>494</sup> Im einfachsten Modell ist das Objekt (in der Terminologie der Krankheitserreger) entweder infizierbar (*susceptible*) oder infiziert (*infected*) – dieses Modell wird daher *SI-Modell* genannt.<sup>495</sup> Wie sich ein Diffusionsgegenstand ausbreitet, hängt insbesondere von seiner Ausbreitungsrate (*spreading rate*) und von der Netzwerkstruktur, aber auch von (unter Umständen im Modell nicht berücksichtigten) Eigenschaften der Objekte und Beziehungen im Netzwerk ab. Für einige simple Modelle von Diffusionsprozessen auf einfach beschreibbaren Netzwerkstrukturen kann man mathematische Aussagen über das Ausbreitungsverhalten des Diffusionsgegenstands herleiten;<sup>496</sup> kleinteiligere Modelle lassen sich oft nur anhand von Simulationen studieren. In diesen Simulationen kann man beobachten, wie sich der Zustand der einzelnen Objekte mit der (als diskrete Schritte dargestellten) Zeit verändert (Abb. 3.46, S. 217).

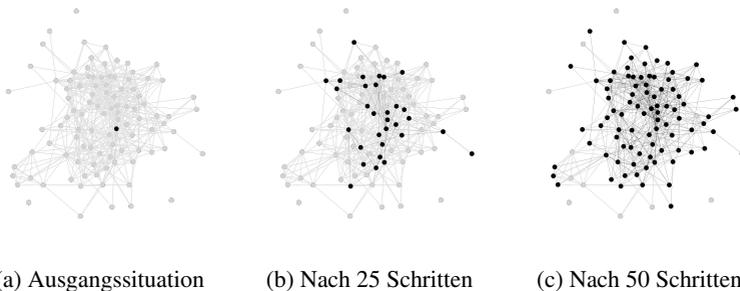


Abbildung 3.46: Ausbreitung einer Information im Unternehmensgraphen<sup>497</sup>

<sup>493</sup> Grundlegend zur Diffusion von Innovationen Rogers, *Diffusion of Innovations*, 2003 (Erstauflage 1962); im juristischen Kontext früh aufgegriffen von Walker, *The American Political Science Review* 63 (1969), S. 880 (883).

<sup>494</sup> Man spricht daher auch von Kompartimentmodellen (*compartmental models*); dazu etwa Brauer, *Compartmental Models in Epidemiology*, in: Brauer, van den Driessche und Wu (Hrsg.), *Mathematical Epidemiology*, 2008, S. 19–79.

<sup>495</sup> Newman, *Networks*, 2010, S. 628–631.

<sup>496</sup> Dazu etwa Barrat, *Barthélemy und Vespignani*, *Dynamical Processes on Complex Networks*, 2008, S. 189–195.

Da die Simulationen stochastische Elemente enthalten, muss man mehrere Simulationen gemeinsam betrachten, wenn man etwas über das erwartbare Ausbreitungsverhalten im betrachteten Netzwerk lernen will (Abb. 3.47, S. 219). Werden die Simulationen auf Graphen durchgeführt, die Teil eines Ensembles sind, so sind sogar für mehrere Graphen jeweils mehrere Simulationen zu starten und die Ergebnisse zusammenzufassen, bevor für das gesamte Ensemble Aussagen zum Verhalten von Diffusionsprozessen getroffen werden können. Man kann dabei nicht nur nach der Anzahl oder dem Anteil der Objekte fragen, die zu einem bestimmten Zeitpunkt infiziert sind, sondern auch untersuchen, wie das Ausbreitungsverhalten mit topologischen Merkmalen des Netzwerks (etwa der Schiefe seiner Gradverteilung oder der Existenz von Gradkorrelationen) zusammenhängt. Je nach Anwendungsfall wird man außerdem mehr als zwei Objektzustände unterscheiden oder weitere Zustandsübergänge zulassen wollen: Manche Krankheiten enden entweder tödlich oder verleihen Immunität (zusätzlicher Zustand *removed*, z.B. im *SIR-Modell: susceptible-infected-removed*); andere können bei infizierten Individuen nach einer Heilung wieder auftreten (Zustandsübergang auch von *infected* zu *susceptible*, z.B. im *SIS-Modell: susceptible-infected-susceptible*); wieder andere haben eine Inkubationszeit (zusätzlicher Zustand *exposed*, z.B. im *SEIR-Modell: susceptible-exposed-infected-removed*).<sup>498</sup> Ferner lässt sich die Wahrscheinlichkeit, dass der Diffusionsgegenstand über eine bestimmte Beziehung übertragen wird, nicht nur von der Ausbreitungsrate, sondern auch von den Eigenschaften einzelner Netzwerkelemente (z.B. einer Beziehungsintensität, die als Gewicht einer Linie modelliert ist) abhängig machen.

<sup>497</sup> Abbildungen 3.46 (a), (b) und (c) (S. 217) zeigen die Ausbreitung einer Information im Unternehmensgraphen; zum Layout bereits Fn. 425 (S. 196). Die Punkte der Unternehmen, denen die Information bekannt ist, sind schwarz markiert, ebenso die Linien, über welche die Ausbreitung der Information im Modell stattgefunden hat. Zu Beginn kennt nur ein Unternehmen die Information, die im Modell in jedem Schritt (*time step*) mit einer Wahrscheinlichkeit von 2 % über die einzelnen Personalverflechtungen zwischen den Unternehmen übertragen wird. Nach 25 Schritten ist die Information insgesamt 31 Unternehmen bekannt; nach 50 Schritten hat sie bereits 88 der 100 Unternehmen erreicht.

<sup>498</sup> Zu den verschiedenen Kompartimentmodellen *Brauer*, *Compartmental Models in Epidemiology*, in: *Brauer, van den Driessche und Wu* (Hrsg.), *Mathematical Epidemiology*, 2008, S. 19–79; zusammenfassend für die Modellierung auf Netzwerken *Kiss, Miller und Simon*, *Mathematics of Epidemics on Networks*, 2017, S. 4–12; instruktiv außerdem *Keeling und Eames*, *Journal of the Royal Society Interface* 2 (2005), S. 295–307.

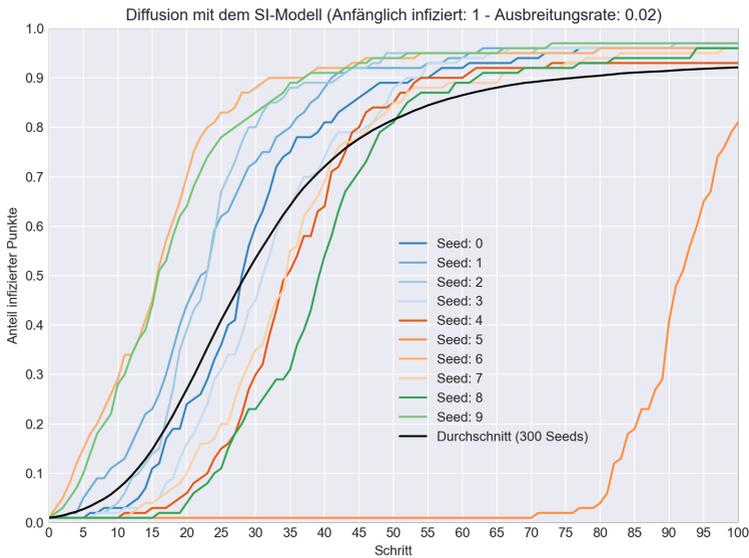


Abbildung 3.47: Ausbreitung eines Diffusionsgegenstands (z.B. einer Information) im Unternehmensnetzwerk<sup>499</sup>

Welchen Platz die skizzierten Modelle in der juristischen Netzwerkforschung haben, hängt davon ab, welches Diffusionsphänomen betrachtet wird. Auf juristischen Zitiernetzwerken und sozialen Netzwerken aus juristischen Akteuren lässt sich die Ausbreitung von Ideen modellieren, um die sie treibenden Kräfte besser zu verstehen.<sup>500</sup> Unter welchen Umständen wird die an einer Universität gelehrt Mindermeinung zur Mehrheitsmeinung?<sup>501</sup> Wie gelangt eine Rechtsfigur von einer Nation in die andere?<sup>502</sup> Wie breiten sich Versicherungen für Führungskräfte (D&O-Versicherungen) in Unternehmen aus?<sup>503</sup> Zur Beantwortung dieser Fra-

<sup>499</sup> Abbildung 3.47 (S. 219) zeigt, wie sich ein Diffusionsgegenstand im Unternehmensgraphen mit einer Rate von 2 % ausbreitet, wenn zu Anfang genau ein Unternehmen „infiziert“ war, für verschiedene Instanzen eines einfachen SI-Modells sowie den Durchschnitt aus 300 Modellen.

<sup>500</sup> Einen Überblick über Modelle zur Diffusion von Ideen geben *Vitanov und Ausloos*, Knowledge Epidemics and Population Dynamics Models for Describing Idea Diffusion, in: *Scharnhorst, Börner und van den Besselaar* (Hrsg.), Models of Science Dynamics, 2012, S. 69–125.

<sup>501</sup> Ansätze zur Modellierung von „Intellectual Diffusion Across the American Legal Academy“ finden sich bei *Katz u.a.*, Journal of Legal Education 61 (2011), S. 76 (96–100).

<sup>502</sup> Diese Frage wird von Juristen typischerweise unter dem Stichwort „Legal Transplants“ diskutiert; dazu insbesondere *Watson*, Legal Transplants, 1993; aus der jüngeren Literatur (mit sehr unterschiedlichen Herangehensweisen) beispielsweise *Garoupa und Ogus*, The Journal of Legal Studies 35 (2006), S. 339–363; *Langer*, The American Journal of Comparative Law 55 (2007), S. 617–676; *Spamann*, BYU Law Review 2009, S. 1813–1877.

<sup>503</sup> Ähnliche Fragestellungen werden außerhalb der Rechtswissenschaft bereits für juristisch interessante Phänomene untersucht, allerdings ohne die explizite Modellierung von Diffu-

gen muss nicht nur ein geeignetes Modell für den Diffusionsprozess gefunden werden, in das zahlreiche Annahmen einfließen. Vielmehr ist auch die Definition des Netzwerks problematisch, auf dem der Diffusionsprozess ablaufen soll, und die Schwierigkeiten von Netzwerkdefinition und Prozessdefinition hängen miteinander zusammen.

Versuchte man beispielsweise, die Ausbreitung juristischer Innovationen (z.B. eines neu geschaffenen Grundrechts auf Gewährleistung der Vertraulichkeit und Integrität informationstechnischer Systeme)<sup>504</sup> auf einem gewöhnlichen Rechtsprechungsziernetzwerk zu modellieren, so erhielt man mit einem Modell des Diffusionsprozesses, das von gleichen Übertragungswahrscheinlichkeiten für alle Zitierbeziehungen ausgeht, vermutlich kaum sinnvolle Resultate – der Anteil der nicht innovationserheblichen Zitate (z.B. routinemäßige Zitate von BVerfGE 18, 85 – Spezifisches Verfassungsrecht)<sup>505</sup> an allen Zitaten scheint dafür zu groß.<sup>506</sup> Hingegen könnte ein ähnliches Vorgehen bei Patentziernetzwerken durchaus Erfolg versprechen, da hier rechtliche Vorgaben und ökonomische Anreize dafür sorgen, dass ein relativ enger Sachzusammenhang zwischen Zitierendem und Zitiertem besteht.<sup>507</sup> Für manche Prozesse der Ausbreitung von Ideen im Recht sind reine Diffusionsmodelle vielleicht sogar insgesamt untauglich: Welchen Einfluss etwa die Änderung der Rechtsprechung eines Gerichts zu einer bestimmten Sachfrage auf die Rechtsprechung anderer Gerichte hat, hängt nicht nur von deren Fallaufkommen ab (ein Gericht kann seine Rechtsprechung nur ändern, wenn ihm ein geeigneter Fall zur Entscheidung vorgelegt wird), sondern auch von institutio-

---

sionsprozessen auf Netzwerken. Beispiele: *Davis*, *Administrative Science Quarterly* 36 (1991), S. 583–613 (*Poison Pills*); *Bizjak, Lemmon und Whitby*, *The Review of Financial Studies* 22 (2009), S. 4821–4847 (*Option Backdating*); *Fiss, Kennedy und Davis*, *Organization Science* 23 (2012), S. 1077–1099 (*Golden Parachutes*). Zur thematisch einschlägigen Literatur aus der Netzwerkforschung im Recht bereits oben S. 22–28.

<sup>504</sup> Grundlegend BVerfGE 120, 274 – Online-Durchsuchungen; aufgegriffen etwa in BVerfGE 124, 43 – Beschlagnahme von E-Mails und BVerfGE 133, 277 – Antiterrordateigesetz.

<sup>505</sup> Zur Rolle von BVerfGE 18, 85 – Spezifisches Verfassungsrecht im Zitiernetzwerk des BVerfG unten S. 269–278.

<sup>506</sup> Die Diffusion würde in einem solchen Modell nicht über die als Pfeile dargestellten Zitate, sondern in entgegengesetzter Richtung erfolgen: von der zitierten Entscheidung als Träger der juristischen Innovation zur zitierenden Entscheidung, die sie aufgreift. So wird die Diffusion der juristischen Innovation von BVerfGE 120, 274 – Online-Durchsuchungen zu BVerfGE 124, 43 – Beschlagnahme von E-Mails und BVerfGE 133, 277 – Antiterrordateigesetz im Beispiel des Grundrechts auf Gewährleistung der Vertraulichkeit und Integrität informationstechnischer Systeme (Fn. 504, S. 220) gerade dadurch angezeigt, dass die Entscheidungen BVerfGE 124, 43 und BVerfGE 133, 277 die Entscheidung BVerfGE 120, 274 zitieren.

<sup>507</sup> Zu Patentnetzwerken bereits oben S. 26–27. In vielen Rechtsordnungen ist vorgesehen, dass bei der Patentanmeldung andere Patente angegeben werden müssen, die für Erfindungen angemeldet wurden, welche die Neuheit der zu patentierenden Erfindung beeinträchtigen könnten. Der Patentanmelder hat kaum Anreize, viele andere Patente zu zitieren: Je mehr andere, nicht zwingend zu nennende Erfindungen er angibt, desto höher das wahrgenommene Risiko, dass seine Patentanmeldung scheidert.

nellen Vorgaben (beispielsweise der Gerichtshierarchie)<sup>508</sup>, deren Auswirkungen sich möglicherweise besser durch deterministische als durch stochastische Prozesse beschreiben lassen.

Geht es um die Ausbreitung von unerwünschten Diffusionsgegenständen – seien es Krankheiten, Computerviren oder sensible Daten –, so kann die juristische Netzwerkforschung sich an der Klärung der Frage beteiligen, mit welchen Mitteln Diffusion verhindert oder zumindest so weit eingedämmt werden kann, dass keine Epidemien ausbrechen (im wörtlichen oder übertragenen Sinne). Hier steht das Verhalten des Diffusionsprozesses infolge gezielter (z.B. regulatorischer) Interventionen im Mittelpunkt der Analyse. Legt man ein Diffusionsmodell zugrunde, bei dem eine Heilung infizierter Objekte grundsätzlich möglich ist, so können die Interventionen sowohl *präventiv* als auch *repressiv* sein; außerdem können sie bei den Objekten oder bei den Beziehungen ansetzen, also *objektadressiert* oder *beziehungsadressiert* sein.

Ein anschauliches Beispiel bietet die Regulierung von Infektionskrankheiten. Die klassische medikamentöse Behandlung ist als Intervention objektadressiert und repressiv (es geht um die Heilung bereits Infizierter); Impfpflichten sind objektadressiert und präventiv.<sup>509</sup> Eine Quarantänepflicht ist beziehungsadressiert und repressiv (bestehende Beziehungen, die potenzielle Übertragungswege darstellen, werden bewusst gekappt);<sup>510</sup> die Strafbarkeit der Infizierung eines anderen mit einer Infektionskrankheit als Körperverletzung lässt sich als beziehungsadressiert und präventiv auffassen (das Entstehen potenzieller Übertragungswege soll weniger wahrscheinlich gemacht werden).<sup>511</sup> Ersetzt man Menschen durch Finanzinstitute, soziale Interaktionen durch finanzielle Verflechtungen und Infektionskrankheiten durch Finanzierungsengpässe, so gelangt man zu Diffusionsprozessen von finanziellem Stress, die auch in der ökonomischen und finanzwissenschaftlichen Literatur verstärkt auf Netzwerken modelliert und im Dunstkreis des Schlagworts „Systemisches Risiko“ diskutiert werden.<sup>512</sup> Auch hier stellt sich die Frage, mit welchen Regulierungswerkzeugen (präventiv oder repressiv, objektadressiert oder

<sup>508</sup> Die Gerichtshierarchie bewirkt in Deutschland zwar keine rechtliche, wohl aber eine faktische Bindung der unteren Gerichte an die Rechtsprechung höherer Gerichte. Der Umfang der faktischen Bindungswirkung ließe sich vielleicht anhand von Zitierdaten abschätzen.

<sup>509</sup> In Deutschland ist eine allgemeine Impfpflicht nach derzeitiger Rechtslage nicht vorgesehen und ihre etwaige Vereinbarkeit mit dem Recht auf körperliche Unversehrtheit (Art. 2 II GG) ist umstritten; dazu *Schaks und Kahnert*, MedR 33 (2015), S. 860–866 (mit Zusammenfassung der Rechtslage auf S. 862–863).

<sup>510</sup> In Deutschland ist die Quarantänepflicht in § 30 Infektionsschutzgesetz (IfSG) geregelt.

<sup>511</sup> Speziell zur Strafbarkeit der Infektion eines anderen mit dem HI-Virus *Eser und Sternberg-Lieben*, StGB § 223, in: *Eser* (Hrsg.), Schönke/Schröder Strafgesetzbuch, 2014 (Rn. 7).

<sup>512</sup> Beispiele: *Iori u.a.*, Journal of Economic Dynamics & Control 32 (2008), S. 259–278; *Gai, Haldane und Kapadia*, Journal of Monetary Economics 58 (2011), S. 453–470; *Battiston u.a.*, Scientific Reports 2 (2012), S. 541-1–541-6; *Summer*, Annual Review of Financial Economics 5 (2013), S. 277–297; *Kenett und Havlin*, Mind & Society 14 (2015), S. 155–167.

beziehungsadressiert) die Systemstabilität im Regelfall und die schnelle Erholung im Störfall gewährleistet werden können.

Dass es für die Wirksamkeit von Interventionen nicht nur auf das *Wie* der Intervention (Interventionsmodus), sondern auch auf das *Wo* (Interventionsstelle) ankommt, zeigt sich deutlich, wenn man untersucht, wie sich Netzwerkstrukturen durch die Eliminierung einzelner Objekte oder Beziehungen verändern.<sup>513</sup> Diese Eliminierung gleicht den erwähnten Interventionen: Ein Mensch, der geimpft ist, fällt als infizierender Kontakt eines Nicht-Geimpften aus;<sup>514</sup> ein Finanzinstitut, das mit einem anderen nicht finanziell verflochten ist, kann nicht direkt finanziellen Stress auf dieses übertragen. Doch wie sollten die Individuen für gezielte Impfkationen ausgewählt werden – zufällig oder systematisch nach Kriterien, die an ihre Position im Netzwerk anknüpfen? Und welche Verflechtungsbeziehungen zwischen Finanzinstituten sollten unterbunden (oder auch befördert) werden? Sollte man alle Finanzinstitute und Verflechtungsbeziehungen gleich behandeln – oder bestünde eine materielle Gleichbehandlung womöglich darin, systematisch nach topologischen Kriterien zu differenzieren?

Bei der Suche nach Antworten zu diesen Fragen helfen Modelle von Störprozessen auf Netzwerken. Durch sie lässt sich unter anderem zeigen, dass die Anfälligkeit eines Netzwerks für Störungen (bzw. umgekehrt seine Robustheit gegenüber Störungen) wesentlich von zwei Faktoren abhängt: der Art der Störung und der Netzwerktopologie.<sup>515</sup> Netzwerke mit homogener Gradverteilung, die unter anderem durch Zufallsgraphen modelliert werden können, sind anfälliger für den zufälligen Ausfall einzelner Objekte oder Beziehungen (im Impfbeispiel etwa die zufällige Auswahl einzelner Individuen für die Immunisierung) als Netzwerke mit heterogener, sehr schiefer Gradverteilung, für deren Analyse beispielsweise Kumulations- oder Konfigurationsmodelle herangezogen werden können.<sup>516</sup> Demgegenüber reagieren Netzwerke mit sehr schiefer Gradverteilung viel sensibler auf die gezielte Eliminierung einzelner Objekte oder Beziehungen anhand von topologischen Kriterien (beispielsweise Grad oder Zwischenzentralität). Schon die Entfernung weniger stark vernetzter Objekte kann hier zum Zerfall des gesamten Netzwerks führen, während ihm der zufällige Ausfall einzelner Netzwerkelemente kaum etwas anhaben kann (denn dass zufällig einer der *hubs*

---

<sup>513</sup> Albert, Jeong und Barabási, *Nature* 406 (2000), S. 378–382; Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 116–120; Newman, *Networks*, 2010, S. 594–615.

<sup>514</sup> So jedenfalls die vereinfachende Annahme vieler Modelle; die medizinische Situation ist komplizierter.

<sup>515</sup> Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 116–120.

<sup>516</sup> Albert, Jeong und Barabási, *Nature* 406 (2000), S. 378–382; Barrat, Barthélemy und Vespignani, *Dynamical Processes on Complex Networks*, 2008, S. 118. Zu den verschiedenen Netzwerkmodellen oben S. 195–211.

ausfällt, ist bei vielen *spokes* sehr unwahrscheinlich).<sup>517</sup> Netzwerke mit schiefer Gradverteilung sind also robust, aber zugleich fragil (*robust yet fragile*).<sup>518</sup>

Aus juristischer Perspektive bedeutet dieser Befund zunächst nur, dass regulatorische Maßnahmen zur Bekämpfung unerwünschter Diffusion die Struktur des Netzwerks berücksichtigen sollten, über das diese Diffusion mutmaßlich stattfindet. Welche juristischen Schlussfolgerungen daraus zu ziehen sind, hängt vom konkreten Diffusionsphänomen ab, ließe sich aber rechtswissenschaftlich erforschen.<sup>519</sup> Die größte Schwierigkeit liegt auch hier wieder in einer angemessenen Netzwerkdefinition. Denn Infektionskrankheiten verbreiten sich nicht nur durch soziale Kontakte, sondern auch aufgrund geographischer Nähe (z.B. über Luft und Wasser) oder durch Kontakte mit Tieren oder Gegenständen. Finanzinstitute sind nicht nur untereinander, sondern mit vielen anderen Unternehmen über viele verschiedene Beziehungen verbunden, deren Einfluss auf die Stabilität des Finanzsystems noch weitgehend ungeklärt ist. Ein Forscher, der in einem Diffusionsmodell für finanziellen Stress nur Eigen- und Fremdkapitalbeziehungen zwischen Finanzinstitutionen betrachtet, kommt zu ganz anderen Ergebnissen als ein Forscher, der auch andere Industriezweige und weitere Beziehungsarten einbezieht. Dies deutet bereits auf eine der größten Herausforderungen für die juristische Netzwerkforschung hin: die Vielfalt von Objekten und Beziehungen, die bei der Definition eines Netzwerks berücksichtigt werden können (Mehrschichtigkeit, hierzu sogleich).

---

<sup>517</sup> *Albert, Jeong und Barabási*, *Nature* 406 (2000), S. 378–382; zusammenfassend *Albert und Barabási*, *Reviews of Modern Physics* 74 (2002), S. 47 (86–90) sowie *Newman*, *Networks*, 2010, S. 609–615. Allgemeine Resultate für Ausfallprozesse auf Zufallsgraphen mit beliebiger Gradverteilung finden sich bei *Callaway u.a.*, *Physical Review Letters* 85 (2000), S. 5468–5471.

<sup>518</sup> Die Terminologie wird vor allem im Zusammenhang mit dem Internet verwendet; siehe *Carlson und Doyle*, *Physical Review E* 60 (1999), S. 1412–1427 (1426); *Carlson und Doyle*, *PNAS* 99 (2002), S. 2538–2545; *Doyle u.a.*, *PNAS* 102 (2005), S. 14497–14502; *Alderson und Doyle*, *IEEE Transactions on Systems, Man, and Cybernetics* 40 (2010), S. 839–852; für Rechtssysteme aufgegriffen von *Ruhl*, *Indiana Law Journal* 89 (2014), S. 559–603; *Ruhl und Katz*, *Iowa Law Review* 101 (2015), S. 191 (239–240). Bei näherem Hinsehen ist die Fragilität von Netzwerken mit schiefer Gradverteilung weniger dramatisch, als es auf den ersten Blick erscheinen mag; dazu *Zweig*, *Network Analysis Literacy*, 2016, S. 379–380.

<sup>519</sup> Einige Forscher leiten aus ihren Analyseergebnissen explizite Handlungsempfehlungen ab, die in die juristischen Überlegungen einbezogen werden können; speziell für Interventionen mit dem Ziel, die Ausbreitung von Krankheiten zu verhindern oder einzudämmen, beispielsweise *Dezső und Barabási*, *Physical Review E* 65 (2002), 055103(R)-1–055103(R)-4; *Pastor-Satorras und ben-Avraham*, *Physical Review E* 65 (2002), S. 036104-1–036104-8; *Cohen, Havlin und ben-Avraham*, *Physical Review Letters* 91 (2003), S. 247901-1–247901-4; *Holme*, *European Physics Letters* 68 (2004), S. 908–914; *Salathé und Jones*, *PLoS Computational Biology* 6 (2010), e1000736-1–e1000736-11.

## E. Herausforderung: Mehrschichtigkeit

Die bisherigen Ausführungen haben sich auf Netzwerke mit einer Art von Objekten und einer Art von Beziehungen konzentriert. Diese machen den Großteil der bisher untersuchten Gegenstände in der juristischen Netzwerkforschung aus; für sie sind auch die Visualisierungs-, Quantifizierungs- und Modellierungsmethoden am weitesten entwickelt. Entsprechend sind die Beispielnetzwerke aus Flatland konstruiert: Das Zitiernetzwerk der flatländischen Rechtsprechung besteht aus Gerichtsentscheidungen (Objekte) und den zwischen ihnen verlaufenden Zitaten (Beziehungen). Das Verflechtungsnetzwerk der flatländischen Unternehmen setzt sich zwar aus Unternehmen und Führungsfiguren (Objekte) sowie Mitgliedschaften in Kontrollgremien (Beziehungen) zusammen, doch sind hier fast ausschließlich die Projektionen auf eine Art von Objekten (Unternehmensnetzwerk und Personalnetzwerk) betrachtet worden.

Die meisten realen Phänomene lassen sich allerdings genauer erfassen, wenn man mehrere Arten von Objekten und Beziehungen in ihrer abstrakten Darstellung als Netzwerke zulässt: Gerichtsentscheidungen zitieren nicht nur andere Gerichtsentscheidungen, sondern auch Gesetzestexte oder rechtswissenschaftliche Publikationen; sie sind Teil eines Netzwerks juristischer Texte verschiedenster Gattungen, das unter anderem über Autorenbeziehungen mit einem sozialen Netzwerk aus Rechtswissenschaftlern, Rechtspraktikern und anderen juristischen Akteuren verbunden ist, das wiederum viele verschiedene Beziehungsarten enthält. Unternehmen sind nicht nur durch Personal-, sondern auch durch Kapitalverflechtungen miteinander verbunden; will man ihr Handeln verstehen, muss man womöglich auch Personal unterhalb der Führungsebene und dessen soziale Beziehungen, Lieferanten- und Kundenbeziehungen sowie rechtliche Beziehungen (insbesondere Vertragsbeziehungen) berücksichtigen. Die Abstraktionen, die zur Untersuchung der Gerichtslandschaft und der Unternehmenslandschaft in Flatland konstruiert wurden, sind demgegenüber – dem Staatsnamen entsprechend – flach.

Die Netzwerkforschung hat die Komplexität der sie interessierenden Phänomene lange Zeit maximal reduziert (oft bewusst, bisweilen aber auch unbewusst) – nicht zuletzt, um ihre Betrachtung überhaupt erst zu ermöglichen. Mittlerweile ist die Untersuchung von Netzwerken mit einer Art von Objekten und einer Art von Beziehungen allerdings so weit fortgeschritten und das Bewusstsein für die Relevanz von Zusammenhängen zwischen verschiedenen Arten von Netzwerkelementen für das Verhalten komplexer Systeme in der Realität so groß, dass man verstärkt nach Wegen sucht, heterogen zusammengesetzte Netzwerke zu analysieren. Diese Netzwerke werden als *mehrschichtige Netzwerke* (*multilayer networks*) bezeichnet.<sup>520</sup>

<sup>520</sup> Die Terminologie ist uneinheitlich; Überblicke mit zahlreichen Nachweisen bieten *Boccaletti u.a.*, *Physics Reports* 544 (2014), S. 1–122 und *Kivelä u.a.*, *Journal of Complex Networks* 2 (2014), S. 203–271; zum mathematischen Hintergrund *De Domenico u.a.*, *Physical Review X* 3 (2013), S. 041022-1–041022-15. In dieser Arbeit bilden *mehrschichtige Netzwerke* den

Die methodischen Herausforderungen der Analyse mehrschichtiger Netzwerke sind groß: Wie kann man Netzwerke mit mehreren Arten von Objekten und Beziehungen am besten visualisieren?<sup>521</sup> Was bedeuten Wichtigkeit und Zusammengehörigkeit in solchen Netzwerken – und wie lassen sich diese Netzwerke charakterisieren?<sup>522</sup> Wie modelliert man das Zustandekommen mehrschichtiger Netzwerke, wie die Prozesse, die auf ihnen ablaufen?<sup>523</sup> Mit diesen Fragen öffnet sich ein weites Feld, das viel methodisches Innovationspotenzial bietet, zumal es in der juristischen Netzwerkforschung bisher kaum Beachtung gefunden hat.<sup>524</sup>

Die Berücksichtigung von Mehrschichtigkeit bei der Untersuchung juristischer Netzwerke ist auf lange Sicht eine Voraussetzung dafür, dass die juristische Netzwerkforschung mehr zum rechtswissenschaftlichen Diskurs beitragen kann als schöne Bilder (Visualisierung), harte Zahlen (Quantifizierung) und interessante Gedankenspiele (Modellierung). Noch sind allerdings nicht einmal die Grenzen der Analyse einfacher juristischer Netzwerke mithilfe von Darstellungen als schlichte, gerichtete, gewichtete oder Multigraphen ausgelotet; auch die Analyse ihrer Veränderungen über die Zeit steht noch am Anfang.<sup>525</sup> Kurz- und mittelfris-

---

Oberbegriff für alle Netzwerke mit mehreren Arten von Objekten oder Beziehungen; eine in der Literatur bisher besonders intensiv untersuchte Spezialisierung sind *multiplexe Netzwerke* (*multiplex networks*), d.h. Netzwerke mit einer Art von Objekten, zwischen denen mehrere Arten von Beziehungen bestehen können; speziell dazu etwa *Bianconi*, *Physical Review E* 87 (2013), S. 062806-1–062806-15; *Nicosia u.a.*, *Physical Review Letters* 111 (2013), S. 058701-1–058701-5; *Battiston, Nicosia und Latora*, *Physical Review E* 89 (2014), S. 032804-1–032804-14.

<sup>521</sup> Dazu beispielsweise *De Domenico, Porter und Arenas*, *Journal of Complex Networks* 3 (2015), S. 159–176; *Rossi und Magnani*, *Chaos, Solitons & Fractals* 72 (2015), S. 68–76; *Pišková, Šlaban und Šmuc*, *MultiNets: Web-Based Multilayer Network Visualization*, in: *Bifet u.a.* (Hrsg.), *Machine Learning and Knowledge Discovery in Databases*, 2015, S. 298–302.

<sup>522</sup> Dazu beispielsweise *Boccaletti u.a.*, *Physics Reports* 544 (2014), S. 1 (10–26); *Kiveliä u.a.*, *Journal of Complex Networks* 2 (2014), S. 203 (229–234, 237–240), jeweils m.w.N.. Speziell für multiplexe Netzwerke ferner *Battiston, Nicosia und Latora*, *Physical Review E* 89 (2014), S. 032804-1–032804-14; *Baxter u.a.*, *Physical Review E* 94 (2016), S. 062308-1–062308-8; *Vörös und Snijders*, *Social Networks* 49 (2017), S. 93–112.

<sup>523</sup> Dazu beispielsweise *Boccaletti u.a.*, *Physics Reports* 544 (2014), S. 1 (27–98); *Kiveliä u.a.*, *Journal of Complex Networks* 2 (2014), S. 203 (234–237, 240–251), jeweils m.w.N., zu Prozessen auf mehrschichtigen Netzwerken außerdem *Salehi u.a.*, *IEEE Transactions on Network Science and Engineering* 2 (2015), S. 65–83. Speziell für multiplexe Netzwerke ferner *Gómez u.a.*, *Physical Review Letters* 110 (2013), S. 028701-1–028701-5; *De Domenico u.a.*, *PNAS* 111 (2014), S. 8351–8356; *Nicosia u.a.*, *Physical Review Letters* 118 (2017), S. 138302-1–138302-6.

<sup>524</sup> Erste Erkundungen z.B. bei *Bommarito und Katz*, *Physica A* 389 (2010), S. 4195–4200; *Bommarito, Katz und Isaacs-See*, *Virginia Tax Review* 30 (2011), S. 523–557; *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327–349.

<sup>525</sup> Ansätze finden sich etwa bei *Cross, Smith und Tomarchio*, *Emory Law Journal* 57 (2008), S. 1227–1257; *Kim*, *Global Environmental Change* 23 (2013), S. 980–991; *Kim, Cho und Kim*, *Scientometrics* 98 (2014), S. 975–998; *Tarissan und Nollez-Goldbach*, *Temporal Properties of Legal Decision Networks: A Case Study from the International Criminal Court*, in: *Rotolo* (Hrsg.), *Legal Knowledge and Information Systems*, 2015, S. 111–120; *Whalen*, *Michigan State Law Review* 2016, S. 539 (561–565).

tig scheint es daher angebracht, den Fokus auf die systematische Erschließung der Erkenntnispotenziale zu legen, welche die Analyse juristischer Netzwerke mit einer Art von Objekten und einer Art von Beziehungen birgt, und dabei die Einschränkungen im Hinterkopf zu behalten, die mit dieser Vorgehensweise für die Interpretation der Analyseergebnisse einhergehen. Ohne eine solche „Aufholarbeit“ sind problembewusste Anwendung und zielgerichtete Weiterentwicklung komplizierterer Methoden kaum möglich. In diesem Sinne zieht die nachfolgende Fallstudie zum deutschen Recht eine klassische juristische Netzwerkdefinition heran, um einige der in diesem Kapitel vorgestellten methodischen Instrumente auf ihren Nutzen in der Praxis zu überprüfen: Untersuchungsgegenstand ist ein Zitiernetzwerk des Bundesverfassungsgerichts.

## Kapitel 4

# Praxis

In diesem Kapitel wird eine Fallstudie zur Rechtsprechung des Bundesverfassungsgerichts (BVerfG) durchgeführt, um die im dritten Kapitel eingeführten Methoden der juristischen Netzwerkforschung praktisch zu erproben. Im ersten Schritt werden Daten zur Untersuchung zusammengestellt. Bereits hier sind zahlreiche (Wertungs-)Entscheidungen zu treffen: Welche Daten sollen aus welchen Quellen gesammelt werden? In welchem Format sollen die Daten für die Zwecke der Analyse gebracht werden? Wie ist dabei vorzugehen? Diese und ähnliche Fragen behandelt Abschnitt A.

Im zweiten Schritt werden die zusammengestellten Daten analysiert. Die Analyse kann prinzipiell den Gegenstand (hier: die Rechtsprechung des BVerfG) oder die Methoden (hier: die Methoden der juristischen Netzwerkforschung) betonen; dabei kann sie Exploration, Explikation oder Extrapolation zum Ziel haben.<sup>1</sup> Hier stehen tendenziell die Methoden im Vordergrund, deren Praxistauglichkeit in diesem Kapitel getestet werden soll. Die Fallstudie widmet sich allein der Exploration, da der gewählte Gegenstand mit den gewählten Methoden bisher kaum untersucht worden ist. Ihren Kern bildet eine Reihe von Experimenten, deren Aufbau, Ergebnisse und Interpretationsmöglichkeiten in Abschnitt B. zusammengefasst sind.

Im dritten Schritt werden die Ergebnisse der Experimente vor dem Hintergrund des gewählten Datenzuschnitts evaluiert. Die Evaluation betrifft inhaltliche, methodische und praktische Erkenntnisse. Welche Einsichten in die Rechtsprechung des BVerfG konnten mithilfe der erprobten Methoden gewonnen werden? Welche methodischen Werkzeuge haben Resultate geliefert, die für eine oder mehrere Perspektiven der Rechtswissenschaft verwertbar sind, sodass sich eine über die bloße Erkundung hinausgehende Untersuchung lohnen könnte? Welche Schwierigkeiten sind bei der Sammlung, Analyse und Kommunikation der Daten aufgetreten? Diesen Fragestellungen widmet sich Abschnitt C.

---

<sup>1</sup> Zu den Untersuchungszielen der juristischen Netzwerkforschung bereits S. 62–71.

## A. Daten

Will man die Rechtsprechung des BVerfG mit den Methoden der juristischen Netzwerkforschung analysieren, so muss man den Untersuchungsgegenstand präzisieren: Was genau soll betrachtet werden? Angesichts der im zweiten Kapitel geschilderten Möglichkeiten, Netzwerke im Recht zu konstruieren,<sup>2</sup> drängen sich zuvor allerdings zwei weitere Fragen auf. Erstens: Warum sollte man gerade *Rechtsprechung* untersuchen? Und zweitens: Warum bietet es sich an, speziell *Rechtsprechung des BVerfG* zu analysieren?

Eine subjektive Antwort auf beide Fragen lautet: Weil man sich für die Rechtsprechung des BVerfG interessiert. Es gibt allerdings auch einige überindividuelle Gründe. Für eine Untersuchung von Rechtsprechung lässt sich anführen, dass man damit im internationalen Diskurs anschlussfähig ist, in dem Rechtsprechungsziernetzwerke seit geraumer Zeit untersucht, gleichzeitig aber Lücken und Defizite in der bisherigen Forschung beklagt werden.<sup>3</sup> Für eine Untersuchung der Rechtsprechung gerade des BVerfG fällt zunächst ins Gewicht, dass das BVerfG ein deutsches Gericht ist und deutsche Rechtsprechung bisher mit den Methoden der juristischen Netzwerkforschung kaum analysiert wurde.<sup>4</sup> Hinzu kommt, dass es sich um ein höchstes nationales Gericht handelt, sodass etwaige Untersuchungsergebnisse mit den zum Teil bereits vorliegenden Ergebnissen für andere höchste nationale Gerichte (z.B. den *U.S. Supreme Court*) und für supranationale Gerichte (z.B. den EuGH oder den EGMR) verglichen werden können.<sup>5</sup> Auch aus der Perspektive der deutschen Rechtswissenschaft scheint es schließlich naheliegend, mit der Untersuchung der Rechtsprechung desjenigen Gerichts zu beginnen, bei dem die deutsche Rechtsprechung „zusammenläuft“ – nicht nur, weil sich alle anderen Gerichte mehr oder weniger stark an der Rechtsprechung des BVerfG orientieren, sondern auch, weil seine Rechtsprechung für Vertreter aller rechtswissenschaftlichen Fächer potenziell relevant ist. Hierin liegt auch die Hoffnung begründet, dass alle Leser (nicht nur diejenigen, die sich als „Verfassungsrechtler“ identifizieren) etwas für sie Interessantes in den unten geschilderten Experimenten finden mögen. Ein letzter, wesentlicher Grund dafür, die juristische Netzwerkforschung zum deutschen Recht bei der Rechtsprechung des BVerfG zu beginnen, ist

---

<sup>2</sup> Dazu S. 28–61.

<sup>3</sup> Dazu beispielsweise *Whalen*, Michigan State Law Review 2016, S. 539 (548–551, 554). Zur Literatur außerdem oben S. 13–14.

<sup>4</sup> In der Anlage am nächsten kommt dieser Studie ein Beitrag von Dezember 2014 auf der Website des Alexander von Humboldt Instituts für Internet und Gesellschaft (HIIG), der Zitate zwischen 1.394 Entscheidungen aus dem Projekt „Deutschsprachiges Fallrecht (DFR)“ erkundet (abrufbar unter [www.hiig.de/blog/farbenfrohe-rechtsprechung-verweisungsanalyse-von-bverfge/](http://www.hiig.de/blog/farbenfrohe-rechtsprechung-verweisungsanalyse-von-bverfge/)). Demgegenüber konzentriert sich der Beitrag von *Engst u.a.*, JZ 72 (2017), S. 816–826, auf die Richter des BVerfG.

<sup>5</sup> Einen solchen Vergleich unternehmen etwa *Derlén und Lindholm*, German Law Journal 18 (2017), S. 647–686, für den *U.S. Supreme Court* und den EuGH; kritisch dazu *Frankenreiter*, German Law Journal 18 (2017), S. 687–693.

schließlich die Praktikabilität: Hier ist es möglich, auch mit begrenzten Ressourcen Daten zusammenzustellen und diese auf Basis einer klaren Netzwerkdefinition zu analysieren.

Vor diesem Hintergrund kann nun die Präzisierung des Untersuchungsgegenstands erfolgen: Was genau soll gemeint sein, wenn „die Rechtsprechung des BVerfG“ mit den Methoden der juristischen Netzwerkforschung untersucht wird? Diese Frage lässt sich in zwei Schritten beantworten. Der erste Schritt ist die Netzwerkdefinition (I.): Durch welche Objekte und Beziehungen soll „die Rechtsprechung des BVerfG“ als Untersuchungsgegenstand repräsentiert werden? Im zweiten Schritt erfolgt die Netzwerkkonstruktion (II.). Hier wird die Netzwerkdefinition operationalisiert, wobei wiederum ein mehrschrittiger Prozess durchlaufen wird. Am Ende des zweiten Schritts liegen – unter anderem in Form von Graphen – strukturierte Daten zur Rechtsprechung des BVerfG vor, mit denen im Rahmen der Experimente gearbeitet werden kann.

### *I. Netzwerkdefinition*

Orientiert man sich bei der Netzwerkdefinition daran, wie andere Studien zu Netzwerken in der Rechtsprechung ihren Untersuchungsgegenstand konkretisieren, so liegen im hiesigen Kontext zunächst zwei Festlegungen nahe. Erstens: Die Objekte des Netzwerks (d.h. Punkte im Graphen) sind Entscheidungen des BVerfG. Zweitens: Asymmetrische Beziehungen (d.h. Pfeile im Graphen) sind die Zitate zwischen diesen Entscheidungen.<sup>6</sup>

Damit ist immerhin geklärt, dass nur ein Bruchteil aller Zitate, die sich in den Entscheidungen des BVerfG finden lassen, in die Analyse einfließt. Denn das BVerfG zitiert in seinen Entscheidungen nicht nur andere BVerfG-Entscheidungen, sondern beispielsweise auch Entscheidungen anderer Gerichte, Gesetzestexte und Literaturbeiträge (Abb. 4.1, S. 230). Die Einschränkung der betrachteten Zitatmenge auf Zitate zwischen BVerfG-Entscheidungen zielt darauf ab, ein geschlossenes Netzwerk zu definieren, ohne dabei die Menge der Objekte rekursiv erweitern zu müssen:<sup>7</sup> Alle betrachteten Beziehungen sollen sowohl ihren Start- als auch ihren Endpunkt in der ursprünglich angegebenen Objektmenge haben, und die Objektmenge ist hier gemäß der ersten Festlegung eine Menge einzelner Entscheidungen des BVerfG.<sup>8</sup>

<sup>6</sup> Beispiele für ähnliche Festlegungen in der einschlägigen Literatur zur Rechtsprechung anderer Gerichte: *Fowler u.a.*, *Political Analysis* 15 (2007), S. 324 (325 – für den *U.S. Supreme Court*); *Lupu und Voeten*, *British Journal of Political Science* 42 (2012), S. 413 (424 – für den EGMR); *Derlén und Lindholm*, *Journal of International Economic Law* 20 (2017), S. 257 (259 – für den EuGH).

<sup>7</sup> Zum Rekursionsproblem bereits S. 59–61.

<sup>8</sup> Wollte man, ausgehend von BVerfG-Entscheidungen, z.B. auch Zitate von BGH-Entscheidungen als Beziehungen ins Netzwerk aufnehmen, so stellte sich die Frage, ob man die Zitate, die von den zitierten BGH-Entscheidungen ausgehen, ebenfalls einbeziehen sollte – und falls ja, ob man auch die Zitate der von den BGH-Entscheidungen zitierten Dokumente einbezie-

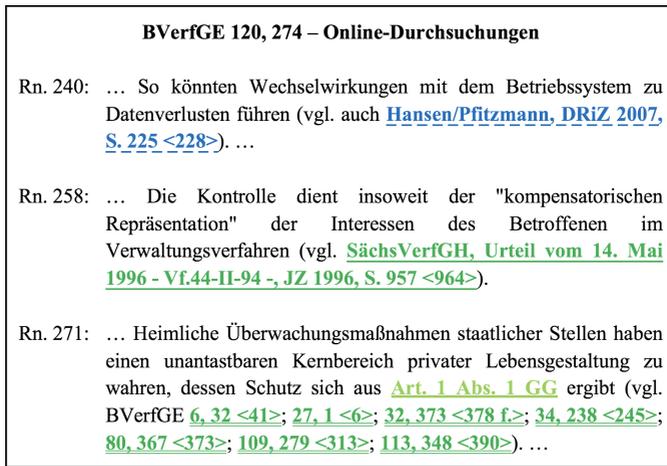


Abbildung 4.1: Zitierte Quellen in einer Entscheidung des BVerfG<sup>9</sup>

### 1. Objekte: Entscheidungen in BVerfGE 1–140

Die Definition der Objektmenge als Menge von BVerfG-Entscheidungen reicht allerdings nicht aus, um Klarheit über die zu erfassenden Objekte zu schaffen. Hierfür muss vielmehr spezifiziert werden, *welche* BVerfG-Entscheidungen eingeschlossen werden sollen. Intuitiv wird man hier „Alle Entscheidungen!“ antworten (und sich fragen: „Was denn auch sonst?“). Das ist allerdings problematisch, denn es ist äußerst unklar, was „alle Entscheidungen“ eines Gerichts sind.<sup>10</sup> Inspiziert man die Netzwerkdefinitionen in der Literatur genauer, so heißt „alle Entscheidungen“ dort häufig „alle Entscheidungen, die dem Forscher bei Durchführung der Studie zur Verfügung standen“.<sup>11</sup> Welche Entscheidungen davon umfasst sind, hängt dann insbesondere von der Datenquelle ab. Überträgt man diese Grundsätze

hen sollte, ... Die Beschränkung auf Zitate zwischen BVerfG-Entscheidungen vermeidet solche Probleme.

<sup>9</sup> Gesetzestexte (hier: Grundgesetz) sind hellgrün und dick unterstrichen; Rechtsprechung ist dunkelgrün und doppelt (Entscheidungen des BVerfG) oder einfach (andere Entscheidungen) unterstrichen; Literatur ist blau und gestrichelt unterstrichen. Wo hier und im Folgenden Randnummern zu den in BVerfGE veröffentlichten Entscheidungen zitiert werden, sind damit die *juris*-Randnummern gemeint; der zitierte Entscheidungstext entspricht der *juris*-Formatierung. Grund dafür ist, dass der Fallstudie im Folgenden die *juris*-Texte zugrunde gelegt werden; dazu S. 236–240. Soweit nicht abweichend gekennzeichnet, stammen die angegebenen Entscheidungsnummern von der Website des Projekts „Deutschsprachiges Fallrecht (DFR)“; Einzelheiten in Fn. 13 (S. 231) sowie auf S. 248–249.

<sup>10</sup> Dazu bereits oben S. 76–79; zu Eingrenzungsmöglichkeiten außerdem S. 32–33.

<sup>11</sup> Beispiele: *Fowler u.a.*, *Political Analysis* 15 (2007), S. 324 (327–328); *Smith*, *San Diego Law Review* 44 (2007), S. 309 (312–314); *Cross, Smith und Tomarchio*, *Emory Law Journal* 57 (2008), S. 1227 (1237–1238).

auf die Rechtsprechung des BVerfG, so könnte eine Definition der Objektmenge also lauten: „Alle Entscheidungen, die bei der Suche nach dem BVerfG als Gericht am {Datum der Suche} in {Datenbank des Vertrauens} zur Verfügung standen.“ Diese Definition hat aber den Nachteil, dass sie die Erfassungskriterien des Datenbankanbieters zu den ausschlaggebenden Faktoren der Mengendefinition macht und gleichzeitig auf eine konsequente Anwendung dieser Kriterien vertraut. Dass dieses Vertrauen enttäuscht wird, ist allerdings nicht unwahrscheinlich.<sup>12</sup> Letztlich weiß man beim geschilderten Vorgehen also gerade nicht, welche Entscheidungen die Netzwerkdefinition einschließt.

Gibt es also überhaupt Möglichkeiten, eine Menge von BVerfG-Entscheidungen zu definieren, deren Vollständigkeit sich überprüfen lässt? Ja, die gibt es. Allerdings geht die Sicherstellung von Vollständigkeit in Bezug auf die gewählte Definition auf Kosten der Vollständigkeit in Bezug auf die (unbekannte) Grundgesamtheit aller BVerfG-Entscheidungen. Eine Option ist der Rückgriff auf die bekannteste Sammlung von Entscheidungen des Gerichts: BVerfGE. Dann lässt sich die Menge der erfassten Objekte etwa so definieren: „alle Entscheidungen des BVerfG, die in den Bänden {Startband}–{Endband} der Sammlung BVerfGE veröffentlicht wurden“. Damit treten die Kriterien für eine Aufnahme in die Sammlung an die Stelle der Erfassungskriterien der Datenbankanbieter. Nach § 31 I der Geschäftsordnung des BVerfG (BVerfGGO) werden die Entscheidungen des Plenums (§ 16 I BVerfGG) und der Senate „in einer vom Gericht autorisierten Sammlung der Entscheidungen des Bundesverfassungsgerichts veröffentlicht“, sofern die Veröffentlichung nicht per Beschluss ausgeschlossen wird (§ 31 II BVerfGGO), und die autorisierte Sammlung ist in der Gerichtspraxis BVerfGE. Diese Sammlung enthält also keine repräsentative Stichprobe aller BVerfG-Entscheidungen, sondern eine Auswahl „besonders wichtiger“ Entscheidungen. Das ist immer noch nicht ideal, als Ausgangspunkt einer Analyse aber zumindest vertretbar, sofern das Zustandekommen der untersuchten Entscheidungsmenge bei der Interpretation der Analyseergebnisse berücksichtigt wird. Nach alledem soll für die Zwecke der hiesigen Untersuchung folgende Definition der Objektmenge zugrunde gelegt werden:

„Objekte sind alle Entscheidungen des BVerfG, denen sich in den Bänden 1–140 der Sammlung BVerfGE eine Fundstelle des Formats ‚{Band}, {Startseite}‘ zuordnen lässt (BVerfGE-Entscheidungen).“<sup>13</sup>

<sup>12</sup> Dazu *Coupette und Fleckner*, JZ 73 (2018), S. 379 (380–383).

<sup>13</sup> Wo im Folgenden Randnummern zu den BVerfGE-Entscheidungen zitiert werden, sind die *juris*-Randnummern gemeint (die Sammlung selbst hat Randnummern erst ab dem 132. Band). Die zu den Entscheidungen angegebenen Namen stammen grundsätzlich aus dem Projekt „Deutschsprachiges Fallrecht (DFR)“. Für Entscheidungen, die dort keinen Namen haben, wird hilfsweise auf die mittlerweile verfügbare Entscheidungsliste auf der Website des BVerfG zurückgegriffen; die so gefundenen Namen werden in [eckige] Klammern gesetzt. Entscheidungen, die auch auf der BVerfG-Website keinen Namen tragen, werden bei Bedarf von der Verfasserin selbst benannt; die so vergebenen Namen werden in {geschweifte} Klammern gesetzt. Zur Quellenwahl für die Entscheidungsamen noch unten S. 248–249.

## 2. Beziehungen: Zitate im Format „{Band}, {Startseite}“

Auf dieser Basis kann nun präzisiert werden, welche Beziehungen in der Netzwerkdefinition berücksichtigt werden sollen. Entsprechend der zweiten Festlegung lautet die grobe Antwort hierauf: Zitate zwischen den einzelnen Entscheidungen sollen als asymmetrische Beziehungen erfasst werden. Aber welche Zitate genau – und was zählt überhaupt als Zitat? In der Literatur findet man hier oft die Antwort „alle Zitate“, was nicht selten zu lesen ist als „alle Zitate, die derjenige, der den Datensatz erstellt hat, mit seinen Methoden aus den verfügbaren Dokumenten extrahiert hat“.<sup>14</sup> Diese Einschränkung lässt sich kaum vermeiden. Im Verhältnis zum Umfang der Auswirkungen, welche die Beziehungsdefinition auf die Ergebnisse späterer Analysen hat, sind die Ausführungen hierzu in vielen Publikationen aber erstaunlich kurz.<sup>15</sup> Hinzu kommt, dass unter „alle Zitate“ meist „alle Zitate in binärer Zählung“ verstanden wird.<sup>16</sup> Auch werden die Zitierdaten bisher in der Regel auf der Ebene einzelner Gerichtsentscheidungen, nicht aber auf der Ebene einzelner Zitierblöcke erhoben.<sup>17</sup>

Was bedeutet dieser Befund für die Definition der Beziehungen im Rahmen der hiesigen Fallstudie? Zunächst sollen Zitate nicht in binärer Zählung, sondern so erfasst werden, dass sie auch gewichtet gezählt werden können. Das bedeutet, dass jedes Zitat, das bestimmte (sogleich zu konkretisierende) Kriterien erfüllt, als eigenständiges Zitat erfasst werden soll – unabhängig davon, ob es auf eine Entscheidung zeigt, die in der zitierenden Entscheidung bereits zitiert wird. Diese Festlegung hat drei Gründe: Erstens lässt sich die binäre Zählung aus der gewichteten Zählung ableiten. Zweitens ist nicht bekannt, inwieweit binäre und gewichtete Zählung von Zitaten im Rahmen juristischer Netzwerkstudien zu unterschiedlichen Ergebnissen führen, sodass sich die detailliertere Erfassung allein zur Untersuchung dieser Fragestellung lohnt. Und drittens könnte es sein, dass binäre und gewichtete Zählung – sollten sie bei der Methodenanwendung zu unterschiedlichen Ergebnissen führen – verschiedene Informationen über das

<sup>14</sup> Beispiele: *Fowler und Jeon*, *Social Networks* 30 (2008), S. 16 (18); *Clark, Lauderdale und Katz*, *Political Analysis* 20 (2012), S. 329 (338); *Šadl und Olsen*, *Leiden Journal of International Law* 30 (2017), S. 327 (336–337).

<sup>15</sup> Dazu bereits Kapitel 2 Fn. 139 (S. 61).

<sup>16</sup> Zur Erinnerung: Bei der binären Zählung von Zitaten wird für jede zitierende Entscheidung lediglich erfasst, welche anderen Entscheidungen sie *mindestens einmal* zitiert. Demgegenüber wird bei der gewichteten Zählung von Zitaten für jede zitierende Entscheidung erfasst, welche anderen Entscheidungen sie *wie oft* zitiert. Beispiele für die binäre Zählung in netzwerkanalytischen Studien zu Rechtsprechungszitaten finden sich in Kapitel 3 Fn. 11 (S. 78); dort auch zu den (noch) wenigen Ausnahmen.

<sup>17</sup> Der Verfasserin ist keine Studie zu einem Rechtsprechungszitiernetzwerk bekannt, die Daten zu Zitierblöcken erfasst hätte. Das muss nicht heißen, dass es keine solche Studie gibt (gegenteilige Hinweise werden gern entgegengenommen). Wer aber davon ausgeht, dass sich schon die gewichtete Zählung nicht lohnt, der wird die Erhebung von Daten auf der Ebene von Zitierblöcken kaum in Betracht ziehen.

untersuchte Netzwerk zum Ausdruck bringen, die für verschiedene juristische Fragestellungen relevant sind.<sup>18</sup>

Darüber hinaus ist klarzustellen, welche Muster im Text von dem Verfahren, das zur Zitatextraktion eingesetzt wird, auf jeden Fall als Zitate erkannt werden sollen. Dabei erfolgt zwangsläufig eine Abwägung zwischen der Übersichtlichkeit der Spezifikation und der durch sie erreichten Abdeckung. Denn während der überwiegende Teil aller Zitate einem einheitlichen Muster (mit über die Zeit leicht variierender Notation) entspricht, folgt der Rest vielen verschiedenen Mustern. Da schon die Objektmenge als Menge der Entscheidungen mit Fundstellen in BVerfGE definiert wurde, liegt es nahe, für die hiesigen Zwecke die Übersichtlichkeit zu betonen und folgende Definition von Beziehungen zugrunde zu legen:

„Asymmetrische Beziehungen sind alle Vorkommen von BVerfGE-Fundstellen im Format ‚{Band}, {Startseite}‘ in Entscheidungen mit ebensolchen Fundstellen; sie verlaufen von der Fundstelle der Entscheidung, in der sie gefunden wurden, zur gefundenen Fundstelle.“

Diese Festlegung hat weitreichende Konsequenzen. Zunächst setzt sie stillschweigend voraus, dass die Fundstellen in BVerfGE eindeutig sind. Das ist aber nur bei Angaben im Format ‚{Band}, {Startseite}–{Endseite}‘ der Fall (*Langfundstelle*). In Zitaten gibt das BVerfGE die zitierten Entscheidungen hingegen regelmäßig nur im Format ‚{Band}, {Startseite}‘ an (*Kurzfundstelle*). Einige der in BVerfGE abgedruckten Entscheidungen sind allerdings so kurz, dass zwei Entscheidungen auf derselben Seite der Sammlung beginnen.<sup>19</sup> Zitate zu dieser Startseite lassen sich also nicht eindeutig einer zitierten Entscheidung zuordnen.<sup>20</sup> Wer bei der Lektüre einer einzelnen Entscheidung ein Zitat zu einer zweideutigen Kurzfundstelle findet, wird meist dem Kontext entnehmen können, welche der in Betracht kommenden Entscheidungen im betroffenen Zitat gemeint ist (entweder anhand einer konkret zitierten Seite oder anhand des Inhalts der Entscheidung). Eine solche „Auflösung im Einzelfall“ würde die Netzwerkdefinition allerdings erheblich verkomplizieren und stünde außer Verhältnis zur geringen Zahl der betroffenen Entscheidungen.<sup>21</sup> Ohne Einzelfallprüfung kann das Problem auf verschiedene Weisen gehandhabt werden. Für die Zwecke der hiesigen Untersuchung wurde die Festlegung getroffen, zwei Entscheidungen mit derselben Kurzfundstelle wie eine Fundstelle zu behandeln, sie also im Netzwerk durch ein Objekt zu repräsentieren

<sup>18</sup> Allgemein *Coupette und Fleckner*, JZ 73 (2018), S. 379 (384).

<sup>19</sup> Der Verfasserin sind in BVerfGE 1 bis BVerfGE 140 vier solcher Fälle bekannt: BVerfGE 25, 156–156 und BVerfGE 25, 156–157; BVerfGE 87, 152–152 und BVerfGE 87, 152–153; BVerfGE 104, 42–42 und BVerfGE 104, 42–51; BVerfGE 104, 357–357 und BVerfGE 104, 357–370.

<sup>20</sup> In der Terminologie relationaler Datenbanken: Die Fundstellen im Format ‚{Band}, {Startseite}‘ sind keine Schlüssel(kandidaten).

<sup>21</sup> Konkret sind vier Kurzfundstellen nicht eindeutig, dazu bereits Fn. 19 (S. 233).

(Auslegung der obenstehenden Definition der Objektmenge) und eingehende wie ausgehende Zitate auf dieses eine Objekt zu aggregieren.<sup>22</sup>

Bedeutsamer dürfte die mit der Beziehungsdefinition einhergehende Entscheidung sein, alle nicht im Standardformat auftretenden Zitate aus der Betrachtung auszuschließen. Diese Einschränkung betrifft insbesondere Zitate von Entscheidungen, die zum Zitierzeitpunkt noch keine BVerfGE-Fundstelle haben (und dann z.B. nach ihren Aktenzeichen zitiert werden). Dadurch, dass die obige Beziehungsdefinition dieses Zitierformat nicht erfasst, wird die Relevanz einzelner Entscheidungen für andere Entscheidungen, die wenig später ergehen, tendenziell unterschätzt.<sup>23</sup>

### 3. Zusatzinformationen: Merkmale

Aus den vorstehenden Erwägungen ergibt sich zunächst eine minimale Netzwerkdefinition: Es ist geklärt, was Objekte und Beziehungen sein sollen (wie diese Objekte und Beziehungen in Punkte und Pfeile zu übersetzen sind, wird im Rahmen der Erläuterungen zur Netzwerkkonstruktion behandelt). Wenn man sich allerdings (wie hier) in der Position befindet, selbst Daten zusammenzustellen, lohnt es sich, neben Objekten und Beziehungen auch einige Merkmale zu erheben, die für spätere Analysen relevant sein könnten.

Dies sind für die Entscheidungen in BVerfGE neben der Langfundstelle und den Informationen, die sich daraus entnehmen lassen (Band, Startseite, Endseite, Kurzfundstelle und Entscheidungslänge [Schätzung]),<sup>24</sup> insbesondere das Entscheidungsdatum sowie das Aktenzeichen und die darin enthaltenen Informationen (Senat, Registerzeichen, Eingangsjahr und Eingangsnummer).<sup>25</sup> Um die Inter-

<sup>22</sup> Eine andere Möglichkeit wäre gewesen, im Kollisionsfall eine der jeweils zwei betroffenen Entscheidungen (z.B. die zuerst abgedruckte) ganz aus der Analyse ausschließen und dabei in Kauf zu nehmen, dass manche Zitate bei späterer manueller Inspektion sachlich unpassend erscheinen. Der praktische Unterschied dürfte gering sein.

<sup>23</sup> Eine manuelle Sichtung der Entscheidungen lässt vermuten, dass das Problem für Entscheidungen jüngerer Datums größer ist als für ältere Entscheidungen. Um das Problem zu beheben, kann man neben dem Standard-Zitierformat weitere Zitierformate in die Beziehungsdefinition aufnehmen und dann als Bedingung hinzufügen, dass das zitierte Aktenzeichen sich auf eine BVerfGE-Fundstelle abbilden lässt. Dadurch kommen im angewendeten Extraktionsverfahren allerdings einige Schritte hinzu.

<sup>24</sup> Es gibt verschiedene Möglichkeiten, die Länge einer Entscheidung zu definieren, in der Reihenfolge abnehmender Genauigkeit etwa als Anzahl der Zeichen, Anzahl der Wörter (aber was genau ist ein Wort?), Anzahl der Randnummern oder Anzahl der Druckseiten. Alle Definitionen haben Vor- und Nachteile; aus der Langfundstelle entnehmen lässt sich allerdings lediglich eine Schätzung für die Anzahl der Druckseiten ( $\{\text{Endseite}\} - \{\text{Startseite}\}$  bzw. 1, falls die Startseite gleich der Endseite ist).

<sup>25</sup> Der Verfasserin ist genau eine in BVerfGE 1–140 veröffentlichte Entscheidung bekannt, der im Aktenzeichen keine Zahl für den Senat zugeordnet ist: BVerfGE 1, 76 (PBvV 1/51). „Aktenzeichen“ heißt außerdem strenggenommen nur der hintere Teil dessen, was üblicherweise mit Aktenzeichen gemeint ist (also etwa das „1/51“ in „PBvV 1/51“). Hier wird der Sprachgebrauch der juristischen Datenbanken zugrunde gelegt, welche die Senatsbezeichnung und das

pretation von Analyseergebnissen zu erleichtern, hilft es außerdem, einzelnen Entscheidungen die Namen zuzuordnen, unter denen sie in Fachkreisen bekannt sind, z.B. „Naßauskiesung“ (BVerfGE 58, 300), „Kruzifix“ (BVerfGE 93, 1) oder „Online-Durchsuchungen“ (BVerfGE 120, 274).<sup>26</sup>

Bei den Zitaten ist es hilfreich, für jedes Zitat ein Zeichen mitzuführen, aus dem sich rekonstruieren lässt, in welchem Zitierblock innerhalb der zitierenden Entscheidung das Zitat steht und in welcher Reihenfolge die einzelnen Zitierblöcke innerhalb der zitierenden Entscheidung auftreten (*Zitierblockzeichen*).<sup>27</sup> So kann man später bei Bedarf zwischen Co-Zitation auf der Ebene einzelner Entscheidungen und Co-Zitation auf der Ebene von Zitierblöcken differenzieren.<sup>28</sup> Außerdem wird es möglich, die „Zitierdramaturgie“ in einzelnen Entscheidungen nachzuverfolgen und zu ergründen, inwieweit sich daraus nützliche Informationen entnehmen lassen.<sup>29</sup>

## II. Netzwerkkonstruktion

Wie die im Rahmen der Netzwerkdefinition formulierten Anforderungen praktisch umzusetzen sind, ist eine Frage der Netzwerkkonstruktion. Hier wird die gewählte Netzwerkdefinition operationalisiert. Dazu ist einige Eigenarbeit erforderlich, denn es existiert noch kein Datensatz, welcher der oben hergeleiteten Netzwerkdefinition entspricht. Auch gibt es keine Quelle, in der Zitierdaten zu BVerfGE-Entscheidungen so detailliert aufbereitet und zugänglich sind, dass daraus der gewünschte Datensatz mit wenigen Klicks (bzw. wenigen Zeilen Code) zusammengebaut werden kann. Um zu den strukturierten Netzwerkdaten zu gelangen, die für spätere Analysen benötigt werden, sind daher mindestens drei (aufeinander aufbauende) Schritte erforderlich (Abb. 4.2, S. 236): Die BVerfGE-Entscheidungen müssen gesammelt werden (1.), aus ihnen sind Zitate zu extrahieren (2.), und die so produzierten Daten müssen für die Analyse aufbereitet werden

---

Registerzeichen im Aktenzeichen mitführen. Wo mehrere Aktenzeichen in einer Entscheidung zusammengefasst sind, stellt sich die Frage, welche Metadaten genau zu erfassen sind. Dazu noch unten bei Fn. 75 (S. 249).

<sup>26</sup> Zu den hier zugrunde gelegten Entscheidungsnamen bereits Fn. 13 (S. 231) sowie unten S. 248–251.

<sup>27</sup> Zur Erinnerung: Zitierblöcke sind Sequenzen von Zitaten, die in einem Dokument unmittelbar hintereinander stehen; dazu bereits Abb. 3.1 (S. 77) und begleitender Text. Beispielsweise bilden die sieben Zitate in Rn. 271 von BVerfGE 120, 274 – Online-Durchsuchungen einen Zitierblock (Abb. 4.1, S. 230).

<sup>28</sup> Dazu bereits bei Abb. 3.35 (S. 191). Man kann erwägen, auch die zitierte Seite oder die Randnummer, in der zitiert wurde, mit dem Zitat zu erheben; dann muss das Verfahren für die Zitatextraktion entsprechend angepasst werden. Hier zeigt sich das enge Zusammenspiel von Netzwerkdefinition und Netzwerkkonstruktion.

<sup>29</sup> Anstatt „nur“ die Reihenfolge der Zitierblöcke untereinander im Zitierblockzeichen zu kodieren, könnte man die Position der Zitierblöcke auch zeichengenau erheben oder versuchen, zusätzlich nach einzelnen Teilen der Entscheidung (Sachverhaltschilderung, Parteivortrag, Zulässigkeit, Begründetheit) oder nach Randnummern zu differenzieren.

(3.). Im Folgenden wird zu jedem dieser drei Schritte eine Umsetzungsmöglichkeit geschildert, die dem tatsächlichen Vorgehen der Verfasserin nahe kommt.<sup>30</sup>

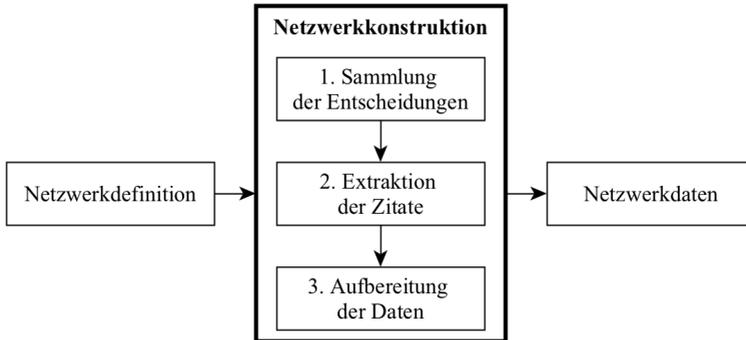


Abbildung 4.2: Schritte bei der Netzwerkkonstruktion

### 1. Sammlung der Entscheidungen

Die Sammlung der BVerfGE-Entscheidungen beginnt mit der Wahl der Datenquelle. Es liegt nahe, hier auf eine Quelle zurückzugreifen, die auf die Sammlung BVerfGE beschränkt ist. Eine digitale Version der BVerfGE kann unter anderem vom Verlag bezogen werden, allerdings als CD-ROM (!), auf der die Daten zwar für die Interaktion über eine graphische Benutzeroberfläche aufbereitet, aber für eine unmittelbare automatisierte Verarbeitung unbrauchbar gespeichert sind.<sup>31</sup> Es bleibt daher im Wesentlichen eine Quelle, die eine vollständige digitale Abdeckung der BVerfGE-Entscheidungen verspricht: *juris*.

Die Möglichkeiten, über *juris* systematisch an Gerichtsentscheidungen zu gelangen, haben sich während der Entstehungsphase dieser Arbeit verschlechtert; das nachfolgend geschilderte Vorgehen funktioniert nur mit der „alten“ Benutzerober-

<sup>30</sup> Das tatsächliche Vorgehen lässt sich kaum übersichtlich darstellen und seine Schilderung brächte vor allem die Erkenntnis, dass *Data Mining* – der Bergbau der Wissensarbeiter – weit weniger glamourös ist, als oft angenommen wird. Die Umsetzung der einzelnen Schritte, wie sie im Folgenden skizziert wird, ist eine „glattegezogene“ Fassung dessen, was tatsächlich unternommen wurde, um die in den Experimenten untersuchten Daten zusammenzustellen. Diese Fassung ist selbst wiederum aus zahlreichen Irrwegen und Fehlversuchen hervorgegangen. Sie hat in der konkreten Durchführung einige Schwachstellen, auf die im Folgenden hingewiesen wird. Die Darstellung des gewählten Vorgehens erfolgt auch deshalb so ausführlich, um eine Diskussion über *Best Practices* bei der Arbeit mit juristischen Daten (insbesondere in der juristischen Netzwerkforschung) zu ermöglichen.

<sup>31</sup> Auf der Website des Verlags findet sich außerdem folgender, im 21. Jahrhundert befremdliche Hinweis: „Systemvoraussetzung: Microsoft Windows bis einschließlich Version 10. Läuft weder unter Linux noch unter macOS.“.

fläche, nicht in der (vom Datenbankanbieter propagierten) „neuen Einfachheit“.<sup>32</sup> Über die alte Benutzeroberfläche ist es möglich, die Volltexte mehrerer hundert Dokumente gemeinsam in einer PDF-Datei herunterzuladen, nachdem man eine Suche durchgeführt und die jeweils auf einer Seite angezeigten Treffer alle gemeinsam markiert hat. Sucht man also nach der Rechtsprechung des BVerfG in bestimmten Jahreszeiträumen, so kann man sich dadurch mit wenig Aufwand die in *juris* verfügbare Rechtsprechung des BVerfG in Form von mehreren Sammel-PDFs verfügbar machen (Schritt 1: Download).<sup>33</sup> In der „neuen Einfachheit“ scheint *juris* diese Möglichkeit abgeschafft zu haben.

Liegen zur über *juris* verfügbaren Rechtsprechung des BVerfG Sammel-PDFs vor, so lassen sich diese in Sammel-TXTs umwandeln (Schritt 2: Konvertierung).<sup>34</sup> Welche Werkzeuge dafür in Betracht kommen, hängt vom Betriebssystem des Computers ab, auf dem die Konvertierung erfolgen soll; im konkreten Fall wurde ein Windows-Tool von *SomePDF* (Version 2.0.0.0) verwendet.<sup>35</sup> Die Sammel-TXTs können dann in Einzel-TXTs mit je einer Entscheidung aufgeteilt werden, indem man an wiederkehrenden Textbausteinen trennt, die den Übergang zwischen zwei Entscheidungen signalisieren (Schritt 3: Zerschneiden). Aus diesen Einzel-TXTs lassen sich mithilfe einfacher regulärer Ausdrücke das Entscheidungsdatum und das führende Aktenzeichen ermitteln, sodass die einzelnen TXT-Dateien aussagekräftig nach dem Muster „{Entscheidungsjahr}-{Entscheidungsmonat}-{Entscheidungstag} BVerfG {Aktenzeichen, in dem ‚/‘ durch ‚-‘ ersetzt wurde}“ benannt werden können (Schritt 4: Text Mining + Umbenennung).

Die Dateien, die am Ende von Schritt 4 vorliegen, enthalten nicht nur Entscheidungen, die in BVerfGE veröffentlicht wurden.<sup>36</sup> Um die in BVerfGE veröffentlichten Entscheidungen herauszufiltern, ohne einen Aktenzeichenabgleich vornehmen zu müssen, kann man die im *juris*-Entscheidungskopf annotierten Fundstellen auslesen und daraus eine (noch nachzubearbeitende und zu überprüfende) Liste erstellen, die einen Dateinamen auf die im Kopf der

<sup>32</sup> Als die Zusammenstellung des nachfolgend verwendeten Datensatzes erfolgte (im Juni 2016), war die „alte“ Benutzeroberfläche noch die einzige Benutzeroberfläche.

<sup>33</sup> Die Verfasserin geht davon aus, dass dies für den wissenschaftlichen Gebrauch zulässig ist; siehe auch oben S. 82–85.

<sup>34</sup> TXT ist ein simples Dateiformat, dessen Inhalt als Sequenz von Zeichen in einem bestimmten Zeichensatz (z.B. ASCII oder Unicode) interpretiert wird. Wird eine Datei in das TXT-Format konvertiert, so gehen insbesondere die Informationen zur Darstellung des darin enthaltenen Textes verloren, z.B. Hervorhebungen und Layout.

<sup>35</sup> Für Nutzer von OSX ist die Command-Line-Utility *textutil* eine Alternative; mit OSX oder Ubuntu/Debian kann außerdem die Python-Bibliothek *extract* verwendet werden.

<sup>36</sup> Anders verhält es sich, wenn man die Suche nach Entscheidungen des BVerfG bereits in *juris* selbst auf BVerfGE eingeschränkt hat. Dass die Verfasserin ein anderes Vorgehen wählte, hat zwei Gründe: erstens Misstrauen gegenüber der Annotation von BVerfGE-Fundstellen in *juris* und zweitens Unsicherheit hinsichtlich der Netzwerkdefinition, die sie der Arbeit zugrunde legen würde (die auf S. 229–235 dargestellte Definition ist Ergebnis eines explorativen Prozesses, in dem die Verfasserin festzustellen versuchte, aus welcher Entscheidungsmenge sich ein geschlossenes und sauberes Korpus zusammenstellen lässt).

enthaltenen Entscheidung gefundene BVerfGE-Fundstelle abbildet (*Fundstellenliste*).<sup>37</sup> Das ist etwas umständlich, da im Zuge von Datei-Download als PDF und Konvertierung in TXT die Formatierung des *juris*-Entscheidungskopfes in mehrere Spalten verloren gegangen ist, aber eine Aufgabe, die wiederum mithilfe regulärer Ausdrücke bewältigt werden kann (Schritt 5: Text Mining).<sup>38</sup> Um sicherzustellen, dass die Liste vollständig und korrekt ist, kann sie überprüft und bei Bedarf korrigiert werden (Schritt 6: Prüfung + Korrektur).<sup>39</sup>

<sup>37</sup> Ein Aktenzeichenabgleich ist nur möglich, wenn eine verlässliche Liste, welche die Aktenzeichen von BVerfGE-Entscheidungen auf ihre jeweiligen BVerfGE-Fundstellen abbildet, digital vorliegt. Als die Verfasserin ihre Daten zusammenstellte (im Juni 2016), war ihr keine solche Liste bekannt. Mittlerweile bietet das BVerfG auf seiner Website eine Aufstellung der Entscheidungen in BVerfGE an, die auch interessante Metadaten (z.B. Angaben zur Art einzelner Entscheidungen) enthält (abrufbar unter [www.bundesverfassungsgericht.de/DE/Entscheidungen/Entscheidungen/BVerfGE.html](http://www.bundesverfassungsgericht.de/DE/Entscheidungen/Entscheidungen/BVerfGE.html)). Aus dieser Aufstellung kann man eine strukturierte Liste erstellen, indem man den HTML-Quelltext der Webseiten zu den einzelnen „Banddekaden“ ausliest (auf der Website des BVerfG sind dies BVerfGE 1 ff. [Sonderfall: erste neun Bände], BVerfGE 10 ff., BVerfGE 20 ff., . . . , BVerfGE 140 ff. [Sonderfall: letzte, noch nicht abgeschlossene Banddekade]). Eine Prüfung dieser Liste durch die Verfasserin im November 2017 ergab allerdings einige Fehler und Lücken, welche die Verlässlichkeit der Aufstellung in Zweifel ziehen: Neben einigen Tippfehlern konnten doppelt erfasste Entscheidungen und falsch zugeordnete Metadaten festgestellt werden; einige Entscheidungen (vornehmlich am Ende einzelner Bände), darunter etwa das Apotheken-Urteil (BVerfGE 7, 377–444), fehlten ganz. Hinzu kommt, dass für Entscheidungen, die mehrere Aktenzeichen zusammenfassen, oft nicht alle Aktenzeichen angegeben sind (es steht dann „u.a.“), was einen umfassenden Aktenzeichenabgleich unmöglich macht.

<sup>38</sup> Die Fundstellen der Entscheidungen in BVerfGE 140 waren manuell nachzutragen, da sie zum Zeitpunkt des Entscheidungsdownloads noch nicht in *juris* vermerkt waren.

<sup>39</sup> Die Umsetzung des geschilderten Verfahrens durch die Verfasserin erfolgte in zwei Schritten. Im ersten Schritt (Januar 2017) wurde zunächst wie angegeben eine rohe Fundstellenliste erstellt und manuell auf offensichtliche Fehler (Resultat der Formatierungsprobleme) durchkämmt, die sofort korrigiert wurden. Die Absicherung der Vollständigkeit erfolgte danach heuristisch, indem für jeden Band der BVerfGE manuell geprüft wurde, ob die aus den Dateien extrahierten Fundstellen eine vollständige Seitenabdeckung des jeweiligen BVerfGE-Bands erreichen. Parallel zur Vollständigkeitsprüfung erfolgte die Prüfung auf Korrektheit, wobei neben einigen durch das Extraktionsverfahren bedingten Fehlern auch mehrere schon in *juris* fehlerhafte Fundstellenangaben zu korrigieren waren. Die einzige dabei zunächst als fehlend identifizierte Seite, die sich nicht als Aufzeichnungs- oder Aufbereitungsfehler herausstellte, erwies sich tatsächlich als leer (BVerfGE 96, 408). Im zweiten Schritt (November 2017) konnte die Verfasserin mithilfe der vom BVerfG bereitgestellten Fundstellenliste (Fn. 37, S. 238) eine zweite Validierung der Fundstellen vornehmen. Dazu wurde untersucht, welche Fundstellen sich *nicht* in beiden Quellen (selbst erstellte Liste und Liste des BVerfG) finden, und für alle Abweichungsfälle wurden die Fundstellen in den gedruckten BVerfGE-Bänden manuell nachgeschlagen. Hierbei stellten sich (neben einigen Fehlern in der Liste des BVerfG) weitere fünf Fehler in den extrahierten Langfundstellen heraus: BVerfGE 3, 213–224 wurde fälschlich als BVerfGE 3, 213–225 erfasst; BVerfGE 25, 156–156 als BVerfGE 25, 156–157; BVerfGE 38, 347–347 als BVerfGE 38, 347–348; BVerfGE 113, 113–128 als BVerfGE 113, 114–128 und BVerfGE 140, 316–317 als BVerfGE 140, 316–316. Diese Fehler ließen sich im Rahmen der heuristischen Prüfung (Januar 2017) nicht finden, da sie keine unvollständige Seitenabdeckung hervorriefen. Sie wurden im November 2017 korrigiert

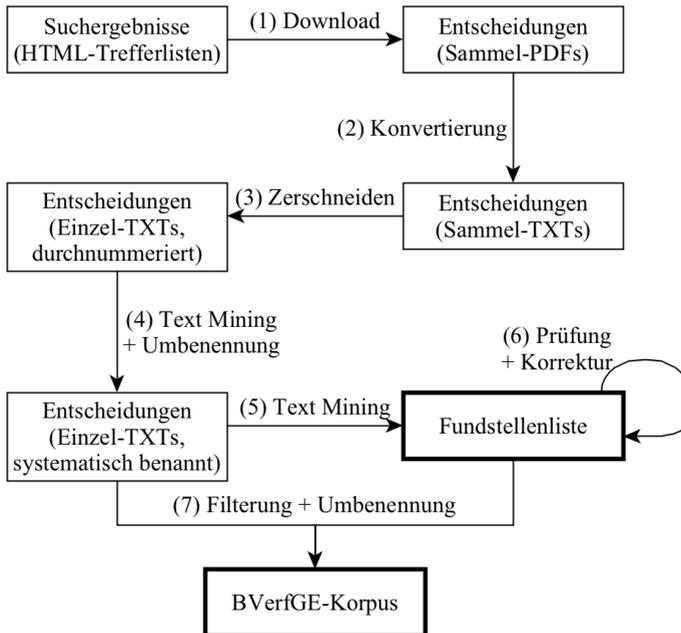


Abbildung 4.3: Vorgehen bei der Sammlung der BVerfGE-Entscheidungen

Mithilfe der überprüften und korrigierten Fundstellenliste können nun aus der Menge aller BVerfG-Entscheidungen die BVerfGE-Entscheidungen der Bände 1 bis 140 ausgesondert (präziser: in einen speziell dafür vorgesehenen Ordner kopiert) werden. Dabei kann man dem Dateinamen zur schnelleren Übersicht für Menschen die Kurzfundstelle anhängen (Schritt 7: Filterung + Umbenennung).<sup>40</sup> Die so zusammengestellten Dateien bilden das *BVerfGE-Korpus*, auf dem alle weiteren Vorbereitungs- und Analyseschritte aufbauen.<sup>41</sup> Das BVerfGE-Korpus enthält 3.277 Entscheidungen mit 3.273 verschiedenen Kurzfundstellen; das zu

(gemeinsam mit allen Folgefehlern in den bereits abgeleiteten Dateien). Im Übrigen stimmten die Fundstellen in der digitalen Liste des BVerfG mit denen in den gesammelten Dateien überein (mit Ausnahme der Entscheidungen, die in der Liste des BVerfG fehlten, und eines Tippfehlers in der Endseite einer Entscheidung in dieser Liste).

<sup>40</sup> Die Kombination aus Dateiname und Kurzfundstelle ist eindeutig, sodass es hier trotz Dopplungen bei den Kurzfundstellen nicht zu Kollisionen kommt.

<sup>41</sup> Die in den Dateien enthaltenen Entscheidungstexte befinden sich in der (aufbereiteten) Form, in der sie von *juris* zur Verfügung gestellt werden. Für zukünftige Untersuchungen wäre es vorzuziehen, auf die Originaltexte zurückzugreifen. Dies setzt allerdings voraus, dass die Originaltexte in einem Format vorliegen, das ihre automatisierte Weiterverarbeitung ermöglicht, und dass sie in dieser Form der Analyse durch Wissenschaftler zugänglich sind.

konstruierende Zitiernetzwerk hat also 3.273 Objekte.<sup>42</sup> Das skizzierte Vorgehen ist in Abbildung 4.3 (S. 239) zusammengefasst.

## 2. Extraktion der Zitate

Das BVerfGE-Korpus enthält 3.277 TXT-Dateien mit Entscheidungstexten zu 3.273 Kurzfundstellen in BVerfGE 1–140. Um die hier gewählte Netzwerkdefinition mit Zitaten als asymmetrische Beziehungen zu operationalisieren, müssen die Zitate aus den Entscheidungstexten extrahiert werden. Dabei sollen für jede zitierende BVerfGE-Entscheidung alle Zitate anderer BVerfGE-Entscheidungen erfasst werden, die dem Muster ‚{Band}, {Startseite}‘ folgen. Die Erfassung soll außerdem mit Zitierblockzeichen geschehen, anhand derer später nachvollzogen werden kann, welche Entscheidungen innerhalb desselben Zitierblocks angeführt werden.<sup>43</sup>

Diesen Anforderungen kann mit einem regelbasierten Verfahren entsprochen werden, das auf regulären Ausdrücken (*regular expressions*, *Regex*) basiert. Reguläre Ausdrücke sind eines der ältesten Werkzeuge der Textverarbeitung und dennoch so mächtig, dass viele Suchprobleme in juristischen Texten damit zufriedenstellend gelöst werden können. Mithilfe regulärer Ausdrücke lassen sich Mengen von Zeichenketten anhand abstrakter Muster definieren. Ein Muster kann sowohl konkrete Zeichen als auch Platzhalter für eine Menge von Zeichen oder Angaben zur gewünschten Anzahl der Zeichen enthalten. Durchsucht man einen Text nach einem regulären Ausdruck, so werden alle Zeichenketten als Treffer behandelt, die seinem Muster entsprechen. Wie ein Muster auszusehen hat und wie es interpretiert wird, hängt davon ab, in welcher Programmiersprache man arbeitet.<sup>44</sup>

<sup>42</sup> Zu den vier nicht eindeutigen Kurzfundstellen bereits oben Fn. 19 (S. 233). Nach der manuellen Prüfung und dem späteren Abgleich mit der Liste von der Website des BVerfG (dazu Fn. 39, S. 238) geht die Verfasserin davon aus, alle Entscheidungen des BVerfG, die in BVerfGE 1–140 veröffentlicht wurden, in das Korpus eingeschlossen zu haben. In der Liste des BVerfG tauchten jedenfalls (auch vor Vornahme der in Fn. 39 [S. 238] erläuterten Korrekturen) keine Entscheidungen auf, die im BVerfGE-Korpus fehlten (vielmehr waren in der Liste des BVerfG nach dem Abgleich mit dem BVerfGE-Korpus umgekehrt einige Fundstellen zu ergänzen).

<sup>43</sup> Dazu oben S. 232–235.

<sup>44</sup> Die Umsetzungen (*Implementierungen*) regulärer Ausdrücke in den einzelnen Programmiersprachen sind von der Menge der regulären Ausdrücke zu unterscheiden, die in der theoretischen Informatik zur Angabe einer regulären Sprache eingeführt werden. Letztere bilden lediglich eine Teilmenge der regulären Ausdrücke, die in den *Regex*-Modulen vieler Programmiersprachen (beispielsweise Python oder Java) angegeben werden können. Zu regulären Ausdrücken in der theoretischen Informatik *Schöning*, *Theoretische Informatik – kurz gefasst*, 2008, S. 28–31; ein Überblick zu regulären Ausdrücken in der hier verwendeten Programmiersprache Python findet sich in der offiziellen Dokumentation (abrufbar unter [docs.python.org/3/howto/regex.html](https://docs.python.org/3/howto/regex.html)).

In der Regex-Syntax der Programmiersprache Python könnte ein einfaches Muster zur Identifikation einzelner Zeichenketten, die BVerfGE-Fundstellen sind, beispielsweise lauten:

```
„BVerfGE\s*?\d{1,3},\s*?\d{1,3}“
```

Das bedeutet in etwa: „Suche nach der Zeichenkette ‚BVerfGE‘, gefolgt von null oder mehr (\*), aber möglichst wenigen (?) Leerzeichen (\s),<sup>45</sup> gefolgt von ein bis drei Ziffern (\d{1,3}, für die Bandangabe), einem Komma, wieder null oder mehr, aber möglichst wenigen Leerzeichen und schließlich wieder ein bis drei Ziffern (für die Angabe der Startseite)“. Diesem Muster entspricht unter anderem die Fundstelle „BVerfGE 1, 208“. Die Realität der TXT-Dateien, die aus den *juris*-Dokumenten generiert sind, sieht allerdings mitunter so aus (BVerfGE 99, 1 – Bayerische Kommunalwahlen, Rn. 34):<sup>46</sup>

```
(vgl. BVerfGE 1, 208 <237 u. 242>; 3, 383 <390 f.>; 4, 31 <39>; 4, 375, <382>; 6, 84 <91>; 11, 266 <271>; 11, 351 <360>; 12, 10 <25>; 12, 73 <76>; 13, 1 <12>; 13, 243 <246>; 18, 172 <180>; 24, 300 <340>; 28, 220 <225>; 34, 81 <98>; 41, 399 <413>; 47, 253 <269>; 48, 64 <79>; 51, 222 <232>; 52, 63 <89>; 57, 43 <56>; 58, 177 <190>; 60, 162 <167>; 69, 92 <106>; 71, 81 <94>; 78, 350 <357>; 85, 148 <157>)
```

Gesucht ist also ein regulärer Ausdruck, mit dem man beispielsweise auch „85, 148“ als Fundstelle erfassen kann – allerdings nur dann, wenn davor irgendwo „BVerfGE“ auftaucht, gefolgt von einer bis dahin ununterbrochenen Reihe anderer Fundstellen, die meist (wenngleich nicht immer) durch Semikola voneinander abgetrennt sind.<sup>47</sup> Außerdem muss man davon ausgehen, dass in den hier zugrunde gelegten Entscheidungstexten teilweise falsche Kommata gesetzt sind (im obigen Block: „4, 375, <382>“) und dass teilweise (unter Umständen mehrere) Seitenzahlen in Klammerzusätzen zitiert werden, in denen auch Text vorkommen kann. Da das BVerfGE-Korpus nicht die Originaltexte, sondern deren *juris*-Versionen enthält, können diese Klammerzusätze nicht nur in [eckigen] Klammern auftreten (so das einheitliche Format in der gedruckten Sammlung), sondern auch in <spitzen> Klammern (Abb. 4.1, S. 230) oder in (runden) Klammern.<sup>48</sup> Die zitier-

<sup>45</sup> Genau genommen werden hier nicht nur Leerzeichen, sondern alle *whitespace characters* erfasst, also etwa auch Tabulatoren oder Zeilenumbrüche.

<sup>46</sup> Hier werden 27 Fundstellen hintereinander zitiert. Das ist der größte Zitierblock, den die Verfasserin im BVerfGE-Korpus finden konnte. Der zweitgrößte besteht aus 19 Fundstellen und findet sich in derselben Entscheidung vier Randnummern weiter (Rn. 38).

<sup>47</sup> Gegenbeispiel: BVerfGE 3, 1 zitiert in Rn. 11 „(BVerfGE 1, 12 und 13)“. Hier ist noch dazu unklar, ob „BVerfGE 1, 12 f.“ (Zitat der Langfundstelle „BVerfGE 1, 12–13“), „BVerfGE 1, 12 [12, 13]“ (implizites Zitat einer Kurzfundstelle mit zwei Seitenangaben) oder – wohl am wahrscheinlichsten – „BVerfGE 1, 12; 1, 13“ (Zitat zweier Kurzfundstellen) gemeint ist. An anderen Stellen steht zwischen zwei Fundstellenangaben kein Semikolon, sondern ein Komma.

<sup>48</sup> Beispiele für Seitenangaben in Entscheidungstexten aus dem BVerfGE-Korpus, die in (runden) Klammern stehen: BVerfGE 16, 64 – Einwohnersteuer (07.05.1963, Zweiter Se-

ten Seiten sollen hier zwar nicht erfasst werden,<sup>49</sup> erschweren aber die Erkennung von Fundstellen innerhalb von Zitierrblöcken.<sup>50</sup>

Ob man mit einem Muster alle Zeichenketten findet, die man zu finden hofft, hängt davon ab, wie gut man die Zeichenketten kennt, nach denen gesucht wird, und wie umfassend man das Muster getestet hat. Dabei können zwei Arten von Fehlern auftreten: Ein *false positive* liegt vor, wenn vom angegebenen Muster eine Zeichenkette erfasst wird, die *nicht* erfasst werden sollte; ein *false negative* liegt vor, wenn von diesem Muster eine Zeichenkette *nicht* erfasst wird, die erfasst werden sollte (Abb. 4.4, S. 242).<sup>51</sup>

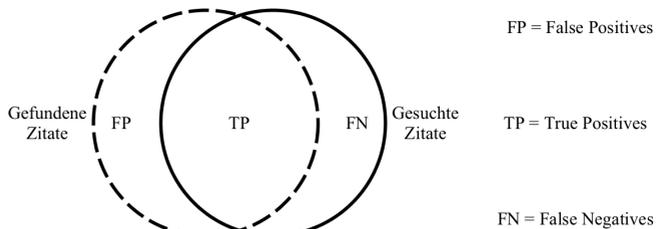


Abbildung 4.4: Verhältnis von gefundenen zu gesuchten Zitaten

nat); BVerfGE 21, 271 – Südkurier (04.04.1967, Erster Senat); BVerfGE 33, 247 – Klagegestop Kriegsfolgen (28.06.1972, Erster Senat). Beispiele für Seitenangaben in [eckigen] Klammern: BVerfGE 7, 18 – Bayerisches Ärztegesetz (28.05.1957, Zweiter Senat); BVerfGE 7, 241 – {Vorprüfungsausschuss} (23.01.1958, Erster Senat); BVerfGE 77, 137 – Teso (21.10.1987, Zweiter Senat). Beispiele für Seitenangaben in <spitzen> Klammern: BVerfGE 3, 58 – Beamtenverhältnisse (17.12.1953, Erster Senat); BVerfGE 101, 151 – Umsatzsteuerbefreiung (10.11.1999, Zweiter Senat); BVerfGE 120, 274 – Online-Durchsuchungen (27.02.2008, Erster Senat), oben Abb. 4.1 (S. 230). Welche Klammern in *juris* für die Seitenangaben verwendet werden, scheint vom Entstehungszeitraum der Entscheidung abhängig zu sein (in der Anfangszeit des Gerichts wurden spitze Klammern genutzt, dann wechselte man zu runden Klammern, schließlich wieder zu spitzen Klammern; die in der Druckversion der BVerfGE verwendeten eckigen Klammern treten nur in Ausnahmefällen auf). Zu Quelle und Notation der angeführten Entscheidungsnamen bereits Fn. 13 (S. 231) sowie unten S. 248–249.

<sup>49</sup> Dazu oben Fn. 28 (S. 235).

<sup>50</sup> Es ist beispielsweise nicht auszuschließen, dass zwei Zahlen als zitierte Seiten einer Entscheidung in einem Klammerzusatz genannt sind, die sich auch als Fundstellen in BVerfGE lesen lassen. Etwa können „91, 93“ als zitierte Seiten für jede Entscheidung auftreten, die in ihrem Band u.a. die Seiten 91 und 93 umfasst (z.B. BVerfGE 100, 59 – Rentenüberleitung II), was ohne Kontext auch als valide Fundstelle BVerfGE 91, 93 – Kindergeld interpretiert werden kann. Das Problem betrifft insbesondere Fundstellen, deren Startseite nur geringfügig größer ist als die Zahl ihres Bands.

<sup>51</sup> Die Terminologie ist in der Literatur zu *Natural Language Processing* und *Data Mining* verbreitet, siehe etwa *Jurafsky und Martin*, *Speech and Language Processing*, 2009, S. 489; *Chen u.a.*, *IEEE Transactions on Knowledge and Data Engineering* 24 (2012), S. 236 (244); *Aggarwal*, *Data Mining*, 2015, S. 340.

Um sicherzustellen, dass ein regulärer Ausdruck die gewünschte Abdeckung liefert, sollte man mit ihm schon in der Formulierungsphase experimentieren und ihn ausgiebig testen (Schritt 1: Experimente + Tests). Außerdem wird das Extraktionsverfahren leichter nachvollziehbar, korrigierbar und optimierbar, wenn man es von vornherein in drei Teilschritte untergliedert: (2a) Extraktion von Zitierblöcken als Listen (für jede Entscheidung separat); (2b) Extraktion von Zitaten als Listen aus den Listen der Zitierblöcke (für jede Entscheidung separat) und (2c) Zusammenführung aller Listen von Zitaten in einer Liste aller Zitate. Für die Zwecke dieser Arbeit wurde zunächst folgender Ausdruck verwendet, um Zitierblöcke zu extrahieren (2a):

```
fundstelle = r'\s*\d{1,3},\s*\d{1,3}(?:\s*?[(<[].*?[]>)])?'
pattern = r'BVerfGE%s(?:[,;]%)s*'%(fundstelle,fundstelle)
prog = re.compile(pattern, flags=re.DOTALL)
```

Aus den durch diesen Ausdruck gefundenen Zitierblöcken konnten die einzelnen Fundstellen extrahiert werden (2b), indem zunächst die (in runden, eckigen oder spitzen Klammern stehenden) Klammerzusätze entfernt und anschließend die verbleibenden Fundstellen mithilfe des Musters `r'\d{1,3},\s*\d{1,3}'` herausgesucht wurden.<sup>52</sup> Aus der Zusammenführung aller im Rahmen der Zitatextraktion erstellten Listen (2c) resultiert zunächst eine „rohe“ Liste aller Zitate im BVerfGE-Korpus, in der noch Selbstzitate (d.h. Annotationen der Entscheidungsfundstelle in *juris*, die vom regulären Ausdruck ebenso erfasst werden) und invalide Fundstellen (d.h. Fundstellen, die nicht in der Fundstellenliste enthalten sind) vorkommen. Diese lassen sich durch Prüfung auf Identität von zitierender und zitierter Fundstelle (findet Selbstzitate) und durch Abgleich mit der Fundstellenliste (findet invalide Fundstellen) herausuchen und entfernen (Schritt 3: Verifikation + Bereinigung).<sup>53</sup> Die so produzierte *Zitatliste* enthält 57.053 Zitate, die in 37.125 Zitierblöcken gefunden wurden, davon 31.373 unterschiedliche Paare von Kurzfundstellen ( $\{\text{Zitierende Entscheidung}\}, \{\text{Zitierte Entscheidung}\}$ ); von 3.063 der 3.273 Kurzfundstellen im BVerfGE-Korpus gehen Zitate aus; 2.905 verschiedene Kurzfundstellen werden zitiert. Das bedeutet auch: Der zu konstruierende Zitiergraph enthält 57.053 oder 31.373 Pfeile, je nachdem, ob man sich für die gewichtete Zählung oder die binäre Zählung von Zitaten entscheidet. Diese Zahlen lassen bereits vermuten, dass die Ergebnisse quantitativer Analysen in Abhängigkeit von der gewählten Zählung sehr unterschiedlich ausfallen können.

Auch wenn man die im Extraktionsverfahren verwendeten regulären Ausdrücke intensiv getestet hat, bleibt doch eine Restunsicherheit hinsichtlich der Qualität der Zitatextraktion, die mit dem dargestellten Verfahren erreicht wurde. Denn obwohl wegen der oben geschilderten Prüfung keine invaliden Fundstellen im Korpus

<sup>52</sup> Tatsächlich kommen noch einige kleinere Zwischenschritte hinzu, die anhand des Codes im Online-Appendix nachvollzogen werden können.

<sup>53</sup> Auf diese Weise hat die Verfasserin 2.581 Selbstzitate und 323 invalide Fundstellen eliminiert.

enthalten sind, so kann es doch auch valide Fundstellen geben, die bei näherer Betrachtung als *false positives* aufgefasst werden können. Das kann insbesondere dann vorkommen, wenn die Entscheidungen wie hier nicht im Original, sondern in aufbereiteter Form ausgewertet werden, da im Aufbereitungsprozess – insbesondere in den Orientierungssätzen – mitunter zusätzliche Fundstellen annotiert werden.<sup>54</sup> Solche Fälle lassen sich im Stadium der Zitatextraktion ohne Hinzuziehung weiterführender Informationen zu den betrachteten Dokumenten nicht erkennen; die regulären Ausdrücke sind „zu weit“, um sie zu vermeiden.<sup>55</sup> Da die im Aufbereitungsprozess hinzugefügten Fundstellen ebenfalls Informationen zu den sachlichen Zusammenhängen zwischen den Entscheidungen enthalten, die Aufschluss über Wichtigkeit, Zusammengehörigkeit und Charakterisierung der BVerfGE-Entscheidungen geben können, werden sie nachfolgend wie Zitate behandelt und somit *nicht* als *false positives* bewertet. Es ist aber auch nicht auszuschließen, dass das spezifizierte Muster entgegen den begründeten Erwartungen doch nicht „so gut wie alle“ Zitate erfasst, sondern in mancher Hinsicht „zu eng“ formuliert ist und daher systematisch *false negatives* produziert.<sup>56</sup>

Es bietet sich daher an, eine *Zufallsstichprobe* aus dem BVerfGE-Korpus zu ziehen (Schritt 4: Stichprobenziehung) und anhand der Zitatliste manuell zu überprüfen, wie viele *true positives* in jeder Entscheidung aus der Stichprobe richtig, wie viele *false positives* falsch und wie viele *false negatives* nicht erfasst wurden (Schritt 5: Manuelle Überprüfung).<sup>57</sup> Für jedes Dokument kann man dann bestimmen, wie hoch der Anteil echter Zitate an allen identifizierten Zitaten ist (*Genauigkeit, precision*) und wie hoch der Anteil korrekt identifizierter Zitate an allen echten Zitaten ist (*Trefferquote, recall*).<sup>58</sup> Will man Genauigkeit und Trefferquote in einer Kennzahl zusammenfassen, so kann man unter anderem ihr harmonisches Mittel bilden, das als *F-Maß (F-measure)* bezeichnet wird.<sup>59</sup>

<sup>54</sup> So ist auf *juris* teilweise die BVerfGE-Fundstelle einer Hauptsacheentscheidung zu einer Entscheidung hinzugefügt, die im einstweiligen Rechtsschutz ergangen ist (z.B. BVerfGE 28, 21 im Orientierungssatz zu BVerfGE 26, 14).

<sup>55</sup> In der Repräsentation der aufbereiteten Daten als Zitiergraph zeigen sich solche Annotationen teilweise als Zyklen im grundsätzlich azyklischen Graphen, etwa wenn die Fundstelle der Hauptsacheentscheidung zur Entscheidung im Verfahren des einstweiligen Rechtsschutzes annotiert wird (wie im Beispiel aus Fn. 54, S. 244).

<sup>56</sup> „Alle“ Zitate wird man mit regelbasierten Verfahren ohnehin kaum finden können, da man z.B. nie alle möglichen Tippfehler oder bewussten Abweichungen vom Standard-Zitiernmuster einplanen kann.

<sup>57</sup> Zur Ziehung von Zufallsstichproben beispielsweise *Kauermann und Küchenhoff*, Stichproben, 2011, S. 11–16, 42–45; *Finkelstein und Levin*, Statistics for Lawyers, 2015, S. 271–277; *Moore, McCabe und Craig*, Introduction to the Practice of Statistics, 2017, S. 191–196.

<sup>58</sup> In der Sprache von Abbildung 4.4 (S. 242): Genauigkeit =  $\frac{|TP|}{|\text{gefundenen Zitate}|}$  und Trefferquote =  $\frac{|TP|}{|\text{gesuchte Zitate}|}$ , wobei  $|X|$  die Anzahl der Elemente in der Menge  $X$ , d.h. ihre *Kardinalität* signalisiert.

<sup>59</sup> Das F-Maß (auch als  $F_1$  bezeichnet) ist definiert als  $F = \frac{2 \cdot \text{Genauigkeit} \cdot \text{Trefferquote}}{\text{Genauigkeit} + \text{Trefferquote}}$ ; siehe etwa *Aggarwal*, Data Mining, 2015, S. 340. Für die Evaluation eines Verfahrens zur Zitatextraktion

Anhand von Genauigkeit, Trefferquote und F-Maß lässt sich die Qualität der Zitatextraktion auf Ebene der einzelnen Entscheidungen in der Stichprobe einschätzen (Schritt 6: Auswertung der Prüfungsergebnisse).

Für diese Fallstudie wurde eine *einfache Zufallsstichprobe*<sup>60</sup> von 33 Entscheidungen (also ca. 1 % der Grundgesamtheit) gezogen.<sup>61</sup> Die in ihren Texten vorkommenden Zitate wurden daraufhin manuell mit den Zitaten in der Zitatliste abgeglichen.<sup>62</sup> Dabei wurde kein einziges *false positive* entdeckt, d.h. alle Zitate, die mit dem beschriebenen Extraktionsverfahren gefunden wurden, sind Zitate im Sinne der oben hergeleiteten Definition (100 % Genauigkeit für alle überprüften Entscheidungen).<sup>63</sup> In 22 der 33 untersuchten Entscheidungen wurden alle gesuchten Zitate gefunden (100 % Trefferquote in  $\frac{2}{3}$  aller Fälle);<sup>64</sup> das arithmetische Mittel der Trefferquoten liegt bei 95,1 %, allerdings sind die Trefferquoten in der Stichprobe schief verteilt.<sup>65</sup>

Von insgesamt 984 in der Stichprobe gesuchten Zitaten wurden 947 (96,2 %) gefunden. Einzelne fehlende Zitate lassen sich auf Aufzeichnungs- oder Aufbereitungsfehler zurückführen.<sup>66</sup> Viele *false negatives* treten allerdings in Grup-

---

wird es beispielsweise von *Chen u.a.*, IEEE Transactions on Knowledge and Data Engineering 24 (2012), S. 236 (244) herangezogen.

<sup>60</sup> Eine einfache Zufallsstichprobe mit  $n$  Elementen aus einer  $N$ -elementigen Grundgesamtheit ist eine Stichprobe, in der jede  $n$ -elementige Teilmenge der Grundgesamtheit mit gleicher Wahrscheinlichkeit gezogen wird. Da es  $\binom{N}{n}$  solcher Teilmengen gibt, wird jede mögliche Stichprobe mit einer Wahrscheinlichkeit von  $\frac{1}{\binom{N}{n}}$  gezogen und jedes Element der Grundgesamtheit gelangt mit einer Wahrscheinlichkeit von  $\frac{1}{N}$  in die Stichprobe; siehe etwa *Kauermann und Küchenhoff*, Stichproben, 2011, S. 12–16.

<sup>61</sup> Eine Aufstellung zu den Ergebnissen der Stichprobenüberprüfung findet sich im Appendix (S. 331), eine detaillierte Dokumentation ist im Online-Appendix abrufbar. Im Online-Appendix sind auch die Ergebnisse der alternativen Zählung zusammengestellt, bei der Fundstellen in Orientierungssätzen nicht als gewünschte Zitate behandelt und daher als *false positives* bewertet werden, sofern sie im Rahmen des Extraktionsverfahrens gefunden wurden.

<sup>62</sup> Dazu wurden aus der nach Entscheidungsdatum geordneten Liste der Dateinamen im BVerfGE-Korpus mithilfe eines Zufallsgenerators 33 Entscheidungen ausgewählt. Zu überprüfen waren danach 19 Entscheidungen des Ersten Senats und 14 Entscheidungen des Zweiten Senats, die insgesamt (aus Zufall) alle Jahrzehnte der BVerfG-Rechtsprechung und alle Banddekaden in BVerfGE abdecken.

<sup>63</sup> Dieses Ergebnis überrascht nicht, wurden doch die Annotationen in Orientierungssätzen explizit in die Definition von *true positives* ein- und alle invaliden Fundstellen im dritten Verfahrensschritt ausgeschlossen.

<sup>64</sup> Daraus folgt, dass der Median der Trefferquoten (und wegen der Genauigkeit von 100 % hier auch der Median der F-Maße) bei 100 % liegt.

<sup>65</sup> Das arithmetische Mittel des (ebenfalls schief verteilten) F-Maßes liegt bei 97,2 %.

<sup>66</sup> Beispielsweise findet sich im TXT-Dokument zu BVerfGE 7, 155 folgende Passage: „BVerfGE 4,115 <129>, 1 331; 4, 219 <238>“ (Rn. 27, Unterstreichung hinzugefügt). Hier wird mit dem beschriebenen Extraktionsverfahren nur die erste Fundstelle erkannt; die letzte Fundstelle hingegen nicht. Ob der unterstrichene Einschub, der genau so im *juris*-Dokument zu finden ist, eine eigene Fundstelle darstellen soll oder eine missglückte bzw. verfälschte Seitenangabe ist, lässt sich nur durch Konsultation des gedruckten BVerfGE-Bands aufklären. Dort lautet die

pen auf (*Burst-Fehler*), die durch (für den Algorithmus) unerwartete Muster in den Zitierblöcken verursacht sind. So wird in BVerfGE 60, 162 teilweise eine abkürzende Schreibweise verwendet, die zu einem vorzeitigen Abbruch der Zitierblöcke führt, etwa: „BVerfGE 3, 19 (27), 383 (394); 4, 375 (381 f.); 6, 84 (98); 12, 132 (133 f.), 135 (137); 14, 121 (135); 24, 300 (341); 30, 227 (246); 41, 399 (421)“ (Rn. 20, Unterstreichung hinzugefügt). Hier wird die Bandnummer nur einmal angegeben, wenn mehrere Entscheidungen aus demselben Band hintereinander zitiert werden. Ein anderes Beispiel ist BVerfGE 42, 143: Dort werden manche Entscheidungen mit Kurzfundstelle und Entscheidungsnamen zitiert, was ebenfalls einen vorzeitigen Abbruch der Zitierblöcke zur Folge hat: „BVerfGE 7, 198 (205ff) - Lüth -; 18, 85 (92f); 30, 173 (187f, 196f) - Mephisto -; 32, 311 (316)“ (Rn. 12, Unterstreichungen hinzugefügt). Die (wiederum schief verteilte) Anzahl der gesuchten Zitate spielt für die Anzahl der *false negatives* kaum eine Rolle (Abb. 4.5, S. 246).

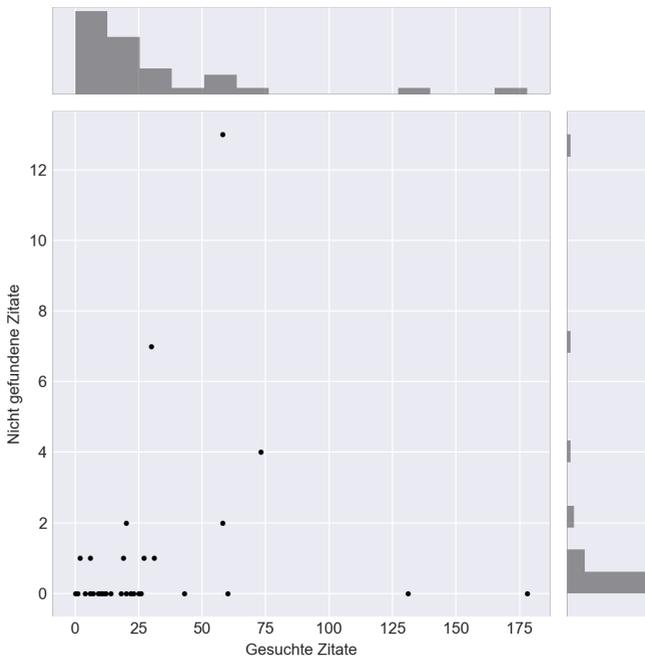


Abbildung 4.5: Gesuchte Zitate und nicht gefundene Zitate (*false negatives*) in der Stichprobe (33 Entscheidungen)

betroffene Passage „BVerfGE 4, 115 [129, 133]; 4, 219 [238]“; Ursache des *false negative* ist in diesem Fall also ein Aufbereitungsfehler im *juris*-Dokument.

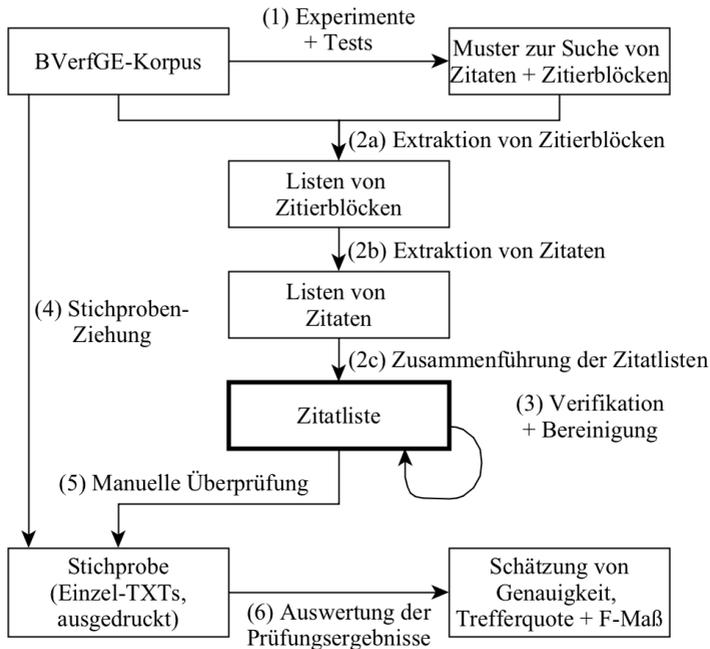


Abbildung 4.6: Vorgehen bei der Extraktion der Zitate aus den BVerfGE-Entscheidungen

Die Ergebnisse der Stichprobenüberprüfung lassen vermuten, dass die Anzahl der *false negatives* in den einzelnen Entscheidungen schief verteilt ist. Daraus folgt, dass von den Stichprobenwerten für Trefferquote und F-Maß nicht mithilfe der Standardverfahren auf die entsprechenden Werte in der Grundgesamtheit geschlossen werden kann.<sup>67</sup> Das ist allerdings für die Zwecke der Fallstudie auch nicht erforderlich, zumal aus der Qualität des Extraktionsverfahrens nicht ohne Weiteres die Qualität der Ergebnisse von Netzwerkmaßen abgeleitet werden kann.<sup>68</sup> Hier kann vielmehr festgehalten werden, dass das beschriebene Verfahren der Zitatextraktion *im Großen und Ganzen* verlässlich erscheint, wenngleich bei der Analyse die anscheinend schiefe Verteilung der *false negatives* über alle

<sup>67</sup> Zu diesen Standardverfahren etwa *Fahrmeir u.a.*, Statistik, 2016, S. 337–364. Zwar kann man auch im vorliegenden Fall ein Modell entwickeln, mit dem sich die Verteilung der Fehler in der Grundgesamtheit anhand der Ergebnisse der Stichprobe schätzen lässt. Hierzu müssten aber Annahmen zur Verteilung dieser Fehler getroffen werden, die mit großer Unsicherheit behaftet wären.

<sup>68</sup> Wie sich insbesondere fehlende Zitate auf die Ergebnisse juristischer Netzwerkstudien auswirken, ist eine offene Forschungsfrage, deren Beantwortung an dieser Stelle nicht versucht wird.

BVerfGE-Entscheidungen im Hinterkopf zu behalten ist.<sup>69</sup> Das Verfahren, mit dem die Zitatliste erstellt wurde, ist in Abbildung 4.6 (S. 247) noch einmal zusammengefasst.

### 3. Aufbereitung der Daten

Im letzten Schritt der Netzwerkkonstruktion werden Fundstellenliste und Zitatliste zu zwei Zitiergraphen zusammengesetzt, von denen einer Mehrfachkanten enthält und für die gewichtete Zitatanzahl steht (*BVerfGE-MultiDiGraph*), während der andere keine Mehrfachkanten enthält und die binäre Zitatanzahl repräsentiert (*BVerfGE-DiGraph*).<sup>70</sup> Um sicherzustellen, dass die Elemente dieser Graphen alle in der Netzwerkdefinition verlangten Merkmale enthalten, ist die Fundstellenliste allerdings zunächst um Metadaten zu ergänzen. Dabei muss insbesondere für die Entscheidungsnummern auf eine externe Quelle zurückgegriffen werden.

Als Quelle von Entscheidungsnummern kommt mittlerweile neben der Entscheidungssammlung aus dem Projekt „Deutschsprachiges Fallrecht (DFR)“ (im Folgenden kurz: DFR-Sammlung) auch die Liste der BVerfGE-Entscheidungen in Betracht, die das BVerfG auf seiner Website zur Verfügung stellt.<sup>71</sup> Letztere wird – nach umfassender Prüfung und Korrektur – für künftige Studien die bevorzugte Quelle sein; in dieser Arbeit werden die Entscheidungsnummern aus der DFR-Sammlung zugrunde gelegt, die sich aus dem HTML-Quellcode der dort bereitgestellten Webseiten auslesen lassen (Schritt 1a: Web Scraping). Nicht alle BVerfGE-Entscheidungen sind dort aufbereitet und die vergebenen Nummern sind nicht immer eindeutig; insgesamt können 1.399 unterschiedliche Entscheidungsnummern zu 1.427 Entscheidungen im BVerfGE-Korpus (darunter zwei Entscheidungen mit zweideutigen Kurzfundstellen) zugeordnet werden (Schritt 1b: Ergänzung).<sup>72</sup> Wo es im Folgenden auf Entscheidungen ankommt, die nicht in der

<sup>69</sup> Dies gilt besonders für die Analyse auf der Mikroebene, die bei den einzelnen Entscheidungen ansetzt.

<sup>70</sup> *MultiDiGraph* und *DiGraph* sind die Namen der entsprechenden Klassen in *networkx*, der Python-Bibliothek, die im Rahmen der Experimente primär zur Netzwerkanalyse verwendet wird. Das „Di“ signalisiert, dass es sich um gerichtete Graphen handelt (diese werden auch als *Digraphen* bezeichnet, siehe nur *Wasserman und Faust*, *Social Network Analysis*, 1994, S. 122), während das „Multi“ anzeigt, dass Mehrfachkanten erlaubt sind.

<sup>71</sup> Die DFR-Sammlung ist abrufbar unter [servat.unibe.ch/dfr/dfr\\_bverfg.html](http://servat.unibe.ch/dfr/dfr_bverfg.html). Zu den Problemen der Liste des BVerfG bereits oben Fn. 37 (S. 238).

<sup>72</sup> Konkret wurden im Rahmen eines ersten Scraping-Vorgangs 1.422 Entscheidungsnummern gefunden (BVerfGE 1–139); die letzten fünf Entscheidungsnummern (BVerfGE 140) wurden im November 2017 manuell hinzugefügt. Die betroffenen zweideutigen Kurzfundstellen sind BVerfGE 87, 152 – Wirtschaftskraft und BVerfGE 104, 357 – Apothekenöffnungszeiten (welcher Entscheidung im BVerfGE-Korpus die Nummern zuzuordnen waren, wurde manuell nachgeschlagen). Die mehrfach vergebenen Entscheidungsnummern sind (sortiert erst absteigend nach der in Klammern stehenden Häufigkeit und dann aufsteigend alphabetisch): „Jugendgefährdende Schriften“ (4), „Untersuchungshaft“ (3), „Sicherungsverwahrung“ (3), „Wiedergutmachung“ (2), „Wahlkampfkostenpauschale“ (2), „Vorratsdatenspeicherung“ (2), „Volksbefragung“ (2),

DFR-Sammlung enthalten sind, wird der im Einzelfall nachgeschlagene Entscheidungsname von der BVerfG-Website zugrunde gelegt und in [eckige] Klammern gesetzt (z.B. BVerfGE 107, 299 – [Handy-Überwachung]). Die Entscheidungen, die auch auf der BVerfG-Website keinen Namen tragen, erhalten von der Verfasserin einen Namen, der in {geschweifte} Klammern gesetzt wird (z.B. BVerfGE 59, 302 – {Buchführungsprivileg}).

Alle weiteren benötigten Informationen lassen sich aus den bereits in der Fundstellenliste vermerkten Zeichenketten (insbesondere den aus Entscheidungsdatum, Aktenzeichen und Kurzfundstelle bestehenden Namen der Dateien im BVerfGE-Korpus) entnehmen (Schritt 1c: Informationsextraktion).<sup>73</sup> Für die aus dem Aktenzeichen hervorgehenden Metadaten wird dabei das in *juris* erstgenannte Aktenzeichen zugrunde gelegt.<sup>74</sup>

Aus der um alle gewünschten Metadaten angereicherten Fundstellenliste, die in späteren Analysen auch *per se* nützlich sein kann und daher als CSV-Datei abgespeichert werden sollte, und der bereits alle gewünschten Metadaten enthaltenden Zitatliste kann schließlich ein gerichteter Graph mit Mehrfachkanten konstruiert werden. Dies geschieht, indem man jede Kurzfundstelle durch einen Punkt und jedes Zitat durch einen Pfeil repräsentiert (Schritt 2: Kombination).<sup>75</sup> Der so erzeugte *BVerfGE-MultiDiGraph* enthält 3.273 Punkte und 57.053 Pfeile.<sup>76</sup>

---

„Vermittlungsausschuss“ (2), „Stichtagsregelung“ (2), „Sozialplan“ (2), „Luftsicherheitsgesetz“ (2), „Lebenslange Freiheitsstrafe“ (2), „Landeslisten“ (2), „Krankenversicherung der Rentner“ (2), „Kontostammdaten“ (2), „Kindererziehungszeiten“ (2), „Häusliches Arbeitszimmer“ (2), „Feuerwehrrabgabe“ (2), „Erbchaftsteuer“ (2), „Ehrengerichte“ (2), „Ehelichkeitsanfechtung“ (2), „Ehegattensplitting“ (2), „Ehegatten-Arbeitsverhältnisse“ (2), „Asylbewerberleistungsgesetz“ (2).

<sup>73</sup> Das gilt auch für die Langfundstellen und die Endseite der Entscheidungen, die sich daraus entnehmen lässt – vorausgesetzt, man hat die „rohe“ Fundstellenliste entsprechend strukturiert. Wer dies (wie die Verfasserin) versäumt hat, muss hier wieder auf eine Datei zurückgreifen, der sich die Langfundstellen entnehmen lassen.

<sup>74</sup> Das ist nicht unproblematisch, da es durchaus vorkommt, dass Verfahren mit verschiedenen Registerzeichen oder Eingangsjahren in einer gemeinsamen Entscheidung münden. Beispielsweise werden in BVerfGE 137, 34 – Existenzsichernder Regelbedarf die Aktenzeichen 1 BvL 10/12, 1 BvL 12/12 und 1 BvR 1691/13 zusammengefasst. Etwaige Statistiken, die aus Registerzeichen und Eingangsjahren abgeleitet werden, sind daher als Schätzungen zu behandeln.

<sup>75</sup> Wird zunächst die Fundstellenliste von oben nach unten durchlaufen, um die Kurzfundstellen und die ihnen zugeordneten Metadaten als Punkte in den Graphen einzufügen, so landet bei zweideutigen Fundstellen die jeweils weiter unten stehende Fundstelle als Punkt im Graphen. Das ist erwünscht, da es sich hier stets um die Entscheidung handelt, die in BVerfGE weiter hinten abgedruckt (und in zwei von vier Fällen deutlich länger als die zuerst abgedruckte Entscheidung) ist.

<sup>76</sup> Selbst wenn man sich nicht für die Metadaten zu den einzelnen Entscheidungen interessiert, würde es für die Graphenkonstruktion nicht genügen, lediglich die Fundstellen aus der Zitatliste auszulesen. Denn es ist nicht auszuschließen, dass manche BVerfGE-Entscheidungen im BVerfGE-Zitiernetzwerk komplett isoliert sind (dass sie also weder andere Entscheidungen zitieren noch von anderen Entscheidungen zitiert werden).

Er wird immer dann benötigt, wenn die gewichtete Zählung von Zitaten in der Analyse genutzt werden soll.<sup>77</sup>

Entfernt man aus dem BVerfGE-MultiDiGraphen zunächst die Mehrfachkanten, so erhält man einen gerichteten Graphen, der ebenfalls 3.273 Punkte und alle zugehörigen Metadaten enthält, aber nur 31.373 Pfeile aufweist (Schritt 3: Entfernung von Mehrfachkanten).<sup>78</sup> Dieser Graph wird im Folgenden als *BVerfGE-DiGraph* bezeichnet. Er ist immer dann heranzuziehen, wenn die Untersuchung auf der binären Zählung von Zitaten aufbauen soll.<sup>79</sup>

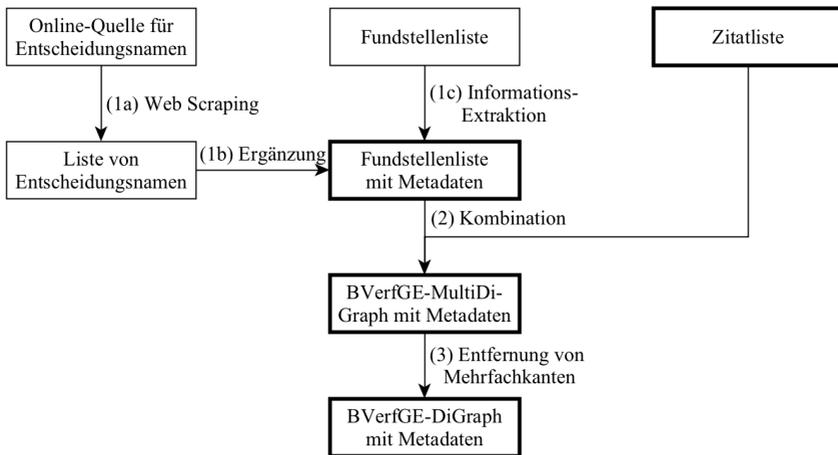


Abbildung 4.7: Vorgehen bei der Aufbereitung der Daten zu den BVerfGE-Entscheidungen

BVerfGE-MultiDiGraph und BVerfGE-DiGraph müssen in einem für Graphen geeigneten Dateiformat gespeichert werden, um jederzeit neu geladen und von verschiedenen Anwendungen genutzt werden zu können. Für diese Arbeit wurde das auf XML basierende Austauschformat GRAPHML gewählt. Abbildung 4.7 (S. 250) fasst das geschilderte Vorgehen zusammen. An seinem Ende liegen insgesamt fünf Datenquellen vor, die hier kollektiv als *BVerfGE-Daten* bezeichnet

<sup>77</sup> Da jedem einzelnen Zitat hier ein Zitierblockzeichen zugeordnet wurde, in dem potenziell interessante Informationen enthalten sind, ist es im vorliegenden Fall nicht möglich, anstelle eines Graphen mit Mehrfachkanten generell auf einen gewichteten Graphen zurückzugreifen (jedes Zitat hat eine Identität, sodass sich mehrere Zitate zwischen zwei Entscheidungen nicht ohne Informationsverlust zusammenfassen lassen). Wo einzelne Methoden allerdings für gewichtete Graphen, nicht aber für Graphen mit Mehrfachkanten verfügbar sind, kann es sich anbieten, den BVerfGE-MultiDiGraphen zum Zwecke der Analyse temporär in einen gewichteten Graphen ohne Mehrfachkanten zu überführen.

<sup>78</sup> Da sich das Zitierblockzeichen hier nicht mehr sinnvoll interpretieren lässt, sollte man dieses Attribut bei den verbleibenden Pfeilen entfernen.

<sup>79</sup> Das ist insbesondere dann der Fall, wenn die verwendeten Maße nur für ungewichtete Graphen ohne Mehrfachkanten definiert sind.

werden: Vier Dateien mit strukturierten Daten (Fundstellenliste mit Metadaten und Zitatliste als CSV-Dateien, BVerfGE-MultiDiGraph und BVerfGE-DiGraph als GRAPHML-Dateien) und das BVerfGE-Korpus als Sammlung unstrukturierter TXT-Dateien, aus denen sich bei Bedarf die Entscheidungstexte auslesen lassen.

## B. Experimente

In diesem Abschnitt wird die Rechtsprechung des BVerfG anhand der BVerfGE-Daten analysiert. Dabei ist stets zu berücksichtigen, dass diese Daten lediglich die in BVerfGE 1–140 veröffentlichte Rechtsprechung umfassen und dass im Prozess ihrer Zusammenstellung einige Entscheidungen getroffen wurden, deren Auswirkungen auf die Analyseergebnisse kaum erforscht sind.<sup>80</sup> Hinzu kommt, dass BVerfGE-Entscheidungen oft unter Rückgriff auf Textbausteine erstellt werden, die gerade auch Zitierblöcke enthalten.<sup>81</sup> Dass eine Entscheidung im konkreten Fall zitiert wird, bedeutet also nicht zwangsläufig, dass eine inhaltliche Auseinandersetzung mit ihr stattgefunden hat, und etwaige Strukturen, die sich aus den Zitaten ergeben, sind auch Produkt eines zentralisierten Wissensmanagementsystems aufseiten des BVerfG.<sup>82</sup>

Entsprechend vorsichtig müssen die Ergebnisse aller Experimente interpretiert werden.<sup>83</sup> Um verschiedene Methoden der juristischen Netzwerkforschung erpro-

<sup>80</sup> Drei dieser Entscheidungen seien nochmals hervorgehoben: Erstens wurde nicht auf die Originaltexte, sondern auf die über *juris* verfügbaren Entscheidungstexte zurückgegriffen. Zweitens wurden aus diesen Texten nicht alle Rechtsprechungszitate, sondern nur Zitate zu BVerfGE-Entscheidungen extrahiert. Drittens wurden lediglich solche Zitate erfasst, die als Fundstellenangaben in BVerfGE formuliert sind, nicht aber Zitate nach Aktenzeichen oder Zeitschriftenfundstellen.

<sup>81</sup> Bei der Erstellung der Entscheidungstexte spielen insbesondere die Wissenschaftlichen Mitarbeiter des BVerfG eine Rolle, welche das Mitglied des Gerichts, dem sie zugewiesen sind, bei dessen dienstlicher Tätigkeit unterstützen (§ 13 I BVerfGG); derzeit sind jedem Richter vier Wissenschaftliche Mitarbeiter zugeordnet. Zur Rolle der Wissenschaftlichen Mitarbeiter am BVerfG beispielsweise *Zuck*, Die Wissenschaftlichen Mitarbeiter des Bundesverfassungsgerichts, in: *van Ooyen und Möllers* (Hrsg.), Handbuch Bundesverfassungsgericht im politischen System, 2015, S. 443–457 (mit kritischer Randbemerkung zu den Textbausteinen auf S. 454).

<sup>82</sup> Damit unterscheidet sich das Rechtsprechungszitiernetzwerk des BVerfG insbesondere von den überwiegend dezentral wachsenden wissenschaftlichen Zitiernetzwerken, die (mindestens) seit *Price*, *Science* 149 (1965), S. 510–515, untersucht werden und die in der allgemeinen Netzwerkforschung besondere Aufmerksamkeit erfahren haben. Zu wissenschaftlichen Zitiernetzwerken etwa *Eom und Fortunato*, *PLoS ONE* 6 (2011), e24926-1–e24926-7; *Radicchi, Fortunato und Vespignani*, *Citation Networks*, in: *Scharnhorst, Börner und van den Besselaar* (Hrsg.), *Models of Science Dynamics*, 2012, S. 233–257; *Yan und Ding*, *Journal of the American Association for Information Science and Technology* 63 (2012), S. 1313–1326. Empirische Unterschiede zwischen wissenschaftlichen Zitiernetzwerken und Rechtsprechungszitiernetzwerken beobachten *Clough u.a.*, *Journal of Complex Networks* 3 (2015), S. 189–203.

<sup>83</sup> Nachfolgend wird grundsätzlich davon ausgegangen, dass die Zitate in den BVerfGE-Entscheidungen nicht sachfremd motiviert sind. Diese Voraussetzung ist auch dann erfüllt,

ben zu können, werden im Folgenden jeweils mehrere Experimente auf Mikro-, Meso- und Makroebene durchgeführt. Auf jeder Ebene wird die für sie abstrakt formulierte Leitfrage wieder aufgegriffen und mit Blick auf die Rechtsprechung des BVerfG als Anwendungsfall konkretisiert.<sup>84</sup> Dabei ist die Reihenfolge der Ebenen gegenüber der Anordnung in den vorigen Kapiteln modifiziert.

Die Experimente beginnen auf der Makroebene mit dem Ziel, die Rechtsprechung des BVerfG, wie sie sich in BVerfGE 1–140 darstellt, zu charakterisieren (I.). Es folgt ein Blick auf die Mikroebene mit dem Versuch, besonders wichtige BVerfGE-Entscheidungen zu identifizieren (II.). Schließlich wird auf der Mesoebene analysiert, welche BVerfGE-Entscheidungen zusammengehören (III.).

### *I. Makroebene: Was ist charakteristisch?*

Bei der Untersuchung eines Netzwerks auf der Makroebene steht die Charakterisierungsfrage (S. 73) im Vordergrund: Welche Eigenschaften sind für das betrachtete Netzwerk als Ganzes charakteristisch? Bezogen auf die Rechtsprechung des BVerfG (repräsentiert durch BVerfGE 1–140), kann man sich dieser Frage in zwei Schritten nähern. Im ersten Schritt erfolgt die Sichtung der Sammlung BVerfGE mit dem Ziel, sie zusammenzufassen (1.). Erst im zweiten Schritt geht es darum, das Zitiernetzwerk zu beschreiben (2.).

#### *1. Die Sammlung sichten: Graphiken und Kennzahlen*

Um sich einen Eindruck von der Rechtsprechung des BVerfG zu verschaffen, die in BVerfGE 1–140 veröffentlicht ist, kann man sowohl Methoden der Visualisierung als auch Methoden der Quantifizierung nutzen. Visualisierungsmethoden erlauben eine graphische Repräsentation der Sammlung, die viele Informationen inkorporiert, ohne sie zu aggregieren (a)). Demgegenüber ermöglichen Quantifizierungsmethoden eine Zusammenfassung in Zahlen, die viele Informationen aggregiert und gerade dadurch kompakt repräsentiert, wobei die Aggregationsergebnisse selbst wiederum graphisch dargestellt werden können (b)).

---

wenn die an der Formulierung der Texte beteiligten Personen auf Textbausteine zurückgreifen, ohne sich mit den darin zitierten Entscheidungen im Einzelnen auseinanderzusetzen, sofern die Textbausteine im konkreten Kontext inhaltlich nicht vollkommen unpassend sind. Denn hier geht es nicht darum, die soziologischen Prozesse offenzulegen, die der Rechtsprechung des BVerfG zugrunde liegen. Ziel ist vielmehr, das Ergebnis dieser Prozesse mit den Werkzeugen der juristischen Netzwerkforschung zu beschreiben. Dafür erweist sich die Sorgfalt, mit der das BVerfG seine Rechtsprechungslinien in Form von Zitaten dokumentiert, als äußerst hilfreich.

<sup>84</sup> Zu den einzelnen Leitfragen oben S. 71–73.

### a) Visuelle Repräsentation

Die Erkundung der BVerfGE-Daten beginnt mit einigen Fragen zu den gesammelten Entscheidungen: Wann wurden sie getroffen? Von welchen Senaten? Wie lang sind sie? Welche Registerzeichen kommen vor? Jeder Jurist, der regelmäßig mit Entscheidungen des BVerfG arbeitet, wird hier einige Intuitionen haben: Wenige Entscheidungen werden im Spätsommer veröffentlicht („Sommerloch“); Erster und Zweiter Senat sind als Spruchkörper in der Sammlung BVerfGE ähnlich stark vertreten; viele Entscheidungen sind eher kurz, einige wenige hingegen sehr lang; besonders häufig ist das Registerzeichen BvR (Verfassungsbeschwerden nach Art. 93 I Nr. 4a und 4b GG).<sup>85</sup> Viele solcher Intuitionen lassen sich bereits anhand visueller Repräsentationen der BVerfGE-Daten überprüfen, in denen jede einzelne Entscheidung dargestellt ist. Es bietet sich an, dazu mehrere verschiedene Graphiken zu nutzen, in denen die interessierenden Merkmale der Daten auf unterschiedliche visuelle Variablen abgebildet werden.<sup>86</sup>

Soll die zeitliche Verteilung der BVerfGE-Entscheidungen im Mittelpunkt stehen, so kann man zunächst das Jahr einer Entscheidung auf der x-Achse und die Position der Entscheidung innerhalb des Jahres auf der y-Achse abtragen.<sup>87</sup> Anschließend lässt sich jeder Entscheidung ein graphisches Symbol zuordnen, dessen *Form* den Spruchkörper, dessen *Größe* die Entscheidungslänge und dessen *Farbe* das Registerzeichen des (hier: in *juris*) führenden Aktenzeichens signalisiert, und die resultierenden Symbole können an die Position in der Graphik gesetzt werden, die sich aus ihrem Entscheidungsdatum ergibt. Abbildung 4.8 (S. 254) zeigt ein mögliches Ergebnis.

Die Darstellung bestätigt zunächst das „Sommerloch“, das offenbar in jüngerer Zeit zunehmend in Bedrängnis gerät. Die in BVerfGE abgedruckten Entscheidungen scheinen tendenziell länger zu werden; im Jahr 1956 sticht die Entscheidung des Ersten Senats zum KPD-Verbot heraus (BVerfGE 5, 85 – KPD-Verbot), die circa 308 Seiten lang ist. Bei den Registerzeichen dominieren visuell BvL (Normenkontrollen auf Vorlage der Gerichte nach Art. 100 I GG) und BvR (Verfassungsbeschwerden nach Art. 93 I Nr. 4a und 4b GG).<sup>88</sup>

<sup>85</sup> Eine Übersicht zu den Registerzeichen des BVerfG, die in den Statistiken des Gerichts als Aktenzeichen bezeichnet werden, findet sich im Appendix (S. 332). Nicht alle der dort aufgeführten Registerzeichen spielen in der Rechtsprechung des Gerichts eine Rolle. Die Legende der Registerzeichen, die in den nachfolgenden Abbildungen teilweise gezeigt wird, enthält lediglich Registerzeichen, die im BVerfGE-Korpus mindestens einmal als führende Aktenzeichen auftreten.

<sup>86</sup> Zu den einzelnen visuellen Variablen bereits Abb. 3.5 (S. 106) und begleitender Text.

<sup>87</sup> Zur Problematik der visuellen Darstellung temporaler Informationen bereits oben S. 110–112 sowie *Coupette und Fleckner*, JZ 73 (2018), S. 379 (387–388).

<sup>88</sup> Das ist auch deshalb bemerkenswert, weil hier auf das Registerzeichen des bei *juris* führenden Aktenzeichens abgestellt wurde, welches das in der alphabetischen Ordnung am weitesten vorne stehende zu sein scheint. Das bedeutet, dass etwa Entscheidungen wie BVerfGE 131, 316 – Landeslisten, in denen die Aktenzeichen 2 BvE 9/11, 2 BvF 3/11 und 2 BvR 2670/11 zusammen-

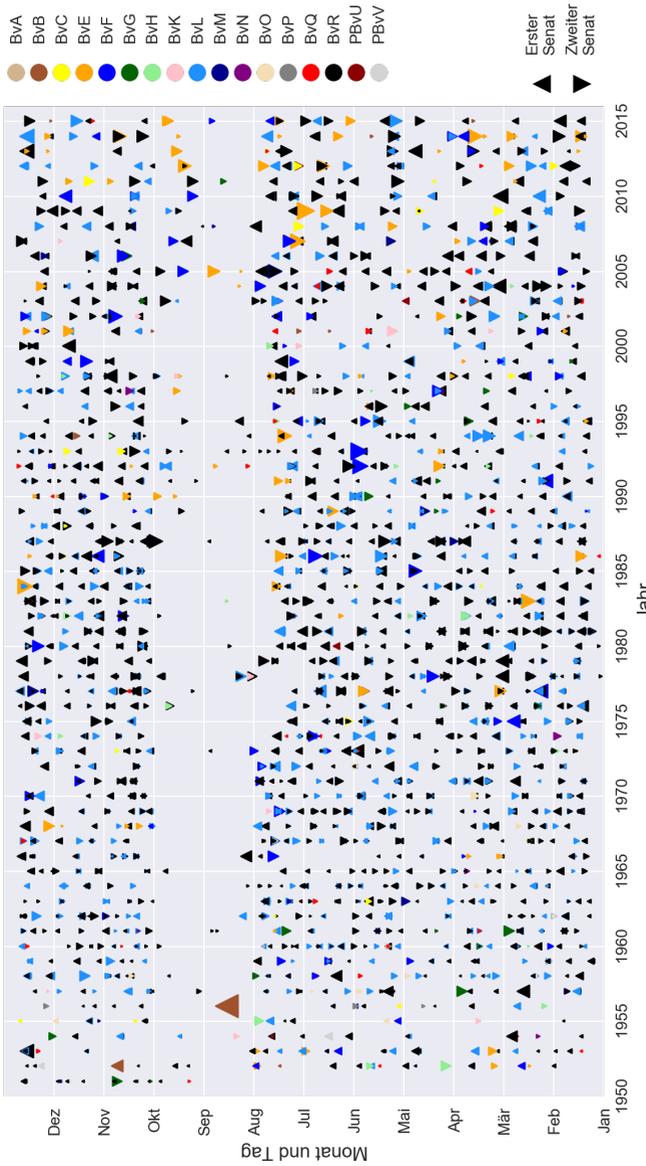


Abbildung 4.8: BVerfGE im Überblick – Betonung der zeitlichen Struktur

gefasst wurden, hier in der Farbe von BvE (Verfassungsstreitigkeiten zwischen Bundesorganen nach Art. 93 I Nr. 1 GG) auftauchen.

Obwohl Abbildung 4.8 (S. 254) viele einzelne Informationen entnommen werden können, reicht sie allein nicht aus, um einen umfassenden Eindruck von den BVerfGE-Daten zu vermitteln. So liegen die Symbole der Entscheidungen, die in geringem zeitlichen Abstand voneinander getroffen wurden, teilweise übereinander, wodurch beispielsweise die einzige Entscheidung ohne Senatszuordnung im Aktenzeichen (hier durch einen Kreis dargestellt), PBvV 1/51 vom 22. November 1951 (BVerfGE 1, 76 – Steuerverwaltung), kaum sichtbar ist. Die Aufteilung der Entscheidungen zwischen den Senaten und die Längenverhältnisse zwischen den Entscheidungen lassen sich anhand der Abbildung nur schwer nachvollziehen; die Struktur der Entscheidungssammlung (ihre Aufteilung in 140 Bände) kommt in der Graphik überhaupt nicht mehr vor. Es bietet sich daher an, noch weitere Visualisierungen zur Exploration heranzuziehen.

Will man statt der zeitlichen Struktur die Einteilung der Sammlung in Bände hervorheben, so besteht eine Möglichkeit darin, den Entscheidungsband auf der x-Achse und die Seitenzahl im Band auf der y-Achse abzutragen. Dann kann man jeder Entscheidung ein Rechteck zuordnen, das sich von ihrer Startseite im Band bis zu ihrer Endseite erstreckt. Bildet man dann das Registerzeichen des führenden Aktenzeichens wieder auf die *Farbe* der einzelnen Rechtecke ab (Abb. 4.9, S. 256), so zeigt sich deutlicher als zuvor, dass die Registerzeichen BvE (Verfassungsstreitigkeiten zwischen Bundesorganen nach Art. 93 I Nr. 1 GG) und – in geringerem Maße – auch BvC (Beschwerden im Wahlprüfungsverfahren nach Art. 41 II GG) in den jüngsten circa 25 Bänden mehr Raum einnehmen als in den älteren Bänden. Das Registerzeichen BvF (Normenkontrollen auf Antrag von Verfassungsorganen nach Art. 93 II Nr. 2 GG) bringt offenbar in regelmäßigen Abständen längere Entscheidungen hervor. Die letzten circa 25 Bände scheinen weniger kleinteilig strukturiert zu sein als die meisten älteren Bände (Bestärkung der Intuition, dass die in BVerfGE abgedruckten Entscheidungen länger werden). Die Entscheidung zum KPD-Verbot zeigt sich nun als langer Balken im fünften Band.

Will man schließlich die Senatsstruktur der Sammlung hervorheben, die aus Abbildung 4.9 (S. 256) nicht mehr ablesbar ist, so kann man statt der Registerzeichen den Spruchkörper in der Farbe der Rechtecke kodieren und die einzige Entscheidung ohne Senatszuordnung im Aktenzeichen rot hervorheben (Abb. 4.10, S. 257). Das Resultat hat (in den Augen der Verfasserin) jedenfalls ästhetischen Charme – und es mutet so an, als ob der Zweite Senat in den letzten circa 40 Bänden etwas mehr Platz in der Sammlung beansprucht hat als der Erste Senat. Alle drei Visualisierungen dienen vor allem dazu, die Aufmerksamkeit des Forschers auf Regelmäßigkeiten oder Unregelmäßigkeiten zu lenken, deren nähere Untersuchung sich lohnen könnte. Sind die Entscheidungen über die Zeit länger geworden – oder trägt der visuelle Eindruck, den die nicht aggregierten Daten vermitteln? Ist der Anteil des Zweiten Senats an der in BVerfGE veröffentlichten Rechtsprechung in den letzten Jahren wirklich gewachsen? Und sind die in BVerfGE veröffentlichten Entscheidungen mit den Registerzeichen BvE oder BvF

tatsächlich regelmäßig umfangreicher als Entscheidungen mit den Registerzeichen BvL oder BvR? Um diese und ähnliche Fragen zu beantworten, kann man die BVerfGE-Daten quantitativ zusammenfassen (hierzu sogleich).

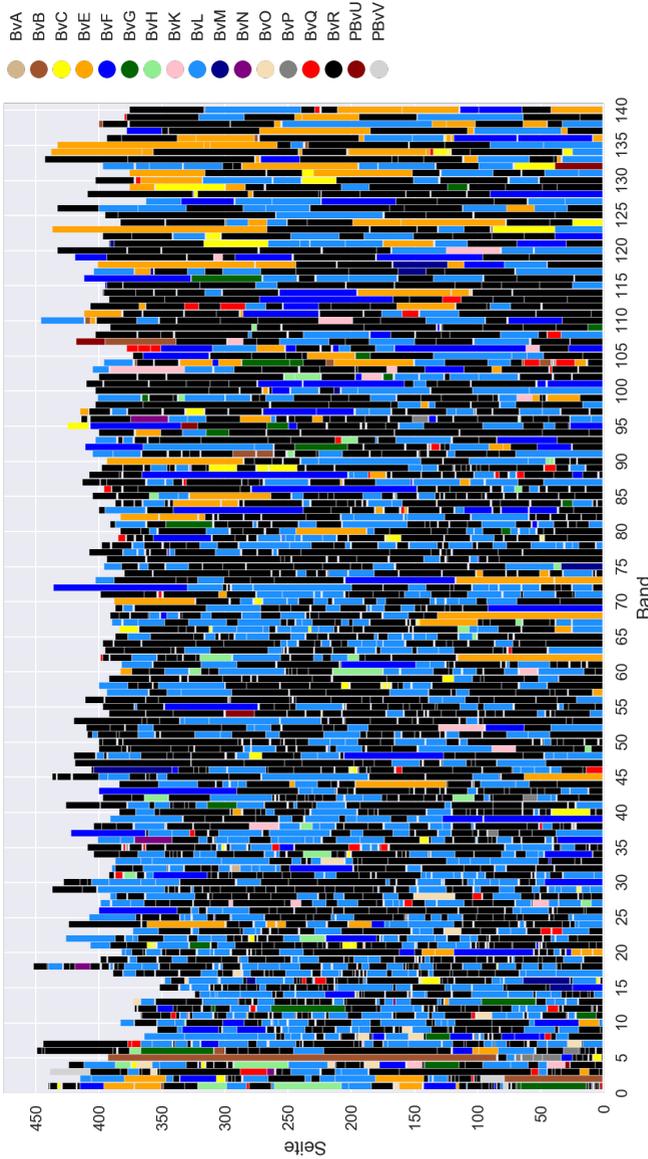


Abbildung 4.9: B VerfGE im Überblick – Betonung der Bandstruktur



Abbildung 4.10: BVerfGE im Überblick – Betonung der Senatsstruktur

*b) Quantitative Zusammenfassung*

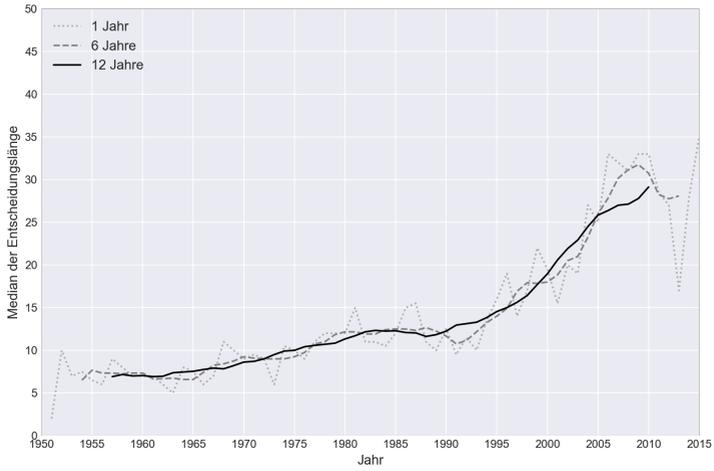
Zur quantitativen Zusammenfassung der BVerfGE-Daten kann grundsätzlich jedes Merkmal eingesetzt werden, das für die BVerfGE-Entscheidungen erhoben wurde. Hier sollen die im vorigen Abschnitt aufgeworfenen Fragen im Vordergrund stehen. Als Datengrundlage wird die um Metadaten angereicherte Fundstellenliste herangezogen.

Um die Entwicklung der Entscheidungslänge über die Zeit zu untersuchen, kann man die Entscheidungen zunächst nach ihrem Entscheidungsjahr oder ihrem Entscheidungsband gruppieren und für jede Gruppe den Median der Entscheidungslängen bilden. Sollen längerfristige Trends sichtbar werden, so bietet es sich an, für verschiedene Fenstergrößen den gleitenden Durchschnitt der ermittelten Mediane zu betrachten.<sup>89</sup> Was genau man sieht, hängt davon ab, wie man die Daten aggregiert (Abb. 4.11, S. 259). Der Befund ist aber eindeutig: Betrachtet man den Zeitraum von 1951 bis heute, so sind die in BVerfGE veröffentlichten Entscheidungen mit der Zeit deutlich länger geworden.

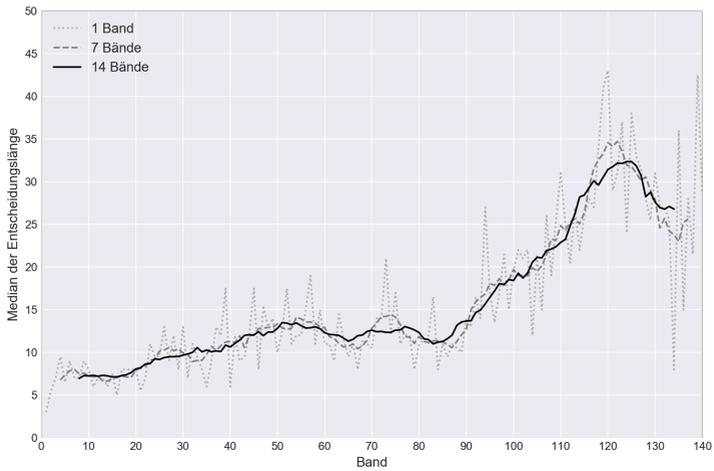
Zur Beantwortung der Frage, ob in den letzten Jahren der Zweite Senat gegenüber dem Ersten Senat in der Sammlung BVerfGE quantitativ an Bedeutung gewonnen hat, kann man die BVerfGE-Entscheidungen nach Entscheidungsjahr oder Entscheidungsband gruppieren, für jede Gruppe die Summe der Entscheidungslängen ermitteln und diese Summen jeweils bis zum Betrachtungszeitpunkt aggregieren. So erhält man für jedes Jahr bzw. jeden Band die Gesamtlänge aller bis dato in BVerfGE veröffentlichten Entscheidungen des Ersten bzw. Zweiten Senats. Stellt man die Differenz zwischen der Gesamtlänge aller BVerfGE-Entscheidungen des Ersten Senats und der Gesamtlänge aller BVerfGE-Entscheidungen des Zweiten Senats graphisch dar, so zeigt sich ein leichtes Übergewicht des Zweiten Senats ab dem Jahr 2000 bzw. ab Band 103 als Fallen der resultierenden Überschusskurve (Abb. 4.12, S. 260).

---

<sup>89</sup> Zum gleitenden Durchschnitt bereits Kapitel 3 Fn. 397 (S. 188).



(a) Nach dem Entscheidungsdatum

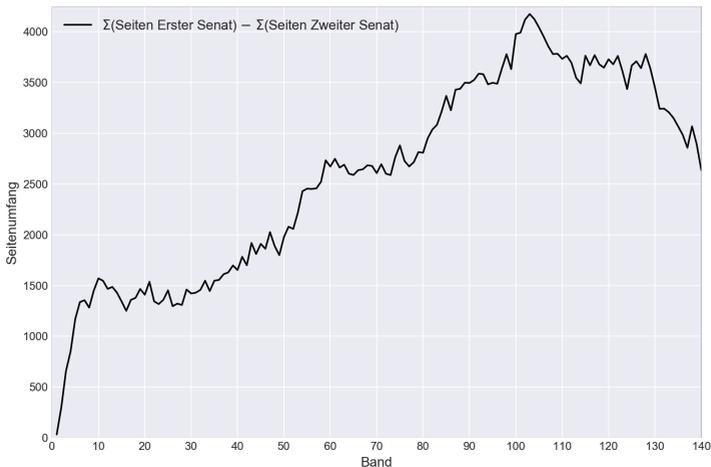


(b) Nach dem Entscheidungsband

Abbildung 4.11: Median der Entscheidungslängen in BVerfGE 1–140



(a) Nach Entscheidungsjahr



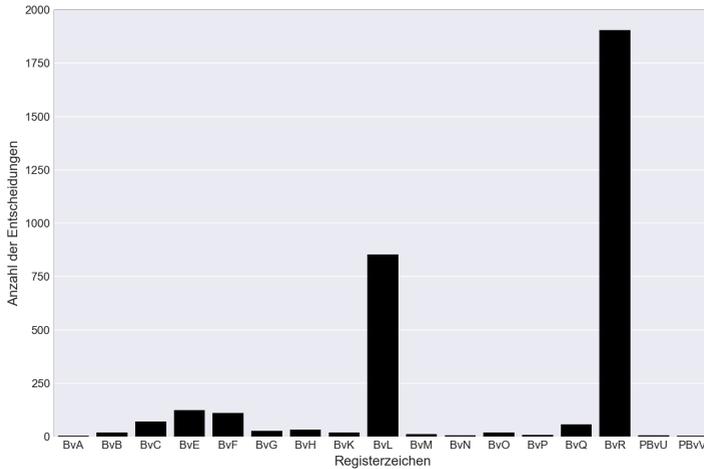
(b) Nach Entscheidungsband

Abbildung 4.12: Differenz zwischen Seitenumfang des Ersten Senats und Seitenumfang des Zweiten Senats<sup>90</sup>

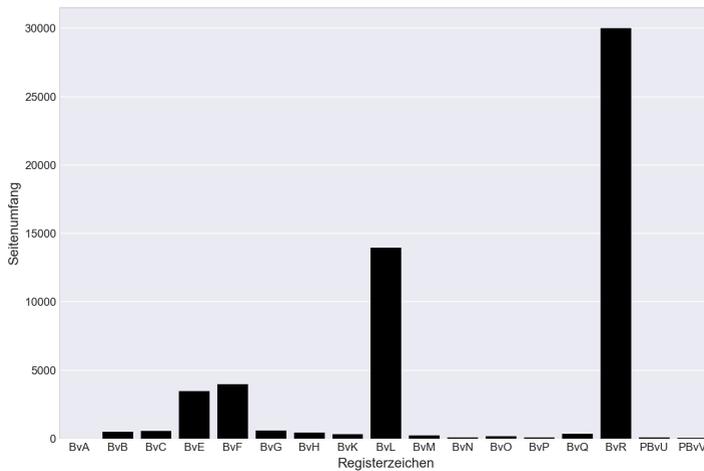
Die Dominanz von BvR und BvL unter den Registerzeichen lässt sich verifizieren, indem man wahlweise die Anzahl oder den Seitenumfang aller Entscheidungen zu einem bestimmten Registerzeichen als Balkendiagramm darstellt (Abb. 4.13, S. 261). Betrachtet man beide Balkendiagramme nebeneinander, so scheinen Ent-

<sup>90</sup> Die Abbildungen zeigen die Differenz der Seitenumfänge für alle Entscheidungen, die bis zum jeweiligen Jahr bzw. Band in BVerfGE veröffentlicht wurden.

scheidungen zu den Registerzeichen BvE und BvF im Verhältnis zu ihrer Anzahl ein größeres Seitenvolumen zu beanspruchen als viele andere Registerzeichen (allen voran – wenig überraschend – BvQ, Einstweilige Anordnungen nach § 32 BVerfGG und [bis 1970] sonstige Verfahren).



(a) Anzahl der Entscheidungen



(b) Seitenumfang der Entscheidungen

Abbildung 4.13: Registerzeichen in BVerfGE 1–140

Will man die Länge der Entscheidungen zu den verschiedenen Registerzeichen noch näher betrachten, so eröffnen *Boxplots* die Möglichkeit, viele zusammenfas-

sende Statistiken gleichzeitig zu visualisieren und dabei außerdem zwischen den Senaten zu differenzieren (Abb. 4.14, S. 262).

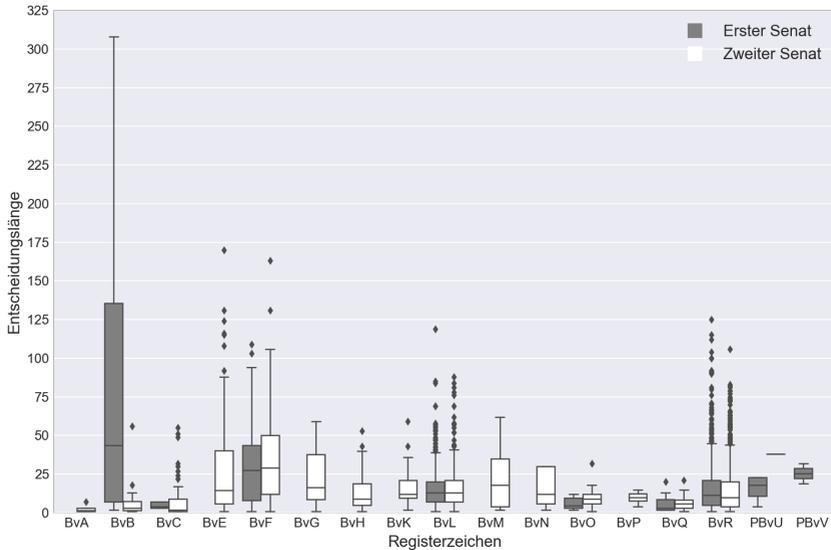


Abbildung 4.14: Entscheidungslänge nach Registerzeichen<sup>91</sup>

Dabei zeigt sich insbesondere, dass einzelne große Werte bei insgesamt kleiner Fallzahl großen Einfluss ausüben können: Die Statistik des Ersten Senats zum Registerzeichen BvB wird von der Entscheidung zum KPD-Verbot (1 BvB 2/51) dominiert. Ansonsten bestätigt sich der Verdacht, dass BvF-Entscheidungen typischerweise länger sind als Entscheidungen zu anderen Registerzeichen, während dies für BvE-Entscheidungen nicht behauptet werden kann. Unter den in der Sammlung insgesamt am stärksten vertretenen Registerzeichen BvR und BvL ist auch die Zahl der im Verhältnis zum Median nach herkömmlichen Maßstäben „atypisch“ langen Entscheidungen am größten. Die Boxplots legen in der Gesamtschau mit den Balkendiagrammen allerdings nahe, dass die Entscheidungslängen insgesamt schief verteilt sind, was ein Blick auf das entsprechende Histogramm bestätigt (Abb. 4.15, S. 263).

<sup>91</sup> Die einzelnen Boxen zeigen die *Quartile* (0,25-Quantil, 0,5-Quantil [Median] und 0,75-Quantil) für die Verteilung der Entscheidungslänge, getrennt nach dem Registerzeichen des führenden Aktenzeichens und dem Spruchkörper. Die Linien außerhalb der Boxen symbolisieren den Rest der jeweiligen Verteilung, der nicht mehr als das anderthalbfache des Interquartilsabstands vom oberen bzw. unteren Quartil der entfernt ist. Der *Interquartilsabstand* (*interquartile range*, IQR) ist die Breite des Intervalls, in dem die mittleren 50 % der beobachteten Werte liegen. Entscheidungen, die mehr als anderthalb Mal länger bzw. kürzer sind als das obere bzw. untere Quartil, sind nach herkömmlichen Maßstäben „Ausreißer“ und hier als Rauten dargestellt.

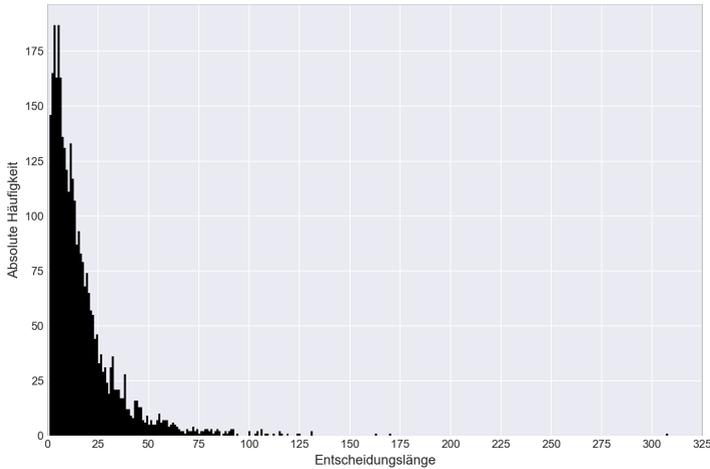


Abbildung 4.15: Schiefe Verteilung der Entscheidungslängen

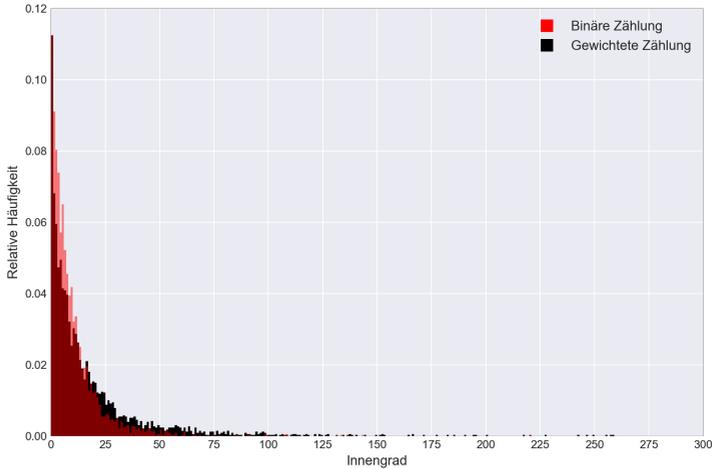
## 2. Das Zitiernetzwerk beschreiben: Verteilungen

Die bisherigen Erkundungen haben lediglich einen ersten Eindruck von den BVerfGE-Entscheidungen vermittelt. Im Zentrum der Fallstudie stehen die Zitierbeziehungen zwischen diesen Entscheidungen. Um das BVerfGE-Zitiernetzwerk auf der Makroebene zu charakterisieren, werden nun einige eindimensionale und zweidimensionale Verteilungen betrachtet.

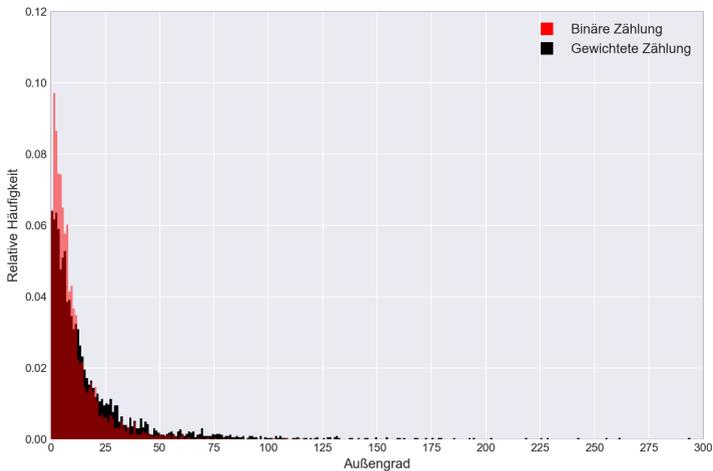
Unter den eindimensionalen Verteilungen sind vor allem die Innengradverteilung und die Außengradverteilung von Interesse, die sich sowohl für den BVerfGE-DiGraph (binäre Zählung von Zitaten) als auch für den BVerfGE-MultiDiGraph (gewichtete Zählung von Zitaten) ermitteln lassen.<sup>92</sup> Angesichts der Resultate, die für andere höchste Gerichte erzielt wurden (z.B. für den *U.S. Supreme Court*, den EuGH oder den EGMR), kann man erwarten, dass beide Verteilungen sehr schief sind: Viele Entscheidungen werden selten zitiert bzw. zitieren kaum, einige wenige Entscheidungen werden sehr häufig zitiert bzw. zitieren sehr viel. Ein Blick auf die Darstellung der Verteilungen als Histogramme bestätigt diese Erwartung (Abb. 4.16, S. 264). Damit sind die arithmetischen Mittel von Innen- bzw. Außengrad (9,59 bei binärer, 17,43 bei gewichteter Zählung) wenig aussagekräftig.

<sup>92</sup> Zur Erinnerung: Der Innengrad einer Entscheidung gibt an, wie häufig (gewichtete Zählung) bzw. von wie vielen verschiedenen Entscheidungen (binäre Zählung) sie zitiert wird. Der Außengrad einer Entscheidung gibt an, wie häufig (gewichtete Zählung) bzw. wie viele verschiedene Entscheidungen (binäre Zählung) sie zitiert. Die Innengrad- bzw. Außengradverteilung gibt für ein Netzwerk aus Entscheidungen an, welche Innen- bzw. Außengrade wie häufig vorkommen. In der Regel erfolgt die Angabe als Prozentsatz (relative Häufigkeit).

Dass Innengrad- und Außengradverteilung bei gewichteter Zählung von Zitaten noch schiefer sind als bei binärer Zählung, ist wenig überraschend.



(a) Innengradverteilung



(b) Außengradverteilung

Abbildung 4.16: Gradverteilungen der BVerfGE-Entscheidungen

Obwohl man in den Histogrammdarstellungen insgesamt wenig erkennen kann, stellt sich die Frage, auf welche Entscheidungen die seltenen, hohen Gradwerte entfallen. Für die Innengradverteilung wird diese Frage im Rahmen der Experimente

auf der Mikroebene beantwortet.<sup>93</sup> Die Entscheidungen mit einem Außengrad von über 200 bei gewichteter Zählung („über 200 Vorkommen von Zitaten“) sind BVerfGE 130, 1 – Verwertungsverbot Wohnraumüberwachung (294), BVerfGE 135, 317 – ESM-Vertrag (261), BVerfGE 135, 155 – [Filmabgabe] (255), BVerfGE 134, 141 – Beobachtung von Abgeordneten (242), BVerfGE 137, 108 – Art. 91e GG (228), BVerfGE 131, 316 – Landeslisten (225), BVerfGE 137, 273 – Katholischer Chefarzt (218) und BVerfGE 129, 208 – TKÜ-Neuregelung (202); jene mit einem Außengrad von über 100 bei binärer Zählung („über 100 verschiedene Entscheidungen werden zitiert“) sind BVerfGE 134, 141 – Beobachtung von Abgeordneten (115), BVerfGE 130, 1 – Verwertungsverbot Wohnraumüberwachung (108) und BVerfGE 137, 108 – Art. 91e GG (102).

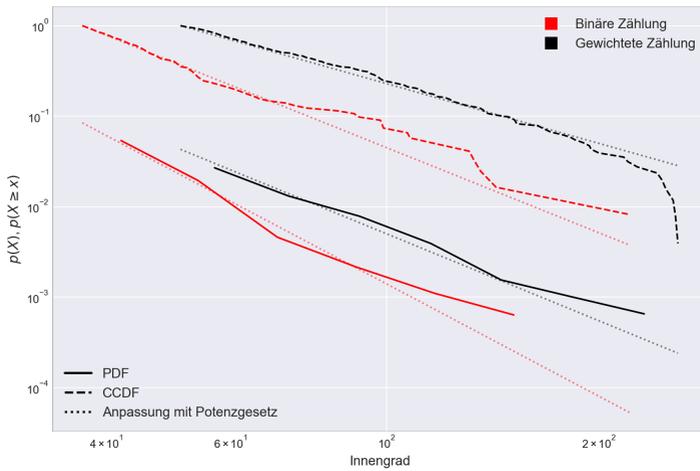
Aus den Histogrammen lässt sich über die Schiefe der abgebildeten empirischen Gradverteilungen hinaus nicht entnehmen, welcher (abstrakten) Verteilungsfunktion sie folgen. Wer sich dafür interessiert, kann auf die logarithmische Darstellung von Wahrscheinlichkeitsdichte (*probability density function*, PDF) und komplementärer kumulativer Verteilungsfunktion (*complementary cumulative distribution function*, CCDF) zurückgreifen und beispielsweise die Anpassung mit einem Potenzgesetz (*power law*) visuell darstellen (Abb. 4.17, S. 266).<sup>94</sup> Welche schiefe Verteilung (Potenzgesetz, Exponentialverteilung, Lognormalverteilung oder eine von mehreren Mischformen) auf die BVerfGE-Daten am besten passt, lässt sich allerdings auch durch quantitativen Vergleich verschiedener Anpassungen nicht ermitteln.<sup>95</sup> Für die juristische Netzwerkforschung scheint es daher sinnvoll, sich zunächst auf die Mechanismen zu konzentrieren, welche die beobachteten Gradverteilungen hervorbringen könnten. Anhand dieser Mechanismen wird sich bestimmen lassen, durch welche abstrakte Verteilung Innengradverteilung und Außengradverteilung juristischer Zitierdaten am besten beschreibbar sind.

---

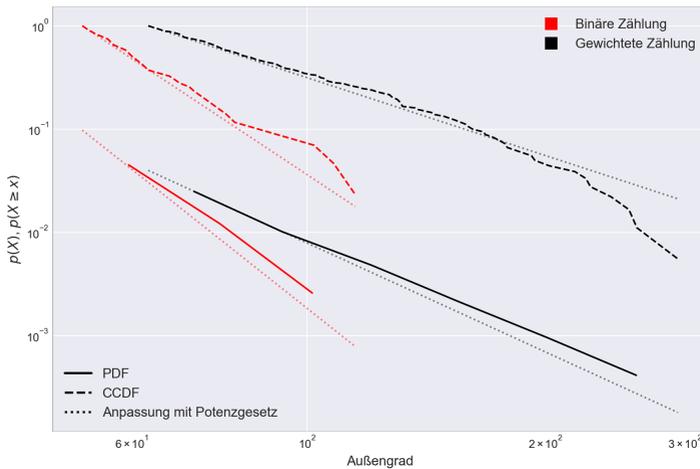
<sup>93</sup> Dazu unten S. 269–281.

<sup>94</sup> Zu den Begrifflichkeiten bereits oben bei Abb. 3.31 (S. 179).

<sup>95</sup> Einzelheiten dazu finden sich im Online-Appendix.



(a) Innengradverteilung



(b) Außengradverteilung

Abbildung 4.17: Gradverteilungen der BVerfGE-Entscheidungen in logarithmischer Darstellung und Anpassung mit Potenzgesetz<sup>96</sup>

<sup>96</sup> Wie an der Beschriftung der x-Achse abgelesen werden kann, wird die Anpassung für jede der Verteilungen erst ab einem bestimmten x-Wert ( $x_{min}$ ) durchgeführt. Dieses  $x_{min}$  ist so gewählt, dass die bestmögliche Anpassung durch ein Potenzgesetz erreicht wird, dessen Exponent ( $\alpha$ ) in der Regel mit dem  $x_{min}$  gemeinsam angegeben wird; Einzelheiten bei *Alstott, Bullmore und Plenz*, PLoS ONE 9 (2014), e85777-1 (3). Die Paare  $(x_{min}; \alpha)$  sind hier: Innengrad, binäre

Mithilfe zweidimensionaler Verteilungen lässt sich unter anderem das „Gedächtnis“ des BVerfG charakterisieren. Welche Rolle spielen ältere Entscheidungen in der Rechtsprechung des Gerichts? Übersetzt in die Sprache der Ziternetzwerke, lautet die Frage: Inwiefern werden ältere Entscheidungen in der Rechtsprechung des Gerichts zitiert? Die Zitierung älterer Rechtsprechung in jüngeren Entscheidungen ist zwar nur ein Proxy für die Wirkung älterer Entscheidungen – schließlich muss mit einem Zitat kein inhaltlicher Einfluss einhergehen und inhaltlicher Einfluss kann auch ohne Zitat vorhanden sein. Wer die Entscheidungen des BVerfG liest, kann jedoch den Eindruck gewinnen, dass das Gericht sich bemüht, durch Zitate der eigenen Rechtsprechung seine Rechtsprechungslinien nachzuzeichnen.

Diese Intuition ist zugleich eine plausible Interpretation der Bilder, die man erhält, wenn man für jedes Zitat erfasst, aus welchem Jahr zitierende und zitierte Entscheidung stammen, und zählt, wie häufig einzelne Paare von Jahren vorkommen (Vermischung der BVerfGE-Entscheidungen nach Entscheidungsjahr). Wird dann die relative Häufigkeitsverteilung der Jahrespaare betrachtet und so normalisiert, dass sie für jedes zitierende Jahr angibt, welcher Anteil aller Zitate dieses Jahres auf ein bestimmtes zitiertes Jahr entfällt, so erhält man die Matrixdarstellungen in Abbildung 4.18 (S. 267). In Abbildung 4.18 (a) (S. 267) zeigt sich deutlich, dass das BVerfG regelmäßig auch ältere Entscheidungen zitiert. Die stärkere Konzentration der Weißtöne nahe der Diagonalen in Abbildung 4.18 (b) (S. 267) legt nahe, dass in der intensiveren Auseinandersetzung (angedeutet durch mehrmalige Zitierung einzelner Fundstellen in derselben Entscheidung) eher Entscheidungen der letzten zehn bis fünfzehn Jahre relevant sind.

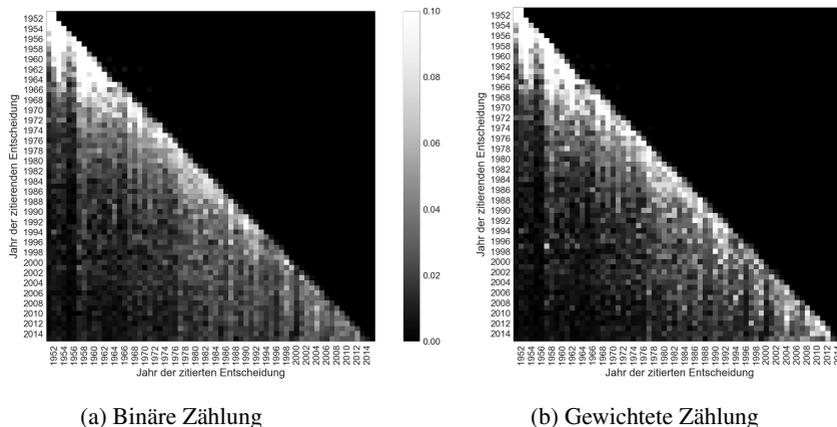
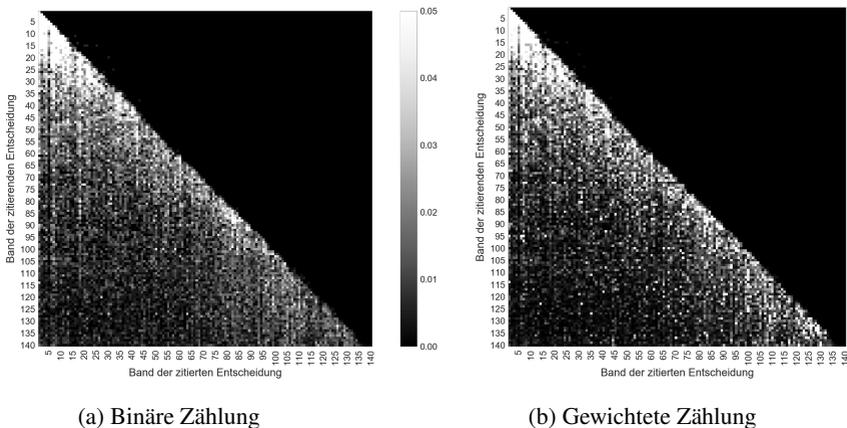


Abbildung 4.18: Vermischung zwischen BVerfGE-Entscheidungen nach Entscheidungsjahr

Zählung (37; 4, 12); Innengrad, gewichtete Zählung (51; 3, 19); Außengrad, binäre Zählung (52; 6, 08); Außengrad, gewichtete Zählung (63; 3, 50).

Da Rechtsprechungszitate regelmäßig nur in die Vergangenheit gerichtet sein können, sind die Matrizen in Abbildung 4.18 (S. 267) bei der hier gewählten Zeilen- und Spaltenanordnung annähernd untere Dreiecksmatrizen, d.h. fast alle Zellen oberhalb der Diagonale haben den Wert Null. Wer aufgrund der Abweichung in der oberen linken Ecke stutzig wird („Werden im Jahr 1951 Entscheidungen aus dem Jahr 1952 zitiert?“), kann feststellen, dass es sich bei den hierfür verantwortlichen „Zitaten“ um *juris*-Annotationen zu den Entscheidungen BVerfGE 1, 74 und BVerfGE 1, 85 handelt, die aufgrund der geringen Zitatzahl im Jahr 1951 insgesamt großes Gewicht erlangen;<sup>97</sup> ein Grund mehr, in späteren Studien nach Möglichkeit auf die Originale der Entscheidungstexte (anstatt auf *juris*) zurückzugreifen. Da die betroffenen Entscheidungen in demselben Band veröffentlicht wurden, taucht die Anomalie in der Darstellung nach Entscheidungsband nicht mehr auf (Abb. 4.19, S. 268). Ansonsten bestätigt die Darstellung nach Entscheidungsband den Eindruck der Darstellung nach Entscheidungsjahr. Für eine weiterführende Einordnung des Befundes zur Charakterisierung der Arbeitsweise des BVerfG böte es sich an, die Bilder für das BVerfG mit den „Gedächtnisstrukturen“ anderer Gerichte, etwa des *U.S. Supreme Court* oder des EuGH, zu vergleichen.



(a) Binäre Zählung

(b) Gewichtete Zählung

Abbildung 4.19: Vermischung zwischen BVerfGE-Entscheidungen nach Entscheidungsband

<sup>97</sup> Der zweite Orientierungssatz zu BVerfGE 1, 74 lautet: „Zur Hauptsacheentscheidung vgl BVerfG, 1952-05-28, 1 BvR 213/51, BVerfGE 1, 322.“ Der dritte Orientierungssatz zu BVerfGE 1, 85 lautet: „Zur Hauptsacheentscheidung vgl BVerfG, 1952-02-20, 1 BvF 2/51, BVerfGE 1, 117.“

## II. Mikroebene: Was ist wichtig?

Auf der Mikroebene steht bei der Untersuchung eines Netzwerks die Wichtigkeitsfrage im Mittelpunkt (S. 72): Welche Elemente des betrachteten Netzwerks sind besonders wichtig? Speziell für die Rechtsprechung des BVerfG (repräsentiert durch BVerfGE 1–140) kommen zwei verschiedene Interpretationen von Wichtigkeit in Betracht. Will man etwa Entscheidungen finden, die im gesamten Zitiernetzwerk eine herausgehobene Rolle spielen, so geht es um Wichtigkeit als *abstrakte* Wichtigkeit (1.). Hat man hingegen zum Ziel, Entscheidungen zu identifizieren, deren Lektüre sich in einem bestimmten Kontext (insbesondere: gegeben eine andere, als relevant vorausgesetzte Entscheidung) lohnen könnte, so steht Wichtigkeit als *konkrete* Wichtigkeit im Vordergrund (2.).<sup>98</sup>

### 1. Zentrale Entscheidungen finden: Abstrakte Wichtigkeit

Wer allgemein nach zentralen Entscheidungen in der Rechtsprechung des BVerfG sucht, geht von einer abstrakten Wichtigkeitsdefinition aus. Legt man dabei das aus BVerfGE 1–140 konstruierte Zitiernetzwerk zugrunde, so drängt sich die Frage auf: Was sind die meistzitierten Entscheidungen des BVerfG innerhalb dieses Zitiernetzwerks? Dies entspricht einer statischen Betrachtung, bei der das Zitiernetzwerk in seinem Zustand nach der Veröffentlichung von BVerfGE 140 als Momentaufnahme angesehen wird (a)). Angesichts der zeitlichen Ordnung aller Entscheidungen liegt jedoch die Vermutung nahe, dass die statische Betrachtung vor allem ältere Entscheidungen als zentral identifiziert – insbesondere bei binärer Zählung von Zitaten. Denn je älter eine Entscheidung, desto mehr nach ihr ergangene Entscheidungen haben die Möglichkeit, sie zu zitieren. Außerdem ist die Grenze „BVerfGE 140“ willkürlich und eine Rangliste der meistzitierten Entscheidungen würde vielleicht ganz anders aussehen, wenn man die Betrachtung bei BVerfGE 100 oder BVerfGE 120 stoppte. Angesichts dieser Schwierigkeiten kann sich eine dynamische Betrachtung lohnen, die mit zeitabhängigen Maßen arbeitet, deren Werte für einzelne Entscheidungen über die Zeit nachverfolgt werden können (b)).

---

<sup>98</sup> Die Experimente konzentrieren sich für beide Interpretationen von Wichtigkeit auf die Identifikation wichtiger BVerfGE-Entscheidungen (d.h. Objekte). Man kann auch nach wichtigen Zitaten (d.h. Beziehungen) oder – sofern entsprechende Daten vorliegen – nach wichtigen Zitierblöcken suchen.

a) *Statische Betrachtung: Meistzitierte Entscheidungen*

Die Frage nach den meistzitierten Entscheidungen lässt sich unterschiedlich beantworten, je nachdem, ob man die binäre oder die gewichtete Zählung von Zitaten zugrunde legt.<sup>99</sup> Tabelle 4.1 (S. 270) und Tabelle 4.2 (S. 271) führen die zehn (bzw. bei Gleichstand auf Rang 10: elf) BVerfGE-Entscheidungen auf, die von den meisten verschiedenen Entscheidungen (binäre Zählung) bzw. insgesamt am häufigsten (gewichtete Zählung) in BVerfGE 1–140 zitiert sind.

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Zitate</i>	<i>Rang</i>
18, 85	Spezifisches Verfassungsrecht	221	1
7, 377	Apotheken-Urteil	143	2
7, 198	Lüth	136	3
1, 14	Südweststaat	133	4
8, 274	Preisgesetz	131	5
6, 32	Elfes	108	6
6, 55	Steuersplitting	108	6
50, 290	Mitbestimmung	107	8
1, 97	Hinterbliebenenrente I	99	9
1, 208	7,5%-Sperrklausel	98	10
4, 7	Investitionshilfe	98	10

Tabelle 4.1: Die meistzitierten Entscheidungen in BVerfGE (binäre Zählung)

Den Spitzenplatz in beiden Ranglisten hätten viele Verfassungsrechtler wahrscheinlich erraten: Die Entscheidung BVerfGE 18, 85 – Spezifisches Verfassungsrecht, die zu den Grenzen der Nachprüfung gerichtlicher Entscheidungen durch das BVerfG immer wieder zitiert wird.<sup>100</sup> Im Apotheken-Urteil (BVerfGE 7, 377) macht das BVerfG grundlegende Ausführungen zur Berufsfreiheit und entwickelt unter anderem die Drei-Stufen-Theorie (Rn. 74–82); Lüth ist ein Klassiker zur Meinungsfreiheit und zur mittelbaren Drittwirkung der Grundrechte. BVerfGE 1, 14 – Südweststaat ist das erste Urteil des BVerfG (gefällt vom Zweiten Senat am 23. Oktober 1951) und betrifft die Kompetenzen der Länder, enthält aber auch allgemeine Aussagen zur Position des BVerfG und findet sich außerdem nur bei binärer Zählung in den „Top 10“. <sup>101</sup> Mit BVerfGE 8, 274 – Preisgesetz etabliert das BVerfG die heutige Lesart des Bestimmtheitsgebots für die Delegation von Rechtssetzungsgewalt an die Exekutive (Art. 80 I GG); BVerfGE 100,

<sup>99</sup> Zur Erinnerung: Bei binärer Zählung von Zitaten ermittelt man für jede Entscheidung lediglich, ob diese eine andere Entscheidung zitiert bzw. von dieser zitiert wird. Bei gewichteter Zählung hält man zusätzlich fest, wie häufig eine Entscheidung eine andere zitiert bzw. von einer anderen Entscheidung zitiert wird.

<sup>100</sup> Zentral sind die Rn. 20–22, in denen das Gericht seinen Maßstab für die Überprüfung einfachgerichtlicher Entscheidungen entwickelt.

<sup>101</sup> Die zehn Entscheidungen, die vor dem 23. Oktober 1951 getroffen wurden, sind Beschlüsse mit den Kurzfundstellen BVerfGE 1, 1; 1, 3; 1, 4; 1, 5; 1, 7; 1, 9; 1, 10; 1, 12; 1, 13; 1, 66.

313 – Telekommunikationsüberwachung I, eine wesentliche Entscheidung zum Fernmeldegeheimnis (Art. 10 GG), kommt nur bei gewichteter Zählung in der Spitzengruppe vor.

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Zitate</i>	<i>Rang</i>
18, 85	Spezifisches Verfassungsrecht	259	1
7, 377	Apotheken-Urteil	257	2
100, 313	Telekommunikationsüberwachung I	255	3
7, 198	Lüth	249	4
8, 274	Preisgesetz	246	5
1, 208	7,5%-Sperrklausel	242	6
65, 1	Volkszählung	227	7
35, 79	Hochschul-Urteil	220	8
109, 279	Großer Lauschangriff	217	9
53, 257	Versorgungsausgleich I	200	10

Tabelle 4.2: Die meistzitierten Entscheidungen in BVerfGE (gewichtete Zählung)

Generell fällt auf, dass hauptsächlich Entscheidungen aus den „frühen“ Bänden in den Tabellen vertreten sind, vor allem bei binärer Zählung. Die beiden jüngsten Entscheidungen in der Tabelle zur gewichteten Zählung betreffen Aspekte des Datenschutzes (BVerfGE 100, 313 – Telekommunikationsüberwachung I; BVerfGE 109, 279 – Großer Lauschangriff). Auch ist der Unterschied zwischen binärer Zählung und gewichteter Zählung für die Entscheidungen verschieden groß: BVerfGE 18, 85 – Spezifisches Verfassungsrecht wird von 221 verschiedenen Entscheidungen und insgesamt 259 Mal zitiert; das Apotheken-Urteil von 143 verschiedenen Entscheidungen und insgesamt 257 Mal. Das legt nahe, dass beide Entscheidungen in unterschiedlichen Situationen zitiert werden, die mit dem Entscheidungsinhalt (prozessuale Grundsätze auf der einen, materielle Grundsätze auf der anderen Seite) einhergehen könnten: Auf „spezifisches Verfassungsrecht“ wird meist beiläufig verwiesen, während jedenfalls ein Teil der Entscheidungen, die das Apotheken-Urteil zitieren, sich in der Sache mit ihm auseinanderzusetzen scheint. Diese Vermutung lädt dazu ein, die Werte der höchstrangigen Entscheidungen nach binärer und gewichteter Zählung zueinander ins Verhältnis zu setzen. Dieses *Zitierverhältnis* zeigt an, wie viel Prozent aller Zitate einer Entscheidung durch *verschiedene* andere Entscheidungen erfolgen (Abb. 4.3, S. 272).

BVerfGE	Entscheidungsname	R(b)	R(g)	Z(b) / Z(g)
18, 85	Spezifisches Verfassungsrecht	1	1	221/259 = 0,85
7, 377	Apotheken-Urteil	2	2	143/257 = 0,56
7, 198	Lüth	3	4	136/249 = 0,55
1, 14	Südweststaat	4	–	133/151 = 0,88
8, 274	Preisgesetz	5	5	131/246 = 0,53
6, 32	Elfes	6	–	108/164 = 0,66
6, 55	Steuersplitting	6	–	108/195 = 0,55
50, 290	Mitbestimmung	8	–	107/185 = 0,58
1, 97	Hinterbliebenenrente I	9	–	99/124 = 0,80
1, 208	7,5%-Sperrklausel	10	6	98/242 = 0,40
4, 7	Investitionshilfe	10	–	98/137 = 0,72
100, 313	Telekommunikationsüberwachung I	–	3	36/255 = 0,14
65, 1	Volkszählung	–	7	85/227 = 0,37
35, 79	Hochschul-Urteil	–	8	42/220 = 0,19
109, 279	Großer Lauschangriff	–	9	31/217 = 0,14
53, 257	Versorgungsausgleich I	–	10	72/220 = 0,32

Tabelle 4.3: Die meistzitierten Entscheidungen in BVerfGE (Vergleich der Zählungen)

Das Verhältnis zwischen Zitaten nach binärer Zählung und Zitaten nach gewichteter Zählung fällt unter den meistzitierten Entscheidungen sehr unterschiedlich aus. Einige Werte sind „auffällig niedrig“, etwa BVerfGE 100, 313 – Telekommunikationsüberwachung I (0,14), BVerfGE 109, 279 – Großer Lauschangriff (0,14) und BVerfGE 35, 79 – Hochschul-Urteil (0,19). Andere Werte sind „auffällig hoch“, z.B. BVerfGE 1, 14 – Südweststaat (0,88), BVerfGE 18, 85 – Spezifisches Verfassungsrecht (0,85), BVerfGE 1, 97 – Hinterbliebenenrente I (0,80). Der Befund deutet darauf hin, dass Entscheidungen, die zu prozessualen Fragen zitiert werden, im Zitiernetzwerk durch ein höheres Verhältnis von binärer zu gewichteter Zählung sichtbar werden.<sup>102</sup> Ebenso scheint es Entscheidungen zu eng umgrenzten Spezialthemen zu geben, mit denen sich einzelne spätere Entscheidungen intensiv auseinandersetzen (in Tab. 4.3: Datenschutz und Hochschulrecht). Dazwischen liegen Entscheidungen zu einzelnen Grundrechten (Apotheken-Urteil: Art. 12 GG; Lüth: Art. 5 I GG; Elfes: Art. 2 I GG; Steuersplitting: Art. 6 I GG; Mitbestimmung: Art. 9 III GG; Investitionshilfe und Versorgungsausgleich I: Art. 14 GG; Volkszählung: Art. 2 I i.V.m. 1 I GG) und zum Staatsorganisationsrecht (Preisge-

<sup>102</sup> Beispielsweise lautet der zweite Leitsatz von BVerfGE 1, 97 – Hinterbliebenenrente I: „Voraussetzung für die Zulässigkeit einer Verfassungsbeschwerde gegen ein erlassenes Gesetz ist, daß der Beschwerdeführer selbst, gegenwärtig und unmittelbar durch das Gesetz, nicht erst mittels eines Vollziehungsaktes, in einem Grundrecht verletzt ist.“ Ein Zitat zu BVerfGE 1, 97 passt bei der Prüfung der Zulässigkeit einer Verfassungsbeschwerde gegen ein Gesetz also fast immer.

setz: Gewaltenteilung; 7,5%-Sperrklausel: Wahlrecht). Sollte sich herausstellen, dass die geschilderten Tendenzen nicht nur für sehr häufig zitierte Entscheidungen beobachtbar sind, dass also Entscheidungsinhalt und Zitierverhältnis tatsächlich korrelieren, könnte das Verhältnis zwischen binärer und gewichteter Zählung zu mindest ein Anhaltspunkt zur retrospektiven Kategorisierung von Entscheidungen sein.

*b) Dynamische Betrachtung: Juristische Zitierkurven*

In Rechtsprechungsziternetzwerken, die aufgrund der zeitlichen Ordnung der Entscheidungen annähernd azyklisch sind, bevorzugt der Innengrad als Wichtigkeitsmaß ältere gegenüber jüngeren Entscheidungen.<sup>103</sup> Dieses Problem lässt sich beheben, indem man eine zeitabhängige Variante des Innengrads zur Bildung der Rangliste zugrunde legt: den temporalen Innengrad, der sowohl für die binäre als auch für die gewichtete Zählung von Zitaten ermittelt werden kann (S. 191–193). Der temporale Innengrad gibt für jede Entscheidung an, welcher Anteil aller nach ihrem Eintritt ins Netzwerk erfolgten (binär oder gewichtet gezählten) Zitate auf eine bestimmte Entscheidung entfällt, und beantwortet daher die Frage: Welcher Anteil aller im Netzwerk vorgefundenen Zitate, die auf Entscheidung X hätten gerichtet sein können, sind tatsächlich auf X gerichtet worden? So wird für jede Entscheidung der gleiche Maßstab angelegt, und alle Entscheidungen haben weitgehend dieselbe Chance, einen relativ hohen Wert zu erreichen.<sup>104</sup>

Tabelle 4.4 (S. 274) und Tabelle 4.5 (S. 275) führen für die binäre und die gewichtete Zitzählung jeweils die zehn Entscheidungen auf, die nach Hinzufügen von BVerfGE 140 zum Zitiernetzwerk unter Zugrundelegung des temporalen Innengrads am wichtigsten erscheinen.<sup>105</sup> In beiden Tabellen tauchen einige Entscheidungen auf, die bei den statischen Ranglisten nicht in Spitzenpositionen vertreten sind. Während die Zusammensetzung von Tabelle 4.4 (S. 274) nach Bänden weiterhin durchmischt ist, sind in Tabelle 4.5 (S. 275) nun ausschließlich jüngere Entscheidungen vertreten.

<sup>103</sup> Umso bemerkenswerter ist, dass BVerfGE 100, 313 – Telekommunikationsüberwachung I und BVerfGE 109, 279 – Großer Lauschangriff es bei gewichteter Zählung trotzdem in die „Top 10“ geschafft haben.

<sup>104</sup> Einen Sonderfall bilden Entscheidungen, die so jung sind, dass sie noch nicht zitiert werden konnten – sei es, weil sie den Gerichten noch nicht bekannt sind (bei einem Zitiernetzwerk mit Entscheidungen nur eines Gerichts wenig und speziell beim BVerfG kaum problematisch) oder weil sie noch keine durch die Zitatextraktion erfasste Fundstelle zugewiesen bekommen haben (im vorliegenden Fall aufgrund der Beschränkung auf BVerfGE-Fundstellen problematisch).

<sup>105</sup> Wie die Berechnung des temporalen Innengrads algorithmisch umgesetzt ist, kann im Online-Appendix nachvollzogen werden. Dabei werden die Entscheidungen nach ihrem Entscheidungsdatum und hilfsweise nach ihrem Band und ihrer Startseite aufsteigend in eine eindeutige Reihenfolge gebracht, sodass mehrere Entscheidungen mit demselben Entscheidungsdatum nicht ganz gleich behandelt werden. Die Umsetzung entspricht daher nicht vollständig der hier angegebenen Definition, die verlangt, dass mehrere Entscheidungen, die an demselben Tag veröffentlicht werden, in der Ordnung aller Entscheidungen dieselbe Position einnehmen.

Tatsächlich liegt die Annahme nahe, dass der temporale Innengrad insbesondere bei gewichteter Zählung ältere Entscheidungen *faktisch* benachteiligt – trotz seiner „Zeitneutralität“ aus mathematischer Sicht. Denn während zur Nachzeichnung von Rechtsprechungslinien auch ältere Entscheidungen relevant sind, kommen für intensive inhaltliche Auseinandersetzungen vor allem jüngere Entscheidungen in Betracht.<sup>106</sup> Ist in einem akut zu entscheidenden Fall eine jüngere Entscheidung inhaltlich einschlägig, so ist aus praktischer Sicht relativ wahrscheinlich, dass sie in der Entscheidung zum aktuellen Fall häufig zitiert wird. Da die zitierte Entscheidung jung ist, steht den vielen auf sie vereinigten Zitaten eine nur kleine Grundgesamtheit potenzieller Zitate gegenüber, sodass ihr temporaler Innengrad schnell hohe Werte annehmen kann. Soll es darum gehen, zu einem bestimmten Zeitpunkt die thematischen Schwerpunkte der Verfassungsrechtsprechung zu identifizieren, so kann dieser Effekt jedoch durchaus erwünscht sein: Zu Europa, Steuern und Datenschutz (Schwerpunkt von Tab. 4.5, S. 275) hat das BVerfG in jüngerer Zeit tatsächlich ausführliche Stellungnahmen abgegeben.<sup>107</sup>

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Temporal</i>
18, 85	Spezifisches Verfassungsrecht	0,00752
117, 1	Erbschaftsteuer	0,00530
123, 267	Lissabon	0,00518
105, 73	Pensionsbesteuerung	0,00481
50, 290	Mitbestimmung	0,00475
7, 377	Apotheken-Urteil	0,00464
126, 400	Steuerliche Diskriminierung eingetragener Lebenspartnerschaften	0,00441
7, 198	Lüth	0,00440
65, 1	Volkszählung	0,00438
8, 274	Preisgesetz	0,00427

Tabelle 4.4: Die meistzitierten Entscheidungen in BVerfGE nach temporalem Innengrad (binäre Zählung)

<sup>106</sup> Diese Tendenz zeigt sich auch auf der Makroebene beim Vergleich der Vermischung zwischen BVerfGE-Entscheidungen nach dem Entscheidungsjahr bzw. -band in binärer bzw. gewichteter Zählung; dazu oben bei Abb. 4.18 (S. 267) und Abb. 4.19 (S. 268).

<sup>107</sup> Beispiele für Europa: BVerfGE 132, 195 – Europäischer Stabilitätsmechanismus; BVerfGE 134, 366 – OMT-Beschluss; BVerfGE 135, 317 – ESM-Vertrag. Beispiele für Datenschutz: BVerfGE 125, 260 – Vorratsdatenspeicherung; BVerfGE 130, 151 – Zuordnung dynamischer IP-Adressen; BVerfGE 133, 277 – Antiterrordateigesetz. Beispiele für Steuern: BVerfGE 133, 377 – Ehegattensplitting; BVerfGE 135, 126 – Degressive Zweitwohnungssteuer; BVerfGE 138, 136 – Erbschaftsteuer.

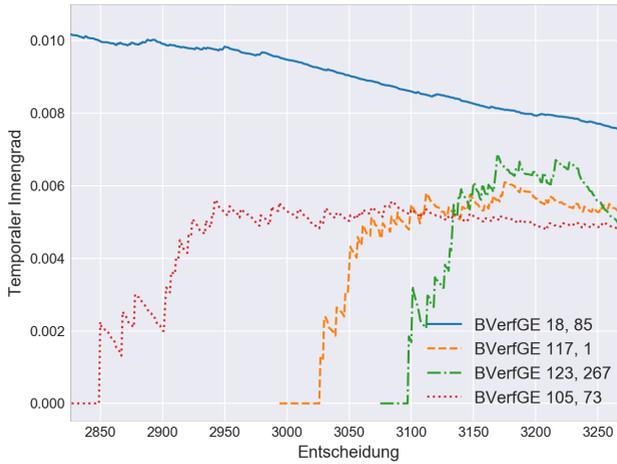
<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Temporal</i>
132, 195	Europäischer Stabilitätsmechanismus	0,01496
123, 267	Lissabon	0,01328
117, 1	Erbschaftsteuer	0,01121
129, 124	EFS	0,01086
126, 400	Steuerliche Diskriminierung eingetragener Lebenspartnerschaften	0,01061
128, 326	EGMR Sicherungsverwahrung	0,01051
109, 279	Großer Lauschangriff	0,01050
100, 313	Telekommunikationsüberwachung I	0,01013
130, 263	[W-Besoldung]	0,00881
125, 175	Hartz IV	0,00831

Tabelle 4.5: Die meistzitierten Entscheidungen in BVerfGE nach temporalem Innengrad (gewichtete Zählung)

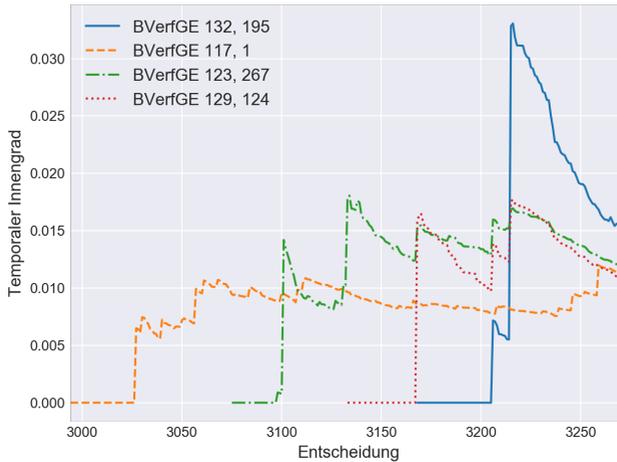
Wie schnell die Werte des temporalen Innengrads sich gerade bei jüngeren Entscheidungen verändern können, zeigt eine weitere Dynamisierung der Perspektive: Der temporale Innengrad ist nicht nur zeitabhängig definiert, sondern er verändert sich auch im Laufe der Zeit. Die Veränderungen lassen sich sichtbar machen, indem man die Entscheidungen nach ihrem Entscheidungsdatum (und bei gleichem Entscheidungsdatum nach ihrer Position in BVerfGE) ordnet, den temporalen Innengrad als Funktion der Anzahl von Entscheidungen definiert, die bereits im Netzwerk vorhanden sind, und die resultierenden *Zitierkurven* zeichnet.

Am Ende von BVerfGE 140 befinden sich 3.273 Entscheidungen im Netzwerk. Abbildung 4.20 (S. 276) zeigt die Veränderung des temporalen Innengrads für die vier Entscheidungen, die am Ende von BVerfGE 140 die höchsten Werte aufweisen, begonnen beim Eintritt der zweitältesten bzw. ältesten dieser Entscheidungen in das Zitiernetzwerk (Abb. 4.20 (a) bzw. (b), S. 276).<sup>108</sup>

<sup>108</sup> In Abbildung 4.20 (a) (S. 276) wurde auf die zweitälteste Entscheidung als untere Grenze der x-Achse abgestellt, da BVerfGE 18, 85 – Spezifisches Verfassungsrecht von den anderen Entscheidungen so weit weg liegt, dass man anderenfalls kaum etwas hätte sehen können.



(a) Binäre Zählung



(b) Gewichtete Zählung

Abbildung 4.20: Zitierkurven der vier am Ende von Band 140 nach temporalem Innengrad wichtigsten Entscheidungen<sup>109</sup>

<sup>109</sup> Die Reihenfolge der Entscheidungen in der Legende entspricht in Abbildung 4.20 (a) (S. 276) der Rangliste des temporalen Innengrads (binäre Zählung) am Ende von BVerfGE 140; in Abbildung 4.20 (b) (S. 276) sind BVerfGE 123, 267 und BVerfGE 117, 1 gegenüber der Rangliste des temporalen Innengrads (gewichtete Zählung) vertauscht, damit die Farben der Zitierkurven dieser Entscheidungen in beiden Abbildungen identisch sind. Auf eine Glättung (z.B. mithilfe eines gleitenden Durchschnitts) wurde verzichtet, da ohne sie die Zeitpunkte besser

Die unterschiedlichen Kurvenverläufe sind schon für sich genommen interessant – möglicherweise eignet sich auch die Form des Kurvenverlaufs als Anhaltspunkt für eine Entscheidungsklassifizierung. Bei der binären Zählung unterscheidet sich die Zitierkurve von BVerfGE 18, 85 – Spezifisches Verfassungsrecht deutlich von den anderen Zitierkurven; bei der gewichteten Zählung ist es die Zitierkurve von BVerfGE 117, 1 – Erbschaftsteuer. Unabhängig vom Zählmodus macht der horizontale Verlauf der Zitierkurve unmittelbar nach dem Eintritt einer Entscheidung ins Zitiernetzwerk deutlich, dass eine Entscheidung in der Regel nicht sofort, sondern erst mit gewissem zeitlichen Abstand in der nachfolgenden Rechtsprechung rezipiert wird.<sup>110</sup> Die Verläufe der Kurven bei gewichteter Zählung zeigen darüber hinaus, dass auch ähnliche Entscheidungsinhalte möglicherweise mit (partiell) synchronen Kurvenverläufen einhergehen: BVerfGE 123, 267 – Lissabon, BVerfGE 129, 124 – EFS und BVerfGE 132, 195 – Europäischer Stabilitätsmechanismus werden offenbar häufig zusammen zitiert – ihre Zitierkurven bewegen sich gemeinsam auf und ab; die lokalen Minima und Maxima befinden sich fast an denselben Stellen. Auch ohne zu wissen, welche zitierenden Entscheidungen für die abrupten Veränderungen in den Zitierkurven der genannten Entscheidungen verantwortlich sind,<sup>111</sup> kann man aus den synchronen Kurvenverläufen daher (in

---

sichtbar sind, an denen sich der temporale Innengrad der einzelnen Entscheidungen abrupt verändert.

<sup>110</sup> Dieser Befund ist zum Teil der hier gewählten Umsetzung des temporalen Innengrads geschuldet, welche die BVerfGE-Entscheidungen in eine strikte Ordnung zwingt. Auch die Beziehungsdefinition, die nur Zitate zu BVerfGE-Fundstellen erfasst, trägt dazu bei. Daneben lässt sich der Befund aber auch sachlich begründen. Zwar ist es nicht unwahrscheinlich, dass mehrere Verfahren zu ähnlichen Themen in kurzem zeitlichen Abstand zueinander beim BVerfG anhängig werden. Da das BVerfG aber mehrere Verfahren zusammenfassen kann und in bestimmten Konstellationen Kammerentscheidungen getroffen werden können, wird es vermutlich eher die Ausnahme sein, dass das Gericht mehrere thematisch eng miteinander zusammenhängende Entscheidungen *als Senatsentscheidungen* trifft. Ein Beispiel für eine solche Ausnahme zeigt sich in Abbildung 4.20 (b) (S. 276): Die unmittelbar aufeinander folgenden, steilen Anstiege der Zitierkurve von BVerfGE 132, 195 – Europäischer Stabilitätsmechanismus lassen sich auf BVerfGE 134, 366 – OMT-Beschluss und BVerfGE 135, 317 – ESM-Vertrag zurückführen. Möglicherweise kann man also auch aus der Frequenz, mit der Senatsentscheidungen zu Anstiegen bestimmter Zitierkurven führen, Rückschlüsse auf die betroffenen Themenbereiche ziehen. Dass es sich bei den Verursachern der Ausnahmesituation um Entscheidungen zu europarechtliche Fragen handelt, überrascht jedenfalls nicht.

<sup>111</sup> Welche Entscheidungen dies sind, kann man in Erfahrung bringen, indem man nachschlägt, welche Entscheidungen sich in der Ordnung nach dem Entscheidungszeitpunkt an den Stellen befinden, wo die jeweiligen Kurven besonders schnell ansteigen. Für die gemeinsamen Bewegungen von BVerfGE 123, 267 – Lissabon, BVerfGE 129, 124 – EFS und BVerfGE 132, 195 – Europäischer Stabilitätsmechanismus sind BVerfGE 134, 366 – OMT-Beschluss und BVerfGE 135, 317 – ESM-Vertrag verantwortlich (s. bereits Fn. 110, S. 277). Den ersten Anstieg bei BVerfGE 123, 267 – Lissabon verursacht BVerfGE 126, 286 – Ultra-vires-Kontrolle Honeywell. Wie sich aus Abbildung 4.20 (b) (S. 276) bereits erahnen lässt, geht der zweite Anstieg bei BVerfGE 123, 267 – Lissabon auf BVerfGE 129, 124 – EFS zurück, der dritte (zugleich erster Anstieg von BVerfGE 129, 124 – EFS) auf BVerfGE 132, 195 – Europäischer Stabilitätsmechanismus.

begrenztem Umfang) Rückschlüsse auf ihren Inhalt ziehen und Wissen, das zu einer der Entscheidungen existiert, auf die anderen Entscheidungen übertragen.<sup>112</sup>

## 2. Leseempfehlungen abgeben: Konkrete Wichtigkeit

Untersuchungen auf der Mikroebene, die eine abstrakte Definition von Wichtigkeit zugrunde legen, sind vor allem von theoretischem Interesse. In der Rechtspraxis kommt es hingegen regelmäßig darauf an, diejenigen Entscheidungen zu identifizieren, die in einem konkreten Kontext von besonderer Relevanz sind. Die juristische Netzwerkforschung kann dabei behilflich sein, indem sie die Struktur des Zitiernetzwerks im unmittelbaren Umfeld einzelner Entscheidungen für Leseempfehlungen nutzbar macht. Das Ziel ähnelt der „More like this“-Funktion, die mittlerweile von vielen wissenschaftlichen Datenbanken angeboten wird:<sup>113</sup> Gegeben eine als relevant identifizierte Gerichtsentscheidung, welche anderen Gerichtsentscheidungen könnten für den Leser interessant sein?

Zur Beantwortung dieser Frage lassen sich neben netzwerkanalytischen Methoden auch zahlreiche Werkzeuge der Textanalyse heranziehen, die dazu bestimmt sind, zueinander „ähnliche“ Dokumente zu identifizieren. Hier soll lediglich gezeigt werden, wie weit man allein auf Basis der Netzwerkstrukturen kommen kann. Ausgangspunkt ist der Befund, dass viele juristische Datenbanken in ihren Annotationen zu einer einzelnen Gerichtsentscheidung zwar aufführen, welche Entscheidungen sie zitiert und von welchen Entscheidungen sie zitiert wird, aber keinerlei „Gewichtung“ anbieten. Dabei ist klar, dass es für die potenzielle Relevanz einer anderen Entscheidung im konkreten Fall einen Unterschied macht, ob diese die betrachtete Entscheidung einmal oder zehnmal zitiert bzw. von der betrachteten Entscheidung einmal oder zehnmal zitiert wird, ob beide Entscheidungen in der späteren Rechtsprechung stets oder nie gemeinsam zitiert werden und ob beide Entscheidungen größtenteils identische oder aber sehr verschiedene andere Entscheidungen zitieren.

Hat man die Zitierdaten mit gewichteter Zählung erfasst oder sogar Daten zu Zitierblöcken erhoben, so kann man die darin enthaltenen Informationen zur Erstellung einer Rangliste von Entscheidungen nutzen, die in einem konkreten Fall (d.h. gegeben eine bestimmte andere Entscheidung) für den Leser besonders relevant sein könnten. Welche Informationen dabei in welcher Weise berücksichtigt

<sup>112</sup> Beispiel: Bekannt ist, dass BVerfGE 123, 267 – Lissabon europarechtliche Fragen behandelt. Die Zitierkurve einer anderen Entscheidung bewegt sich überwiegend synchron zu der Zitierkurve von BVerfGE 123, 267 – Lissabon. Also liegt nahe, dass auch die unbekannte Entscheidung europarechtliche Fragen behandelt.

<sup>113</sup> Beispiele: *ScienceDirect* bietet dem Leser, der Interesse an einer ausgewählten wissenschaftlichen Publikation zeigt, bis zu 100 ähnliche Artikel als Leseempfehlungen an. In der juristischen Datenbank *HeinOnline* steht dem Leser eines Artikels (z.B. in der *Law Journal Library*) ein „More Like This“-Button zur Verfügung, über den er Literaturvorschläge erhalten kann. Bei *JSTOR* kann der Nutzer sogar ein eigenes Dokument hochladen, um auf Basis des Inhalts Vorschläge zu Dokumenten zu erhalten, die ihn interessieren könnten.

werden, hängt davon ab, welche Aspekte man in den Leseempfehlungen betonen möchte. Will man für die Praxis hilfreiche Leseempfehlungen konstruieren, so ist zweitrangig, wie die Rangliste der Leseempfehlungen zustande kommt, solange sie Ergebnisse liefert, die von der Praxis als sinnvoll empfunden werden. Welche Ergebnisse dies sind, lässt sich nur in der praktischen Anwendung herausfinden; das Ranking-Verfahren kann dann in Reaktion auf Nutzerrückmeldungen angepasst werden (d.h. sich selbst anpassen oder sich anpassen lassen).

Eine mögliche Basis für spätere Optimierungen bietet ein Ranking-Verfahren, das, gegeben eine Entscheidung X, auf folgenden Informationen aufbaut: (a) den Nachbarn von X im Zitiernetzwerk (Welche Entscheidungen zitiert X, und von welchen Entscheidungen wird X zitiert?); (b) für jeden Nachbarn Y von X dessen Nachbarn im Zitiernetzwerk (Welche Entscheidungen zitiert Y, und von welchen Entscheidungen wird Y zitiert?); und (c) der Stärke der Verbindung zwischen X und jedem seiner Nachbarn Y im Zitiernetzwerk (Wie häufig wird X von Y bzw. Y von X zitiert?). Daraus lassen sich fünf Metriken zur Bewertung der Relevanz einer mit X benachbarten Entscheidung Y konstruieren:

1. die Stärke der Verbindung zwischen X und Y (unmittelbar aus (c));
2. die Zahl gemeinsamer Nachbarn von X und Y (Größe der Schnittmenge von (a) und (b));
3. der Anteil der Nachbarn von X, die auch Nachbarn von Y sind (2., geteilt durch die Anzahl der Nachbarn von X);
4. der Anteil der Nachbarn von Y, die auch Nachbarn von X sind (2., geteilt durch die Anzahl der Nachbarn von Y); und
5. der durchschnittliche Anteil gemeinsamer Nachbarn von X und Y (arithmetisches Mittel von 3. und 4.).

Für jede dieser Metriken kann man alle Nachbarn der betrachteten Entscheidung X in eine Reihenfolge bringen und für jede mit X benachbarte Entscheidung Y ihre Position in den einzelnen Reihenfolgen festhalten. Diese Positionen lassen sich zu einer Gesamtposition aggregieren, wobei die Ränge in den einzelnen Ranglisten unterschiedlich gewichtet werden können. Soll beispielsweise der Stärke der direkten Verbindung zwischen X und dem jeweiligen Y (1.) und dem Anteil der Nachbarn von X, die auch Nachbarn von Y sind (3.), eine besondere Bedeutung zukommen, so kann man die Ränge der Entscheidungen in diesen Rankings doppelt und alle anderen Ränge einfach gewichten. Indem die Summe aller gewichteten Ränge durch die Anzahl der gewichteten Ränge geteilt wird, lässt sich dann der durchschnittliche Rang einer Entscheidung ermitteln, sodass die Nachbarn von X nach ihrem Durchschnittsrang absteigend sortiert werden können.

Mit diesem System (einfache Gewichtung von 2., 4. und 5., doppelte Gewichtung von 1. und 3.) kann man etwa Lesern von BVerfGE 7, 198 – Lüth die Leseempfehlungen in Tabelle 4.6 (S. 280) unterbreiten; für BVerfGE 120, 274 – Online-Durchsuchungen sind es jene in Tabelle 4.7 (S. 280), für BVerfGE 51, 222 – 5%-Klausel jene in Tabelle 4.8 (S. 281).

<i>BVerfGE</i>	<i>Entscheidungsname</i>
93, 266	„Soldaten sind Mörder“
61, 1	Wahlkampf/CSU : NPD Europas‘
85, 1	Bayer-Aktionäre
42, 143	Deutschland-Magazin
60, 234	Kredithaie
82, 272	Postmortale Schmähkritik
66, 116	Springer/Wallraff
12, 113	Schmid-Spiegel
54, 129	Kunstkritik
10, 118	Berufsverbot I

Tabelle 4.6: Leseempfehlungen für BVerfGE 7, 198 – Lüth

<i>BVerfGE</i>	<i>Entscheidungsname</i>
113, 348	Vorbeugende Telekommunikationsüberwachung
100, 313	Telekommunikationsüberwachung I
109, 279	Großer Lauschangriff
115, 320	Rasterfahndung II
124, 43	Beschlagnahme von E-Mails
129, 208	TKÜ-Neuregelung
107, 299	[Handy-Überwachung]
125, 260	Vorratsdatenspeicherung
106, 28	Mithörrichtung
65, 1	Volkszählung

Tabelle 4.7: Leseempfehlungen für BVerfGE 120, 274 – Online-Durchsuchungen

<i>BVerfGE</i>		<i>Entscheidungsname</i>
6,	84	Sperrklausel
95,	408	Grundmandatsklausel
95,	335	Überhangmandate II
99,	1	Bayerische Kommunalwahlen
11,	351	Reserveliste Nordrhein-Westfalen
1,	208	7,5%-Sperrklausel
121,	266	Landeslisten
47,	253	Gemeindeparlamente
129,	300	Fünf-Prozent-Sperrklausel EuWG
13,	243	Wahlgebietsgröße

Tabelle 4.8: Leseempfehlungen für BVerfGE 51, 222 – 5%-Klausel

Die in den Tabellen 4.6–4.8 (S. 280–281) fokussierten Entscheidungen sind zwar alle verhältnismäßig gut in das BVerfGE-Zitiernetzwerk eingebunden,<sup>114</sup> doch gerade hier steht ein Nutzer juristischer Datenbanken regelmäßig vor dem Problem, in der Menge der annotierten Entscheidungen (oft nur auf Basis einer Fundstellenangabe oder eines Aktenzeichens) diejenigen Entscheidungen zu finden, die für ihn interessant sein könnten. Für isolierte Entscheidungen (solche, die weder zitieren noch zitiert werden) funktioniert das dargestellte Verfahren nicht; allerdings betrifft dies in der Regel nur einen Bruchteil aller Entscheidungen, für den – soweit überhaupt praktisch relevant – auf textbasierte Ähnlichkeitsmaße zurückgegriffen werden könnte.<sup>115</sup>

Inwieweit das Verfahren „in der Breite“ nützliche Resultate liefert, wird sich nur im Praxisbetrieb feststellen lassen. Dass durch die Hinzunahme weiterer Informationen aus dem Zitiernetzwerk (z.B. das Vorkommen zweier Entscheidungen in denselben Zitierblöcken), die Anpassung der Ranking-Gewichte oder die Integration netzwerk- und textbasierter Verfahren noch bessere Ergebnisse erzielt werden können, ist sehr wahrscheinlich. Die oben geschilderte Vorgehensweise bietet aber zumindest einen Ausgangspunkt für ein System, mit dem sich konkrete Leseempfehlungen für Gerichtsentscheidungen auch in juristischen Datenbanken generieren lassen könnten.

<sup>114</sup> Die Entscheidung BVerfGE 7, 198 – Lüth hat 140, die Entscheidung BVerfGE 120, 274 – Online-Durchsuchungen hat 64 und die Entscheidung BVerfGE 51, 222 – 5%-Klausel hat 52 Nachbarn im Zitiernetzwerk (d.h. Entscheidungen, die sie zitieren bzw. von ihr zitiert werden).

<sup>115</sup> Im BVerfGE-Zitiernetzwerk sind 53 Entscheidungen betroffen, also weniger als 2 % der betrachteten Entscheidungen.

### III. Mesoebene: Was gehört zusammen?

Im Zentrum der Untersuchung eines Netzwerks auf der Mesoebene steht die Gruppierungsfrage (S. 73): Welche Elemente des betrachteten Netzwerks gehören zusammen? Bezogen auf die Rechtsprechung des BVerfG (repräsentiert durch BVerfGE 1–140) lässt sich diese Frage in zwei Richtungen konkretisieren. Einerseits bietet es sich an, die BVerfGE-Entscheidungen anhand von strukturellen Kriterien zu ordnen und dabei auf die Algorithmen und mathematischen Definitionen zurückzugreifen, die im Rahmen der Quantifizierung auf der Mesoebene eingeführt wurden (1.). Andererseits kann man aber auch versuchen, inhaltliche Informationen aus den Volltexten der einzelnen Entscheidungen mit strukturellen Informationen aus Zitaten zusammenzubringen, um die Kontexte sichtbar zu machen, in denen bestimmte Stichwörter auftauchen, und die Entwicklung der zugehörigen Diskurse vor dem BVerfG nachzuverfolgen (2.).

#### 1. Rechtsprechung ordnen: Gemeinschaften in BVerfGE-Entscheidungen

Die Ordnung von Rechtsprechung ist eine traditionelle Aufgabe der Rechtswissenschaft.<sup>116</sup> In Deutschland wird diese Aufgabe bislang hauptsächlich mit den Werkzeugen der Rechtsdogmatik wahrgenommen.<sup>117</sup> Die juristische Netzwerkforschung tritt nicht an, diese Werkzeuge zu ersetzen. Sie kann allerdings rechtsdogmatische Arbeiten ergänzen, indem sie Aussagen, die von Experten auf Basis der Analyse ausgewählter Entscheidungen getroffen werden, quantitativ unterfüttert. Und sie kann diese Arbeiten vorbereiten, indem sie Gruppen von Entscheidungen identifiziert, deren Auswahl und Analyse durch Experten sich lohnen könnte. Nachfolgend werden zwei Fragestellungen behandelt, in deren Rahmen die juristische Netzwerkforschung zur Ordnung der Rechtsprechung beitragen kann: Die Suche nach juristischen Problemfeldern, also nach Themen, die immer wieder vor Gericht getragen und dort intensiv diskutiert werden (a)), und die Zuordnung von Entscheidungen zu Rechtsbereichen (b)).

##### a) Problemfelder: Gemeinschaften von $k$ -Cliques

Geht man davon aus, dass BVerfGE-Entscheidungen andere BVerfGE-Entscheidungen unter anderem deshalb zitieren, weil sie inhaltlich einschlägig sind,<sup>118</sup> so müssten sich Themen, zu denen es verhältnismäßig viele BVerfGE-Entscheidungen gegeben hat, im Zitiernetzwerk als Gruppen untereinander stark verbundener Entscheidungen zeigen. Um solche Themen – juristische Problemfelder – zu identifizieren, kann man auf das Konzept der Gemeinschaften von  $k$ -Cliques zurückgreifen.<sup>119</sup> Nach der in dieser Arbeit zugrunde gelegten

<sup>116</sup> Dazu *Bumke*, Rechtsdogmatik, 2017, S. 45–48.

<sup>117</sup> Zur Arbeitsweise der Rechtsdogmatik *Bumke*, Rechtsdogmatik, 2017, S. 113–160.

<sup>118</sup> Zu dieser Annahme bereits oben Fn. 83 (S. 251).

<sup>119</sup> Zu diesem Konzept schon S. 149–150.

Definition ist eine  $k$ -Clique ein vollständig verbundener Subgraph mit  $k$  Punkten: ein Teil eines Graphen, der aus  $k$  Punkten besteht, von denen jeder mit jedem verbunden ist. Eine Gemeinschaft von  $k$ -Cliquen ist die Vereinigung aller benachbarten  $k$ -Cliquen, d.h. die Zusammenfassung aller  $k$ -Cliquen, die ihrerseits über  $k - 1$  Punkte miteinander verbunden sind.<sup>120</sup> Je größer  $k$  gewählt wird, desto weniger Gemeinschaften von  $k$ -Cliquen sind im Zitiergraphen vorhanden und desto eher lassen sich diese Gemeinschaften im Rechtsprechungskontext als juristische Problemfelder deuten.

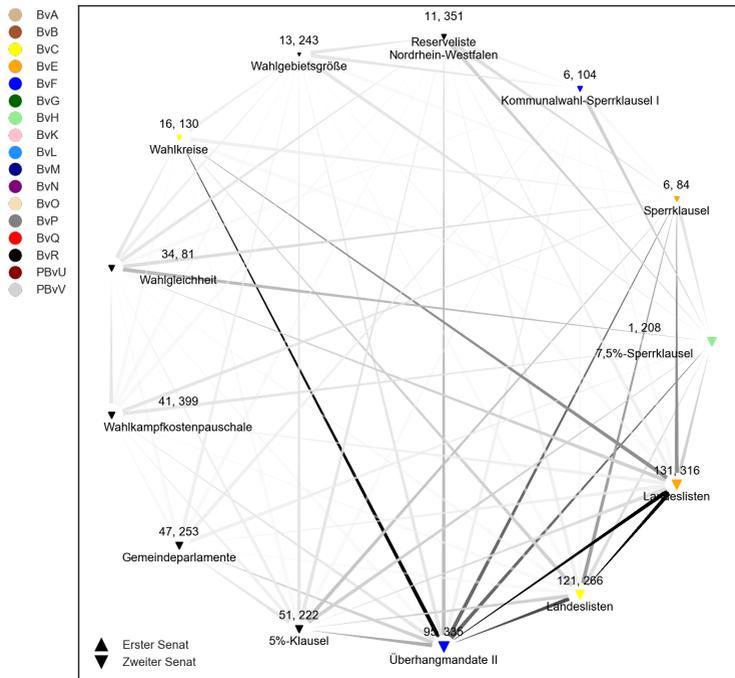
Im Zitiergraphen der BVerfGE-Entscheidungen ist 10 das größte  $k$ , für das Gemeinschaften von  $k$ -Cliquen existieren.<sup>121</sup> Konkret werden fünf Gemeinschaften von 10-Cliquen gefunden, die in den Abbildungen 4.21–4.25 (S. 284–288) dargestellt sind und im Folgenden auch kurz als Gemeinschaften bezeichnet werden.<sup>122</sup> Die Entscheidungen, die den einzelnen Gemeinschaften angehören, sind nach ihrer Position in der Sammlung BVerfGE entgegen dem Uhrzeigersinn angeordnet, beginnend mit der ältesten Entscheidung ganz rechts („im Westen“). Die Form eines Symbols zeigt den Senat an, von dem die Entscheidung getroffen wurde, die Farbe des Symbols indiziert das Registerzeichen des führenden Aktenzeichens (nach der schon für Abb. 4.8 [S. 254] und Abb. 4.9 [S. 256] genutzten Farbkarte), seine Größe steht für die Entscheidungslänge. Die Pfeile zeigen die Zitatrichtung an (vom dickeren zum dünneren Ende), die sich hier aber bereits der Anordnung entnehmen lässt; außerdem ist die Intensität des Schwarztons nach der Anzahl der Zitate zwischen den einzelnen Entscheidungen (gewichtete Zählung) linear skaliert auf dem Intervall  $[-1, 20]$ , sodass Pfeile, die lediglich einfache Zitate repräsentieren, in hellem Grau erscheinen, während Pfeile, die für 20 oder mehr Zitate stehen, schwarz sind.<sup>123</sup>

<sup>120</sup> In einem Zitiergraphen lassen sich  $k$ -Cliquen finden, indem man die Richtungen der Pfeile (Zitate) für einen Moment ignoriert und die „Schablone“ eines vollständig verbundenen Subgraphen mit  $k$  Punkten über den Zitiergraphen „rollt“; dazu *Fortunato*, *Physics Reports* 486 (2010), S. 75 (130).

<sup>121</sup> Das ist das Ergebnis des verwendeten Algorithmus.

<sup>122</sup> Die Reihenfolge der Abbildungen entspricht der Reihenfolge, in der die Gemeinschaften in einem Durchlauf des zugrunde liegenden Algorithmus gefunden wurden; sie hat keine besondere Bedeutung.

<sup>123</sup> Zwar sind Cliquen nur auf ungerichteten Graphen definiert und die Gemeinschaften von 10-Cliquen werden in der ungerichteten Version des BVerfGE-DiGraphen gesucht. Das hindert aber nicht daran, bei der Visualisierung der Resultate die Richtungsinformationen und die zusätzlichen Informationen aus dem BVerfGE-MultiDiGraphen wieder zu berücksichtigen. Die Skalierung beginnt bei  $-1$  statt bei  $0$ , da die einfachen Zitate anderenfalls kaum sichtbar sind.

Abbildung 4.21: Gleichheit bei Wahlen I<sup>124</sup>

Zwei der fünf Gemeinschaften von 10-Cliquen unter den BVerfGE-Entscheidungen, die sich teilweise überlappen (d.h. manche Entscheidungen sind in beiden Gemeinschaften vertreten), problematisieren die Gleichheit bei Wahlen. Die erste Gemeinschaft (Abb. 4.21, S. 284) enthält 13 Entscheidungen. Darunter sind neben einigen Entscheidungen zu Gleichheitsfragen, welche die Aufstellung zu Wahlen betreffen (BVerfGE 11, 351; 41, 399; 47, 253), ausschließlich Entscheidungen, die (jedenfalls auch) die Problematik der Erfolgswertgleichheit behandeln, neben BVerfGE 13, 243 vor allem ältere Entscheidungen zu Sperrklauseln (BVerfGE 1, 208; 6, 84; 6, 104; 34, 81; 51, 222) sowie Entscheidungen zu Überhangmandaten und negativem Stimmgewicht (BVerfGE 16, 130; 95, 335; 121, 266; 131, 316); Letztere nehmen vergleichsweise häufig aufeinander Bezug.

<sup>124</sup> „Landeslisten“ gehört zu den Entscheidungsamen, die in der DFR-Sammlung doppelt vergeben sind – für die hier abgebildeten BVerfGE 121, 266 und BVerfGE 131, 316. Zur Mehrfachvergabe von Entscheidungsamen bereits Fn. 72 (S. 248).

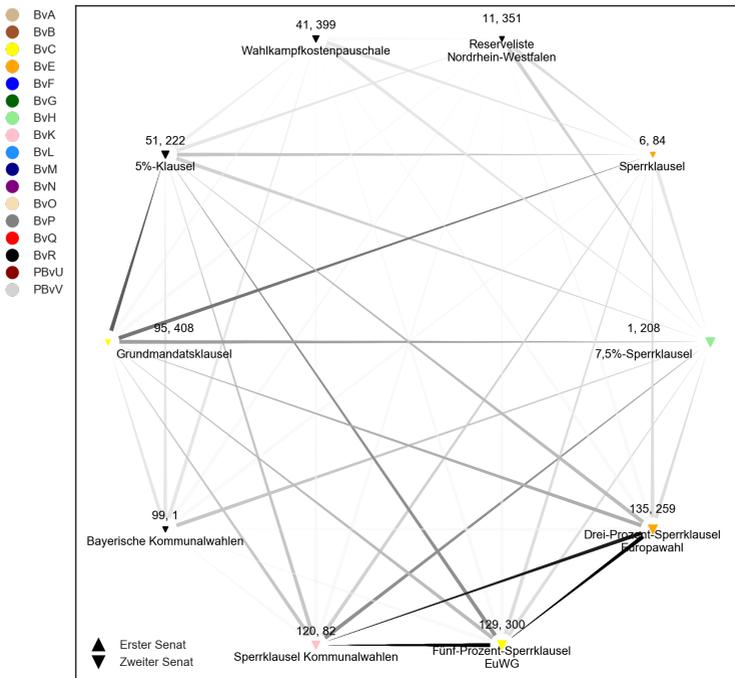


Abbildung 4.22: Gleichheit bei Wahlen II

Die zehn Entscheidungen in der zweiten Gemeinschaft (Abb. 4.22, S. 285) betreffen fast ausschließlich Sperrklauseln und ihre Ausnahmen (BVerfGE 1, 208; 6, 84; 51, 22; 95, 408; 120, 82; 129, 300; 135, 259). Im Gegensatz zur ersten Gemeinschaft sind hier nun auch die jüngeren Entscheidungen zu Sperrklauseln vertreten, die einander besonders häufig zitieren. BVerfGE 11, 351, BVerfGE 41, 399 und BVerfGE 99, 1 thematisieren wieder Gleichheitsprobleme der Wahlaufstellung. Interessant ist, dass in beiden Gemeinschaften viele verschiedene Registerzeichen vertreten sind (BvE, BvF, BvH und BvR; in der zweiten Gemeinschaft zudem BvK). Dass alle Entscheidungen in den ersten beiden Gemeinschaften vom Zweiten Senat stammen, ist hingegen angesichts der Geschäftsverteilung wenig verwunderlich.

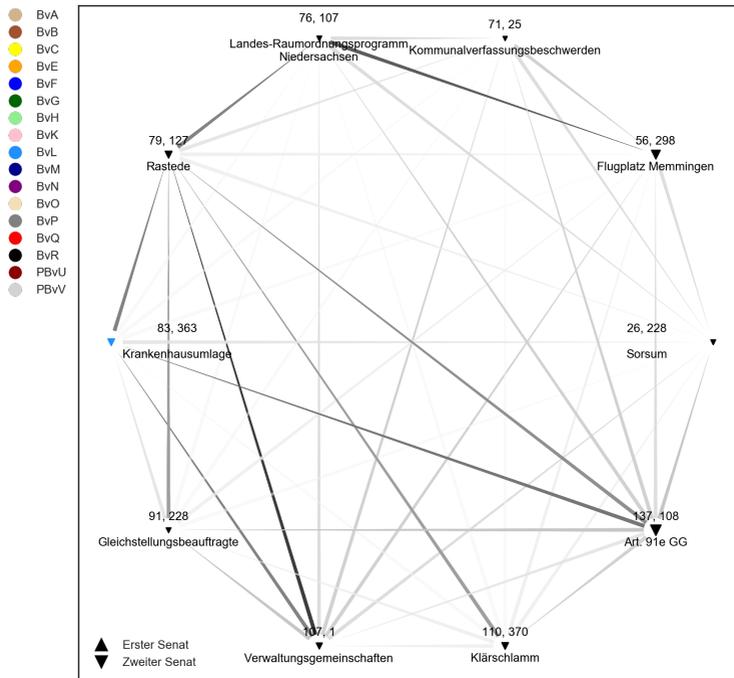


Abbildung 4.23: Kommunale Selbstverwaltung

Auch die dritte Gemeinschaft (Abb. 4.23, S. 286) enthält ausschließlich Entscheidungen des Zweiten Senats, allerdings – mit Ausnahme einer BvL-Entscheidung – nur solche mit führendem Registerzeichen BvR, also Verfassungsbeschwerden. Alle zehn Entscheidungen betreffen die in Art. 28 II GG verankerte Garantie der kommunalen Selbstverwaltung. Vergleicht man das Bild mit den Darstellungen der anderen Gemeinschaften, so sind die Entscheidungen zur kommunalen Selbstverwaltung verhältnismäßig schwach miteinander vernetzt. Ein Grund hierfür könnte sein, dass die vom BVerfG unter dem Dach des Art. 28 II GG behandelten Probleme sich inhaltlich stärker voneinander unterscheiden als die Probleme, die im Zentrum der anderen Gemeinschaften stehen, sodass weniger stark auf die in der älteren Rechtsprechung entwickelten Maßstäbe zurückgegriffen werden kann als etwa bei der Prüfung von Sperrklauseln zu Wahlen (Abb. 4.22, S. 285).

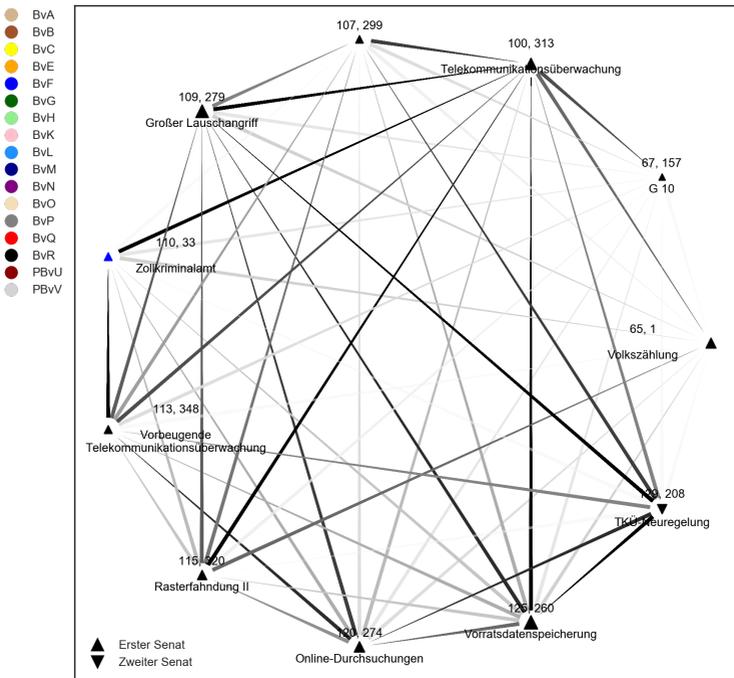


Abbildung 4.24: Datenschutz<sup>125</sup>

Die niedrige Vernetzungsintensität der Entscheidungen zur kommunalen Selbstverwaltung zeigt sich besonders deutlich im Kontrast zu den elf Entscheidungen in der vierten Gemeinschaft, die Probleme des Datenschutzes behandelt, welche vornehmlich an Art. 10 GG und Art. 2 I i.V.m. 1 I GG gemessen werden (Abb. 4.24, S. 287). Die Kombination aus intensiver Bezugnahme auf ältere Entscheidungen und (im Vergleich zu den anderen Gemeinschaften) erhöhter Entscheidungslänge deutet darauf hin, dass das BVerfG sich in Datenschutzfragen einem erhöhten Begründungsdruck ausgesetzt sieht – insbesondere dort, wo es um darum geht, die Chancen und Risiken des technologischen Fortschritts in Ausgleich zu bringen. Auffällig ist außerdem, dass es sich bei den Registerzeichen der führenden Aktenzeichen wieder fast ausschließlich um Verfassungsbeschwerden handelt (BvR). Schließlich ist bemerkenswert, dass die jüngste der vertretenen Entscheidungen, BVerfGE 129, 208 – TKÜ-Neuregelung, zugleich die einzige Entscheidung des Zweiten Senats ist, welcher die vorangegangenen Entscheidungen des Ersten Senats intensiv rezipiert.

<sup>125</sup> Die in der Graphik unbenannte Entscheidung ist BVerfGE 107, 299 – [Handy-Überwachung].

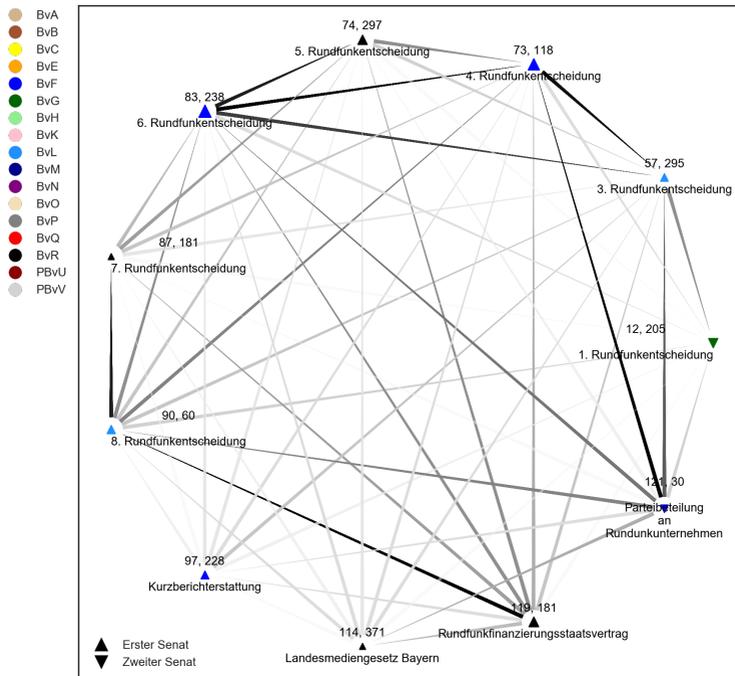


Abbildung 4.25: Rundfunkordnung

Die fünfte Gemeinschaft enthält elf Entscheidungen zur Rundfunkordnung des Grundgesetzes. Diese Entscheidungen stammen alle vom Ersten Senat – mit Ausnahme der ältesten und der jüngsten Entscheidung, die vom Zweiten Senat getroffen wurden. Mit BvF, BvG, BvL und BvR sind verhältnismäßig viele verschiedene Registerzeichen in der Gemeinschaft vertreten. Dass die (vom Zweiten Senat stammende) 2. Rundfunkentscheidung in der Abbildung fehlt, hängt auch damit zusammen, dass sie nicht von allen nachfolgenden Rundfunkentscheidungen zitiert wird.

Wie die letzte Beobachtung zeigt, ist das Konzept der Gemeinschaften von  $k$ -Cliques mit Blick auf „fehlende“ Zitate relativ streng. Es eignet sich daher allenfalls zur Identifikation von Problemfeldern, in denen eine jedenfalls formal (d.h. im Sinne der Existenz von Zitaten) kohärente Rechtsprechung existiert, und auch dort nur für Gerichte, die sehr sorgfältig zitieren. Größere Robustheit gegenüber einzelnen Zitierversäumnissen und bewussten Auslassungen bieten die heuristisch arbeitenden Verfahren der algorithmischen Gemeinschaftserkennung, mit deren Hilfe sich die Rechtsprechung des BVerfG in Rechtsbereiche einteilen lässt (hierzu sogleich).

b) *Rechtsbereiche: Algorithmische Gemeinschaftserkennung*

Um die BVerfGE-Entscheidungen einzelnen Rechtsbereichen zuzuordnen, können Verfahren zur Gemeinschaftserkennung eingesetzt werden. Von den vielen Algorithmen, die dabei in Betracht kommen,<sup>126</sup> wird hier der informationstheoretisch motivierte *Infomap*-Algorithmus erprobt, der in der juristischen Netzwerkforschung schon mehrfach zum Einsatz gekommen ist.<sup>127</sup> Dieser Algorithmus lässt sich auf gerichtete und ungerichtete Graphen anwenden, die auch gewichtet sein können. Für die Strukturierung der BVerfGE-Entscheidungen stellt sich daher zunächst die Frage, welche Version des Zitiergraphen zur Gemeinschaftserkennung zu verwenden ist. Der *Infomap*-Algorithmus simuliert die Bewegung eines Zufallsläufers auf dem Graphen, den er als Eingabe erhält. Aufgrund der Annotationen in den juristischen Datenbanken können sich Juristen bei der Recherche in beide Richtungen über Zitate zwischen Gerichtsentscheidungen bewegen. Die vom Algorithmus simulierte Zufallsbewegung lässt sich daher genau dann als juristische Recherche interpretieren, wenn man die ungerichtete Version eines Zitiergraphen zugrunde legt. Hier wird daher von der ungerichteten Version des BVerfGE-DiGraphen ausgegangen.<sup>128</sup> Sie lässt sich um die Zählinformationen des BVerfGE-MultiDiGraphen anreichern, indem die Häufigkeit jedes Pfeils im BVerfGE-MultiDiGraphen als Kantengewicht im ungerichteten Graphen repräsentiert wird. Im Folgenden wird also auf einem ungerichteten Graphen *ohne* Mehrfachkanten, aber *mit* Kantengewichten gearbeitet.<sup>129</sup>

<sup>126</sup> Dazu oben S. 147–159.

<sup>127</sup> *Derlén und Lindholm*, Europarättslig Tidskrift 2012, S. 462 (466–472); *Derlén u.a.*, Europarättslig Tidskrift 2013, S. 517 (524–535); *Derlén und Lindholm*, European Law Journal 20 (2014), S. 667 (677–679) – dort referenziert unter dem Namen der zugrunde liegenden Funktion (*Map Equation*).

<sup>128</sup> Bei der Umwandlung des gerichteten in einen ungerichteten Graphen gehen teilweise „Zitate“ verloren, die im gerichteten Graphen Teile von Zyklen sind. Aufgrund der Zeitabhängigkeit von Zitaten sollte es eigentlich keine Zyklen geben. Diese können aber unter anderem aus *juris*-Annotationen resultieren – etwa wenn zu einer Entscheidung im Eilverfahren die Fundstelle der Hauptsacheentscheidung hinzugefügt wird. Während der BVerfGE-DiGraph 3.273 Punkte und 31.373 Pfeile hat, weist dessen ungerichtete Version lediglich 31.347 Linien auf. Diese Abweichung im Promillebereich sollte allerdings die Ergebnisse eines Algorithmus zur Gemeinschaftserkennung nicht wesentlich beeinflussen.

<sup>129</sup> Wollte man den juristischen Recherchevorgang nach dem derzeitigen Stand der Technik realistisch modellieren, so müsste man auf einem Graphen arbeiten, der keine Linien mit Gewichten, sondern für jedes Zitat einen Pfeil in die Zitatrichtung und einen Pfeil in die Gegenrichtung hat. Die Pfeile in die Zitatrichtung müssten nach der Anzahl der Zitatvorkommen in der zitierenden Entscheidung gewichtet sein (auf jedes Zitatvorkommen kann man klicken, um zu der zitierten Entscheidung zu gelangen), während die Pfeile in die Gegenrichtung einheitlich das Gewicht 1 haben müssten (jede zitierende Entscheidung ist in der zitierten Entscheidung nur einmal annotiert, unabhängig davon, wie oft auf die zitierte Entscheidung Bezug genommen wird). Dem Zufallsläufer im *Infomap*-Algorithmus stehen hier also mehr Informationen zur Verfügung als dem recherchierenden Juristen: Er kennt die Intensität der Zitierbeziehung zwischen zwei

Der *Infomap*-Algorithmus ist nichtdeterministisch, d.h. mehrere Läufe können verschiedene Ergebnisse liefern. Das heißt auch, dass Entscheidungen, die in einem Lauf derselben Gemeinschaft zugeordnet werden, in anderen Läufen möglicherweise nicht als zusammengehörig identifiziert werden. Gleichzeitig gehören Entscheidungen, die in vielen Läufen derselben Gemeinschaft zugeschlagen werden, mit höherer Wahrscheinlichkeit zusammen als solche, die nur selten in dieselbe Gruppe eingeordnet werden.

Um den Nutzen der algorithmischen Gemeinschaftserkennung für die Strukturierung von BVerfGE-Entscheidungen auf Basis des *Infomap*-Algorithmus zu untersuchen, bietet sich daher ein zweischrittiges Vorgehen an. Im ersten Schritt wird das Ergebnis eines beliebigen Laufs des Algorithmus auf seine Überzeugungskraft aus juristischer Sicht inspiziert. Im zweiten Schritt werden die Ergebnisse für einhundert Läufe des Algorithmus zusammengeführt, um für jede Entscheidung diejenigen anderen Entscheidungen zu ermitteln, die besonders häufig gemeinsam eingeordnet werden.<sup>130</sup>

Führt man den *Infomap*-Algorithmus wie beschrieben auf dem ungerichteten, gewichteten Zitiergraphen aus, so erhält man zunächst für jede Entscheidung eine Zahl, welche die Gemeinschaft repräsentiert, der die Entscheidung zugeordnet worden ist. Diese Zahl hat über die Unterscheidung der einzelnen Gemeinschaften hinaus keine Bedeutung; als Attribut einer Entscheidung ist sie also ein nominales Merkmal, sodass insbesondere kein sinnvoller „Größenvergleich“ durchgeführt werden kann.<sup>131</sup> In dem Lauf, der nachfolgend näher untersucht wird, identifiziert der Algorithmus 247 Gemeinschaften, davon 189 mit zwei oder mehr Mitgliedern.<sup>132</sup> Die größte Gemeinschaft enthält 135 Entscheidungen. Wie Abbildung 4.26 (S. 291) verdeutlicht, ist die Häufigkeitsverteilung der Gemeinschaftsgrößen schief.<sup>133</sup>

---

Entscheidungen unabhängig davon, ob er sich gerade bei der zitierenden oder bei der zitierten Entscheidung befindet.

<sup>130</sup> Die Resultate des zweiten Schritts könnten sich ebenso für Leseempfehlungen fruchtbar machen lassen wie die Metriken zur Bestimmung konkreter Wichtigkeit, die auf der Mikroebene entwickelt wurden (S. 278–281).

<sup>131</sup> Eine Aussage wie „Entscheidung A ist in einer Gemeinschaft mit kleinerer Ziffer als Entscheidung B“ bedeutet daher nichts weiter als „Entscheidung A und Entscheidung B sind unterschiedlichen Gemeinschaften zugeordnet“. Zur Skalierung von Merkmalen bereits oben S. 93–99.

<sup>132</sup> Das Ergebnis dieses Laufs ist im Online-Appendix einsehbar.

<sup>133</sup> Statistische Kennzahlen: Modus: 1; Median: 7; arithmetisches Mittel: 13,25; Standardabweichung: 17,61; unteres Quartil: 2; oberes Quartil: 17.

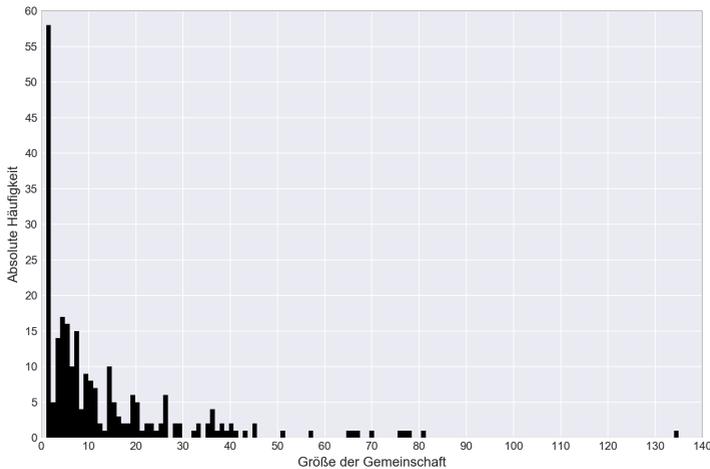


Abbildung 4.26: Häufigkeitsverteilung der Gemeinschaftsgrößen in BVerfGE 1–140

Geht man davon aus, dass die BVerfGE-Zitate vor allem inhaltlich motiviert sind, so müsste die auf ihrer Basis algorithmisch ermittelte Einteilung der BVerfGE-Entscheidungen die inhaltliche Organisation der BVerfG-Rechtsprechung widerspiegeln. Es sollte dann möglich sein, den einzelnen Gemeinschaften Namen zu geben, welche die Schwerpunkte der Entscheidungen zusammenfassen, die in ihnen enthalten sind. Bei dem Versuch, die vom *Infomap*-Algorithmus identifizierten Gemeinschaften zu benennen, erweisen sich die bereits erhobenen Entscheidungsamen als äußerst nützlich; ergänzend können die Namen von der BVerfG-Website herangezogen werden.<sup>134</sup> Für jede Gemeinschaft lässt sich außerdem ein Subgraph des ursprünglichen BVerfGE-DiGraphen erstellen, der nur diejenigen Punkte (Entscheidungen) umfasst, die in der Gemeinschaft enthalten sind. Die Summe von Innen- und Außengrad der Punkte in diesem Subgraphen ist dann ein Maß für die „Repräsentativität“ der einzelnen Entscheidungen in der Gemeinschaft. Wer die Entscheidungen nach dieser Summe absteigend ordnet und die Namen der repräsentativsten Entscheidungen betrachtet, kann viele Gemeinschaften schnell benennen. Die Tabellen 4.9–4.11 (S. 292–293) zeigen dies beispielhaft für die Gemeinschaften 5 (*Beruf* – 77 Entscheidungen), 6 (*Eigentum* – 76 Entscheidungen) und 15 (*Europa* – 23 Entscheidungen); eine Übersicht zu allen Gemeinschaften und den ihnen von der Verfasserin zugeordneten Namen findet sich im Appendix (S. 333–336); eine detaillierte Aufstellung aller Gemeinschaften und ihrer Mitglieder ist im Online-Appendix abrufbar.<sup>135</sup>

<sup>134</sup> Zu den Quellen der Entscheidungsamen bereits oben S. 248–249.

<sup>135</sup> Die Entscheidungen sind absteigend nach der Summe von Innengrad und Außengrad und dann aufsteigend nach Band und Startseite geordnet. Benannt sind alle Gemeinschaften mit fünf oder mehr Mitgliedern. Bei Gemeinschaften mit nur einem Mitglied (*singletons*) ist die

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Innen</i>	<i>Außen</i>	$\Sigma$
7, 377	Apotheken-Urteil	54	0	54
30, 292	Erdölbevorratung	26	13	39
54, 301	[Buchführungsprivileg]	8	13	21
25, 236	[Dentist]	13	7	20
75, 246	[Abschaffung des Rechtsbeistandes]	6	14	20
13, 97	Handwerksordnung	15	3	18
16, 147	Werkfernverkehr	14	4	18
77, 84	Arbeitnehmerüberlassung	8	10	18
9, 73	Arzneifertigwaren	16	1	17
59, 302	{Buchführungsprivileg}	5	12	17
121, 317	Rauchverbot in Gaststätten	3	14	17

Tabelle 4.9: Repräsentative Mitglieder der Gemeinschaft 5 (*Beruf* – 77 Entscheidungen)

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Innen</i>	<i>Außen</i>	$\Sigma$
52, 1	Kleingarten	23	12	35
24, 367	Hamburgisches Deichordnungsge- setz	24	3	27
37, 132	Vergleichsmiete I	23	4	27
58, 300	Naßauskiesung	12	15	27
42, 263	Contergan	13	12	25
21, 73	Grundstücksverkehrsgesetz	23	1	24
31, 229	Schulbuchprivileg	19	3	22
68, 361	Eigenbedarf I	9	12	21
126, 331	[Miterben-Entschädigungsfonds]	0	21	21
70, 191	Fischereibezirke	8	12	20
89, 1	Besitzrecht des Mieters	5	15	20

Tabelle 4.10: Repräsentative Mitglieder der Gemeinschaft 6 (*Eigentum* – 76 Entscheidungen)

Benennung nicht sinnvoll; bei Gemeinschaften mit zwei bis vier Mitgliedern ist sie manchmal sehr leicht (z.B. bei Gemeinschaft 104: BVerfGE 132, 134 – Asylbewerberleistungsgesetz; BVerfGE 137, 34 – Existenzsichernder Regelbedarf; BVerfGE 125, 175 – Hartz IV), manchmal nach kursorischer Inspektion der Entscheidungstexte kaum möglich.

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Innen</i>	<i>Außen</i>	$\Sigma$
73, 339	Solange II	12	3	15
123, 267	Lissabon	8	7	15
37, 271	Solange I	13	1	14
89, 155	Maastricht	10	4	14
129, 124	EFS	5	9	14
140, 317	Identitätskontrolle	0	14	14
126, 286	Ultra-vires-Kontrolle Honeywell	5	7	12
132, 195	Europäischer Stabilitätsmechanismus	3	9	12
58, 1	Eurocontrol I	9	2	11
134, 366	OMT-Beschluss	1	10	11

Tabelle 4.11: Repräsentative Mitglieder der Gemeinschaft 15 (*Europa* – 23 Entscheidungen)

Auch Gemeinschaften wie 0 (*Wahlen* – 135 Entscheidungen), 3 (*Datenschutz* – 38 Entscheidungen), 9 (*Abgeordnete* – 57 Entscheidungen), 11 (*Religion* – 45 Entscheidungen), 24 (*Rundfunk* – 19 Entscheidungen) oder 77 (*Kunst* – 9 Entscheidungen) lassen sich rasch betiteln, ebenso Gemeinschaften, welche die Bewältigung historischer Ausnahmesituationen betreffen, etwa die Gemeinschaften 41 (*Nachkriegszeit [Kriegsfolgen]* – 22 Entscheidungen), 93 (*Berlin* – 10 Entscheidungen) und 110 (*Rentenüberleitung DDR* – 5 Entscheidungen). Die Benennung fällt allerdings nicht immer so leicht. Ein Grund dafür ist, dass zum Teil viele verschiedene Materien in einer Gemeinschaft zusammenkommen, die in der Rechtsprechung des BVerfG ebenso eng miteinander verknüpft sind. Das zeigt sich beispielsweise an Gemeinschaft 7 (Tab. 4.12, S. 293).

<i>BVerfGE</i>	<i>Entscheidungsname</i>	<i>Innen</i>	<i>Außen</i>	$\Sigma$
7, 198	Lüth	51	0	51
66, 116	Springer/Wallraff	12	15	27
20, 162	Spiegel	22	4	26
93, 266	„Soldaten sind Mörder“	8	15	23
61, 1	Wahlkampf/CSU : NPD Europas“	15	5	20
42, 143	Deutschland-Magazin	12	7	19
10, 118	Berufsverbot I	17	1	18
34, 269	Soraya	15	3	18
60, 234	Kredithaie	12	6	18
99, 185	Scientology	3	15	18

Tabelle 4.12: Repräsentative Mitglieder der Gemeinschaft 7 (*Meinung/Presse/APR* – 66 Entscheidungen)

Betrachtet man hier nur die zehn innerhalb der Gemeinschaft am besten vernetzten Entscheidungen, so scheint der Titel *Meinung/Presse/Allgemeines*

*Persönlichkeitsrecht (APR)* passend, und unter den weiteren Gemeinschaftsmitgliedern finden sich ebenso Entscheidungen wie BVerfGE 82, 272 – Postmortale Schmähdikritik, BVerfGE 90, 241 – Ausschwitzlüge und Caroline I–III (BVerfGE 97, 125; 101, 361; 120, 180). Allerdings gehören auch BVerfGE 69, 315 – Brokdorf, BVerfGE 124, 300 – Rudolf Heß Gedenkfeier und BVerfGE 128, 226 – Fraport der betrachteten Gemeinschaft an – Entscheidungen, die viele Juristen in erster Linie mit der Versammlungsfreiheit assoziieren dürften. Die heterogene Gemeinschaftszusammensetzung lässt sich darauf zurückführen, dass Presse- und Meinungsfreiheit sowie Meinungs- und Versammlungsfreiheit in der Rechtsprechung des BVerfG oft zusammen thematisiert und potenzielle Grundrechtsverletzungen an ähnlichen Maßstäben geprüft werden.<sup>136</sup> Insofern spiegelt die Gemeinschaftszusammensetzung hier die Rechtsprechungspraxis des BVerfG wider – mit der Konsequenz, dass sich die Gemeinschaft kaum prägnant benennen lässt.<sup>137</sup>

Weitere Probleme bei der Namensgebung treten auf, wenn viele Entscheidungen in einer Gemeinschaft unbenannt sind, oder wenn mehrere Gemeinschaften Entscheidungen enthalten, deren Namen auf ähnliche Inhalte deuten. Betroffen sind insbesondere Gemeinschaften mit Entscheidungen zu Steuern und staatlichen Leistungen, zu prozessualen Fragen und zu rechtsstaatlichen Grundsätzen. Anhand eines kursorischen Entscheidungsstudiums können hier oft nur sehr allgemeine Namen vergeben werden, beispielsweise *Einstweilige Anordnung (EA)* für die Gemeinschaften 16, 31, 67 und 114. Ebenso scheinen die Entscheidungen zu Art. 3 I GG über die Gemeinschaften 4, 26, 85, 101, 123, 136 und 171 verstreut zu sein, die möglicherweise für unterschiedliche Problemkontexte des allgemeinen Gleichheitssatzes stehen. Im Rahmen von Detailstudien zu einzelnen Gemeinschaften, die ähnliche Namen tragen, könnte man versuchen, diese Gemeinschaften schärfer gegeneinander abzugrenzen.

<sup>136</sup> Zur Verknüpfung von Presse- und Meinungsfreiheit etwa BVerfGE 102, 347 – Schockwerbung I, Rn. 40: „Der – hier in den Schutz der Pressefreiheit eingebettete – Schutz des Art. 5 Abs. 1 Satz 1 GG erstreckt sich auch auf kommerzielle Meinungsäußerungen sowie reine Wirtschaftswerbung, die einen wertenden, meinungsbildenden Inhalt hat (vgl. BVerfGE 71, 162 <175>).“ Ein Beispiel für die Verknüpfung von Meinungs- und Versammlungsfreiheit ist BVerfGE 124, 300 – Rudolf Heß Gedenkfeier, Rn. 45: „Zulässig ist die Verfassungsbeschwerde zunächst hinsichtlich der auf Art. 8 Abs. 1 in Verbindung mit Art. 5 Abs. 1 und 2 GG gestützten Rüge. Als Bestätigung eines Versammlungsverbots kann die angegriffene Entscheidung den Beschwerdeführer in seinem Recht aus Art. 8 Abs. 1 GG verletzen. Dies kann insbesondere dann der Fall sein, wenn die Versammlung in verfassungsrechtlich nicht tragfähiger Weise wegen ihres Inhalts verboten wird. Denn der Inhalt einer Meinungsäußerung, der im Rahmen des Art. 5 GG nicht unterbunden werden darf, kann nicht zur Rechtfertigung von Maßnahmen herangezogen werden, die das Grundrecht des Art. 8 GG beschränken. Dabei richtet sich die Reichweite der Versammlungsfreiheit insoweit nach dem Umfang des von Art. 5 Abs. 1 und 2 GG gewährten Schutzes (vgl. BVerfGE 90, 241 <246>; 111, 147 <154 f.>).“

<sup>137</sup> Zu der möglicherweise atypischen Position von BVerfGE 128, 226 – Fraport in einer Gemeinschaft mit Entscheidungen zu Meinungsfreiheit, Pressefreiheit und allgemeinem Persönlichkeitsrecht außerdem unten S. 298–300.

Will man sich zunächst einen Überblick über das Resultat der Benennungsbemühungen verschaffen, so bietet es sich an, die Zitierbeziehungen, die zwischen den einzelnen Gemeinschaften verlaufen, in Distanzen im zweidimensionalen Raum zu übersetzen. Dazu kann man einen *Gemeinschaftsgraphen* (*community graph*) erstellen, der für jede Gemeinschaft einen Punkt enthält, der alle der Gemeinschaft zugeordneten Entscheidungen repräsentiert.<sup>138</sup> In diesem Graphen sind zwei Gemeinschaften genau dann durch eine Linie miteinander verbunden, wenn eine Entscheidung aus der einen Gemeinschaft eine Entscheidung aus der anderen Gemeinschaft zitiert;<sup>139</sup> das Gewicht der Linie entspricht der Anzahl der Zitate (gewichtete Zählung), die zwischen diesen Gemeinschaften verlaufen.<sup>140</sup> Die kräftebasierte Positionierung der Elemente des Gemeinschaftsgraphen ermöglicht dann erste Schritte in Richtung einer *Kartographie des Verfassungsrechts*.

Abbildung 4.27 (S. 297) zeigt die 189 Gemeinschaften des Gemeinschaftsgraphen, die zwei oder mehr Mitglieder haben. Alle Gemeinschaften mit mehr als zehn Mitgliedern sind durch ihre Namen gekennzeichnet. Zwischen den Gemeinschaften verlaufen 4.709 Linien mit einem Gesamtgewicht von 26.395, d.h. rund 46 % des Gesamtgewichts aller Pfeile im BVerfGE-MultiDiGraphen entfällt auf Beziehungen *zwischen* den identifizierten Gemeinschaften. Diese Zahl ist nicht überraschend, da in vielen BVerfGE-Entscheidungen Fragen zu verschiedenen inhaltlichen Schwerpunkten auftreten und auch Entscheidungen mit materiell-rechtlichem Fokus ihre prozessualen Ausführungen durch Zitate belegen. Gerade deshalb ist jedoch erstaunlich, dass es dem *Infomap*-Algorithmus in den meisten Fällen gelungen ist, juristisch sinnvoll interpretierbare Gemeinschaften zu bilden.

Auch die in Abbildung 4.27 (S. 297) dargestellten Beziehungen und Distanzen zwischen den einzelnen Gemeinschaften scheinen größtenteils plausibel: Der Rechtsbereich *Wahlen* ist eng mit dem Bereich *Abgeordnete* verbunden, der deutliche Beziehungen zu *Verteidigung*, *Untersuchungsausschüsse* und *Europa* aufweist – Bereiche, in denen es regelmäßig auch um die Sicherung parlamentarischer Kontrolle geht. Zwischen *Beruf* und *Freie Berufe* bestehen zahlreiche

<sup>138</sup> In der Literatur scheint der hier als Gemeinschaftsgraph bezeichnete Graph keinen besonderen Namen zu tragen; *Krishnamurthy und Wang*, Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement 2001, S. 19–23, bezeichnen einen ähnlich konstruierten Graphen als *Clustergraphen* (*cluster graph*), was allerdings unglücklich ist, da der Terminus in der Graphentheorie ein anderes Konzept beschreibt; siehe etwa *Shamir, Sharan und Tsur*, Discrete Applied Mathematics 144 (2004), S. 173 (174).

<sup>139</sup> *Pollner, Palla und Vicsek*, Europhysics Letters 73 (2006), S. 478 (478–479) verwenden ebenfalls den Namen *community graph*, meinen damit aber einen Graphen, in dem die Punkte (Gemeinschaften) durch gemeinsame Mitglieder als Linien miteinander verbunden sind.

<sup>140</sup> An einem konkreten Beispiel: Besteht im ursprünglichen Graphen zwischen Punkt a und Punkt b eine Linie mit einem Gewicht von 3, d.h. eine Entscheidung A zitiert eine Entscheidung B dreimal oder umgekehrt, und werden die Entscheidungen A und B vom *Infomap*-Algorithmus den Gemeinschaften 1 respektive 2 zugeordnet, so besteht zwischen den Punkten, die diese Gemeinschaften im Gemeinschaftsgraphen repräsentieren, ebenfalls eine Linie mit dem Gewicht von 3.

Zitierbeziehungen, ebenso zwischen *Steuern* und *Abgaben*; *Rechtsstaat*, *Strafe* und *Bestimmtheit* stehen einander nahe. Die Entfernung von *Kommunale Abgaben* zu *Steuern*, *Abgaben* und *Kommunen* mag verwundern, ebenso die breite Streuung der Gemeinschaften mit prozessuaalem oder rechtsstaatlichem Schwerpunkt, beispielsweise der mit *EA* (d.h. *Einstweilige Anordnung*) oder mit *Gehör* (d.h. *Rechtliches Gehör*) betitelten Gemeinschaften. Die Lage dieser Gemeinschaften könnte auf Eigenheiten des Layout-Algorithmus zurückzuführen sein, nach dem Gemeinschaften ihre Position nicht nur durch Anziehung (von den Gemeinschaften, mit denen sie verbunden sind), sondern auch durch Abstoßung (von den Gemeinschaften, mit denen sie nicht verbunden sind) erhalten.<sup>141</sup> Sie könnte allerdings auch auf inhaltliche Nuancen hindeuten, die erst bei einer Detailbetrachtung einzelner Entscheidungen sichtbar werden.

Da der *Infomap*-Algorithmus nichtdeterministisch ist, stellt das geschilderte Resultat nur eines der möglichen Ergebnisse dar. Damit ist einerseits unklar, ob der Algorithmus hier nur zufällig ein interpretierbares Resultat geliefert hat. Andererseits ist wahrscheinlich, dass die Zusammenfassung der Ergebnisse mehrerer Läufe noch besser interpretierbare Resultate liefert. Vor diesem Hintergrund wurden neben dem oben geschilderten Lauf noch 99 weitere, also insgesamt 100 Läufe unternommen.<sup>142</sup> Für jede BVerfGE-Entscheidung wurde in jedem Lauf die Ziffer der Gemeinschaft festgehalten, der sie in diesem Lauf zugeordnet wurde. Darauf aufbauend lässt sich eine quadratische Matrix erstellen, die für jede Entscheidung eine Zeile und eine Spalte hat und in deren Feldern der Anteil aller Läufe notiert ist, in denen die Zeilen-Entscheidung und die Spalten-Entscheidung *unterschiedlichen* Gemeinschaften zugeordnet wurden (*Dissensmatrix*).<sup>143</sup> Abbildung 4.28 (a) (S. 298) zeigt, wie häufig die (in zehn Klassen eingeteilten) Anteile vorkommen, wenn man alle 5.354.628 Paare verschiedener Entscheidungen betrachtet.<sup>144</sup> Abbildung 4.28 (b) (S. 298) stellt dar, wie häufig die einzelnen Anteile vorkommen, wenn man nur die 76.955 unsicheren Entscheidungspaare betrachtet: Paare von Entscheidungen, die weder immer noch nie in dieselbe Gemeinschaft eingeordnet wurden. An den Bildern wird deutlich, dass die Einteilung der BVerfGE-Entscheidungen durch den *Infomap*-Algorithmus relativ stabil ist.

<sup>141</sup> Dazu bereits S. 112–115.

<sup>142</sup> Die Ergebnisse der 100 Läufe können im Online-Appendix eingesehen werden; der im Vorigen betrachtete Lauf ist dort *run\_0*.

<sup>143</sup> Ermittelt wurde die Hamming-Distanz (*Hamming distance*); siehe *Russell und Norvig*, *Artificial Intelligence*, 2011, S. 752; *Murphy*, *Machine Learning: A Probabilistic Perspective*, 2012, S. 878.

<sup>144</sup> Die Anzahl der Paare verschiedener Entscheidungen ergibt sich wie folgt:  $3.273^2 - \sum_{i=1}^{3.273} i = 5.354.628$ . Das ist die Anzahl der Felder der Dissensmatrix, die oberhalb bzw. unterhalb der Diagonalen liegen, denn die Dissensmatrix ist symmetrisch („in dieselbe Gemeinschaft eingeordnet werden“ ist eine symmetrische Beziehung) und in der Diagonalen stehen nur Nullen (eine Entscheidung kann im Verhältnis zu sich selbst nie „in eine andere Gemeinschaft eingeordnet werden“).

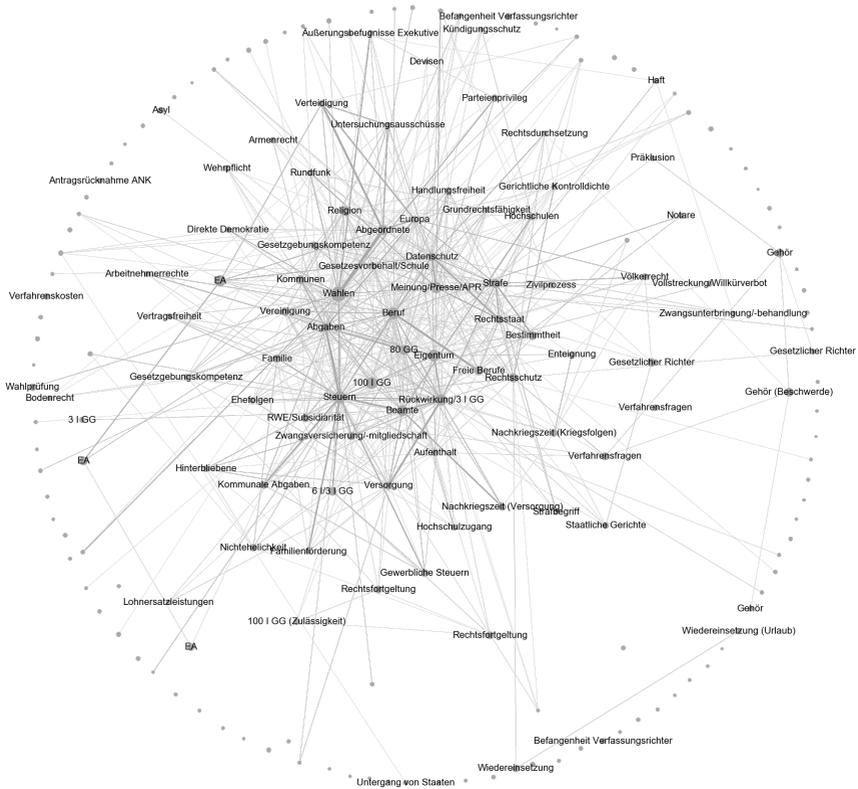
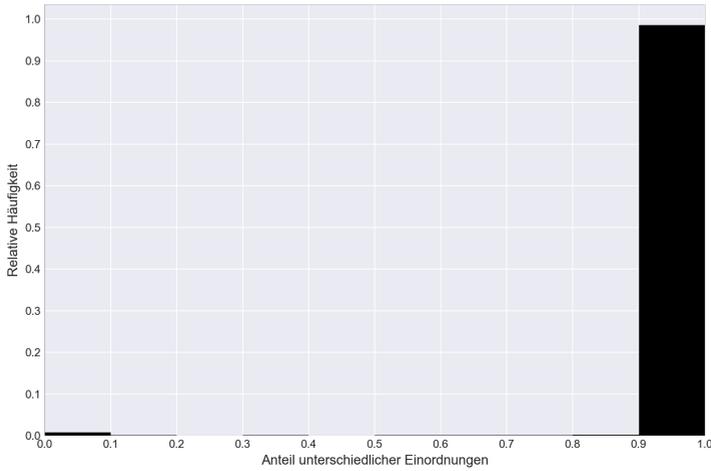
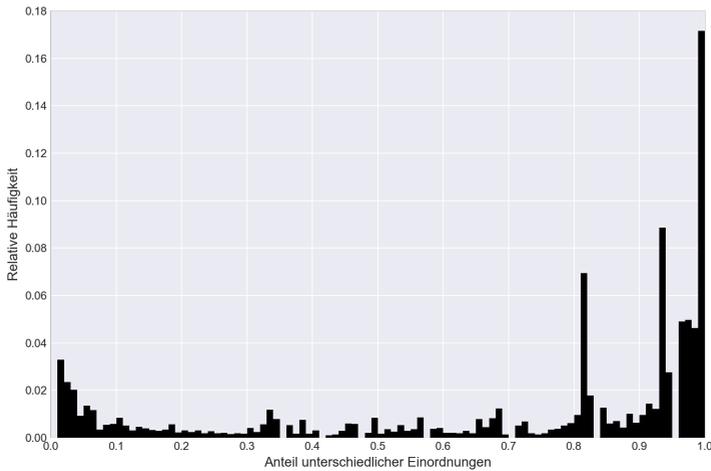


Abbildung 4.27: Gemeinschaftsgraph der Gemeinschaften in BVerfGE 1–140<sup>145</sup>

<sup>145</sup> Die Größe der Punkte ist proportional zur Anzahl der Entscheidungen, die in der jeweiligen Gemeinschaft zusammengefasst sind. Die Stärke der Linien zeigt die Summe der Gewichte aller Linien, die zwischen Entscheidungen verschiedener Gemeinschaften verlaufen. Das Layout wurde mit dem kräftebasierten Fruchterman-Reingold-Algorithmus und *seed 1234* erzeugt, wobei die Abstoßung zwischen den Punkten so weit erhöht wurde, dass alle Gemeinschaftstitel lesbar sind. Wo Normen als (Teile von) Gemeinschaftsnamen dienen, ist auf das Vorstellen von „Art.“ aus Platzgründen verzichtet worden. Bei den als solche erkennbaren Freiheitsrechten (*Beruf*, *Meinung*, *Presse Religion*, *Rundfunk*, *Vereinigung*) ist das Nachstellen von „Freiheit“ unterblieben. Die darüber hinaus genutzten Abkürzungen sind (neben GG für Grundgesetz) ANK (Abstrakte Normenkontrolle), APR (Allgemeines Persönlichkeitsrecht), EA (Einstweilige Anordnung) und RWE (Rechtswegerschöpfung); *Gehör* meint *Rechtliches Gehör*.



(a) Alle Entscheidungspaare (5.354.628)



(b) Unsichere Entscheidungspaare (76.955)

Abbildung 4.28: Relative Häufigkeitsverteilung des Anteils der Zuordnung zu unterschiedlichen Gemeinschaften (100 Läufe)<sup>146</sup>

<sup>146</sup> Abbildung 4.28 (a) (S. 298) zeigt, welcher Anteil der 5.354.628 Paare verschiedener Entscheidungen (y-Achse) wie häufig (x-Achse) in unterschiedliche Gemeinschaften eingeordnet wurde; dabei wurden 10 Klassen (*bins*) gebildet. In Abbildung 4.28 (b) (S. 298) ist dargestellt, welcher Anteil der Zuordnungen (y-Achse) auf die einzelnen Perzentile (x-Achse) entfällt, wenn man nur die 76.955 unsicheren Zuordnungen betrachtet, deren Wert weder 0 noch 1 ist; dabei wurden 100 Klassen gebildet.

Auf Basis der Dissensmatrix kann man auch *Konsensgemeinschaften* (*consensual communities*) zu einem bestimmten Anteil  $p$  als Schwellenwert definieren: Gemeinschaften, deren Mitglieder in einem Anteil von  $p$  aller unternommenen Läufe *alle* in *dieselbe* Gemeinschaft eingeordnet werden (*p-Konsensgemeinschaften*).<sup>147</sup> Diese Gemeinschaften lassen sich ermitteln, indem man aus der Dissensmatrix einen gewichteten Graphen ableitet, in dem jeder Punkt (jede Entscheidung) mit jedem anderen Punkt verbunden ist und das Gewicht jeder Linie dem Anteil aller Läufe entspricht, in denen die betroffenen Entscheidungen in *dieselbe* Gemeinschaft eingeordnet werden (*Konsensgraph*).<sup>148</sup> Die *p-Konsensgemeinschaften* sind die maximalen Cliques<sup>149</sup> des Graphen, den man erhält, wenn man aus dem Konsensgraphen alle Linien entfernt, deren Gewicht kleiner als  $p$  ist (*p-Konsensgraph*).<sup>150</sup> Werden *p-Konsensgemeinschaften* für unterschiedliche Werte von  $p$  ermittelt, so können auf dieser Basis auch hierarchische und überlappende Gemeinschaftsstrukturen untersucht werden; beispielsweise enthält eine 0,9-Konsensgemeinschaft alle 0,95-Konsensgemeinschaften, deren Mitglieder eine Teilmenge der 0,9-Konsensgemeinschaft sind.<sup>151</sup>

Der Mehrwert dieser Vorgehensweise im Vergleich zur Betrachtung eines einzelnen Ergebnisses sei an einigen Beispielen verdeutlicht: Im oben betrachteten Resultat wird BVerfGE 128, 226 – Fraport mit den Entscheidungen zur Meinungs- und Pressefreiheit gruppiert. Untersucht man 100 Iterationen gemeinsam, so stellt sich allerdings heraus, dass die einzige nichttriviale 0,9-Konsensgemeinschaft, in der BVerfGE 128, 226 – Fraport enthalten ist, neben dieser Entscheidung nur

<sup>147</sup> Die Bezeichnung als *consensual communities* findet sich auch bei *Campgiotto, Guillaume und Saifi*, IEEE/ACM ASONAM 2013, S. 272 (272); *Seifi u.a.*, Stable Community Cores in Complex Networks, in: *Menezes, Evsukoff und González* (Hrsg.), Complex Networks, 2013, S. 87 (88–90) sprechen von *consensus communities* bzw. Gemeinschaftskernen (*community cores*). Die hier zugrunde gelegte Definition ist allerdings enger als die der Autoren (hierzu sogleich). Ein verwandtes Konzept, das die Konsensidee noch weiter treibt als hier, ist das Konsensclustering (*consensual clustering*), dazu *Lancichinetti und Fortunato*, Scientific Reports 2 (2012), S. 336-1–336-7.

<sup>148</sup> Das heißt konkret: 1 minus den Wert des entsprechenden Feldes in der Dissensmatrix.

<sup>149</sup> Zur Erinnerung: Cliques sind vollständig verbundene Subgraphen; eine Clique ist maximal, wenn sie nicht Teil einer größeren Clique ist.

<sup>150</sup> Bei *Campgiotto, Guillaume und Saifi*, IEEE/ACM ASONAM 2013, S. 272 (272–273) und *Seifi u.a.*, Stable Community Cores in Complex Networks, in: *Menezes, Evsukoff und González* (Hrsg.), Complex Networks, 2013, S. 87 (89–90) sind die *p-Konsensgemeinschaften* die *Komponenten* des *p-Konsensgraphen*. Den Autoren genügen also maximale verbundene Subgraphen, während die hiesige Definition maximale *vollständig* verbundene Subgraphen verlangt.

<sup>151</sup> Hierarchische Gemeinschaftsstrukturen ergeben sich auch, wenn man die Definition von *Campgiotto, Guillaume und Saifi*, IEEE/ACM ASONAM 2013, S. 272 (272–273) und *Seifi u.a.*, Stable Community Cores in Complex Networks, in: *Menezes, Evsukoff und González* (Hrsg.), Complex Networks, 2013, S. 87 (89–90) zugrunde legt; überlappende Gemeinschaftsstrukturen sind danach allerdings nicht möglich; siehe *Seifi u.a.*, Stable Community Cores in Complex Networks, in: *Menezes, Evsukoff und González* (Hrsg.), Complex Networks, 2013, S. 87 (90).

BVerfGE 69, 315 – Brokdorf enthält; allein BVerfGE 69, 315 – Brokdorf wird also in mehr als 90 von 100 Läufen mit BVerfGE 128, 226 – Fraport zusammen gruppiert. Von den prominenten Entscheidungen aus Tabelle 4.12 (S. 293) sind lediglich BVerfGE 7, 198 – Lüth, BVerfGE 34, 269 – Soraya, BVerfGE 42, 143 – Deutschland-Magazin, BVerfGE 60, 234 – Kredithaie, BVerfGE 61, 1 – Wahlkampf/CSU : NPD Europas‘, BVerfGE 66, 116 – Springer/Wallraff, BVerfGE 93, 266 – ‚Soldaten sind Mörder‘ und BVerfGE 99, 185 – Scientology in derselben 0,9-Konsensgemeinschaft enthalten;<sup>152</sup> BVerfGE 20, 162 – Spiegel wird in mehr als 90 % der Fälle mit BVerfGE 21, 271 – Südkurier, BVerfGE 117, 244 – CICERO und einigen anderen Entscheidungen zur Pressefreiheit zusammengefasst,<sup>153</sup> für BVerfGE 10, 118 – Berufsverbot I schafft es allein BVerfGE 25, 88 – Berufsverbot II über die 90 %-Hürde. Betrachtet man die mit BVerfGE 7, 198 – Lüth regelmäßig zusammengefassten Entscheidungen, so zeigt sich ein Kern von Entscheidungen, in denen die Meinungsfreiheit eine so herausgehobene Rolle spielt, dass diese Entscheidungen stets in dieselben Gemeinschaften eingeordnet werden, während der Bezug zu anderen Entscheidungen mit Schwerpunkten bei der Presse- oder Versammlungsfreiheit offenbar weniger eng ist.

Alle in Tabelle 4.11 (S. 293) unter dem Titel *Europa* zusammengefassten Entscheidungen werden in mehr als 90 % der Fälle gemeinsam gruppiert, ebenso 69 der 76 Entscheidungen aus der Gemeinschaft *Eigentum* (Tab. 4.10, S. 292) und 75 der 77 Entscheidungen aus der Gemeinschaft *Beruf* (Tab. 4.9, S. 292). Auch die Gemeinschaften zur *Religion*, zur *Kunst* und zur *Präklusion* sind nach dem 90 %-Kriterium stabil; in der Gemeinschaft zum *Datenschutz* sind 37 von 38 Entscheidungen in über 90 % der Fälle gemeinsam enthalten. Diese Befunde bestätigen nicht nur Teile des oben näher betrachteten Resultats, sondern sie machen auch stärkere Aussagen über die Gemeinschaftsstruktur der BVerfGE-Entscheidungen, da sie auf einer breiteren Informationsbasis beruhen. Insgesamt scheint der *Infomap*-Algorithmus für die Zuordnung von BVerfGE-Entscheidungen zu Rechtsbereichen also zuverlässig interpretierbare Resultate zu liefern. Da der Algorithmus ausschließlich auf den vom BVerfG geschaffenen Zitierstrukturen in BVerfGE arbeitet, kann man also festhalten: Zur Gruppierung der BVerfGE-Entscheidungen in Rechtsbereiche kann man sich auf die Zitate des BVerfG verlassen.

---

<sup>152</sup> Die 0,9-Konsensgemeinschaft, in der die Entscheidungen zur Meinungsfreiheit zusammengefasst sind, enthält 43 Entscheidungen.

<sup>153</sup> Die 0,9-Konsensgemeinschaft, in der die Entscheidungen zur Pressefreiheit zusammengefasst sind, enthält neun Entscheidungen.

## 2. Begriffskontexte erkunden: Textbasierte Subgraphen

Die zur Ordnung von Rechtsprechung herangezogenen Konzepte gehen von einer *strukturellen* Zusammengehörigkeitsdefinition aus, die auf Zitierbeziehungen basiert. Will man die Rechtsprechung in einem bestimmten Bereich erkunden, so bietet es sich an, zusätzlich eine *semantische*, d.h. auf Inhalten aufbauende Definition heranzuziehen. Semantische Definitionen von Zusammengehörigkeit können mithilfe regulärer Ausdrücke operationalisiert werden, die ausgewählte Stichwörter repräsentieren.<sup>154</sup> Jede Entscheidung im BVerfGE-Korpus kann auf diese regulären Ausdrücke durchsucht werden. Anschließend lassen sich jedem Punkt im BVerfGE-MultiDiGraphen neue Attribute zuordnen, deren Werte den Trefferzahlen zu den einzelnen Stichwörtern entsprechen.<sup>155</sup> Aus dem so angereicherten Zitiergraphen kann man (im Wege der Filterung bzw. Selektion) einen Subgraphen erstellen, der nur diejenigen Entscheidungen als Punkte (und die Zitate zwischen ihnen als Pfeile) enthält, in denen mindestens eine bestimmte Anzahl von Treffern gefunden wird (*textbasierter Subgraph*).<sup>156</sup> Dieser Subgraph ist ebenso der quantitativen und visuellen Analyse zugänglich wie der ursprüngliche Zitiergraph; er ermöglicht aber eine Konzentration auf diejenigen Entscheidungen, die in einem bestimmten Sachkontext potenziell relevant sind.

Im Folgenden werden einige textbasierte Subgraphen visuell erkundet. Die untersuchten Stichwörter sind *Europa*, *Datenschutz* und *Religion* sowie die Trias *Christentum*, *Judentum* und *Islam*. Für jedes Stichwort sind zwei Perspektiven dargestellt: eine Perspektive, die *temporale* Zusammenhänge betont,<sup>157</sup> und eine Perspektive, die *thematische* Zusammenhänge akzentuiert.<sup>158</sup>

Beginnt man mit *Europa*, so lässt sich in der temporalen Perspektive deutlich erkennen, dass die Häufigkeit, mit der sich das BVerfG nicht nur beiläufig zu Europa äußert, in den letzten zehn Jahren stark zugenommen hat. Für den juristischen Beobachter mag diese Einsicht trivial sein – Abbildung 4.29 (S. 302) ermöglicht ihm aber zumindest, auf Basis quantitativer Analysen visuell zu untermauern, was

<sup>154</sup> Zu regulären Ausdrücken bereits oben S. 240–248.

<sup>155</sup> Beispiel: Man durchsucht jede Entscheidung auf das Stichwort *Menschenwürde* und hält die jeweilige Trefferzahl in einem Attribut *menschenwuerde* fest. Führt man die Suche für mehrere reguläre Ausdrücke gleichzeitig durch, die jeweils unterschiedliche Stichwörter repräsentieren, so kann man sie sich auch als Erweiterung des Zitiergraphen zu einem mehrschichtigen Graphen vorstellen, in dem nicht nur Entscheidungen, sondern auch Stichwörter als Punkte enthalten sind, und in dem nicht nur Zitate, sondern auch Stichwortvorkommen als Verbindungen zwischen den Punkten auftauchen.

<sup>156</sup> Reguläre Ausdrücke erfassen nie 100 % dessen, was man durch sie zu erfassen hofft. Verantwortlich dafür sind insbesondere Aufzeichnungs- und Aufbereitungsfehler (dazu oben S. 86–88). Die damit verbundenen Unsicherheiten sind bei der Interpretation aller Resultate im Hinterkopf zu behalten.

<sup>157</sup> Die temporale Perspektive ähnelt Abbildung 4.8 (S. 254). Zur Problematik der visuellen Darstellung temporaler Informationen bereits oben S. 110–112; außerdem *Coupette und Fleckner*, JZ 73 (2018), S. 379 (387–388).

<sup>158</sup> Die thematische Perspektive ähnelt den Abbildungen 4.21–4.25 (S. 284–288).

seine Intuition nahelegt. Auch dass vor allem der Zweite Senat die Rechtsprechung des Gerichts zu Europa prägt, wird für den Kenner des Geschäftsverteilungsplans keine Überraschung sein – andere können es aus Abbildung 4.29 (S. 302) ablesen.

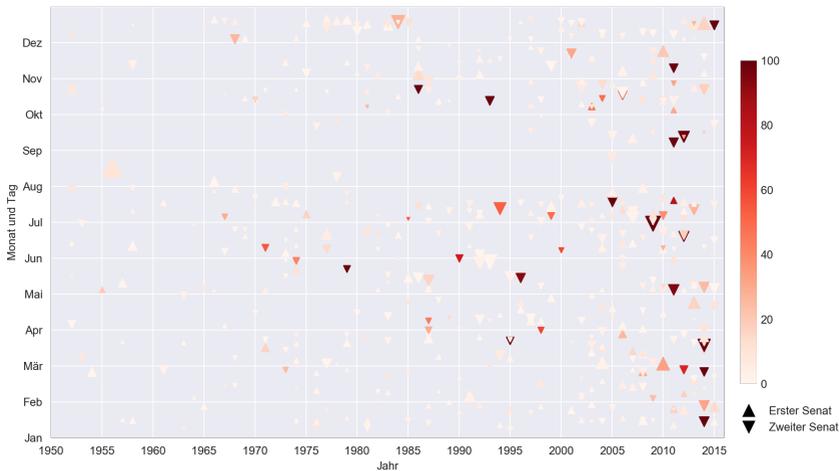


Abbildung 4.29: BVerfGE-Entscheidungen zum Stichwort *Europa*<sup>159</sup>

Wer wissen möchte, welche Entscheidungen sich intensiv mit Europa befassen und wie diese Entscheidungen zueinander stehen, kann auf die thematische Perspektive zurückgreifen (Abb. 4.30, S. 303). Diese wird für große Mengen von Entscheidungen schnell unübersichtlich, sodass man die Darstellung auf Entscheidungen beschränken sollte, die mindestens eine bestimmte Trefferzahl aufweisen. Für das Stichwort *Europa* wurde mit 35 eine Mindestzahl gewählt, die dafür sorgt, dass Abbildung 4.30 (S. 303) gerade noch lesbar ist.<sup>160</sup> Lässt man die Entscheidungen ohne Namen außen vor, so könnte man aus den abgebildeten Entscheidungen

<sup>159</sup> Gesucht wurde nach [Ee]urop[aa], wobei Vorkommen (richtiger oder falscher) Trennstri- che und nachfolgender Leerzeichen (bzw. *whitespace characters*) berücksichtigt wurden. Die Abbildung zeigt alle Entscheidungen mit mindestens einem Treffer. Jede Entscheidung ist durch ein Dreieck repräsentiert, an dessen Ausrichtung sich der entscheidende Senat ablesen lässt. Die Größe des Dreiecks ist proportional zur Länge der Entscheidung, die Farbe signalisiert die Anzahl der Treffer (je dunkler das Rot, desto häufiger wurde der Suchausdruck in der entsprechenden Entscheidung gefunden). Die Farbkarte ist hier im Intervall [0, 100] skaliert, was aber nichts über die maximale Trefferzahl aussagt (Spitzenreiter ist BVerfGE 123, 267 – Lissabon mit 817 Treffern), sondern nur dazu dient, zwischen Entscheidungen mit „relativ vielen“ und solchen mit „relativ wenigen“ Treffern zu differenzieren. Die geschilderten Grundsätze gelten entsprechend für alle nachfolgenden temporalen Darstellungen, bei denen neben einem anderen Suchausdruck allerdings oft auch eine andere Skalierung für die Farbkarte gewählt wurde.

<sup>160</sup> Das einzige Kriterium, nach dem sich die Wahl des Schwellenwerts hier bestimmt, ist die Übersichtlichkeit des Results. Eine Möglichkeit, sie theoretisch zu fundieren, besteht darin, auf Statistiken zur Verteilung der Trefferzahlen zurückzugreifen und etwa den jeweiligen Median

(insbesondere den Entscheidungen, die durch dunkelrote Dreiecke repräsentiert sind) ohne Weiteres einen Grundkurs zu den verfassungsrechtlichen Problemen der europäischen Integration zusammenstellen (nach BVerfGE 140 ergangene Entscheidungen wären zu ergänzen). Die thematische Perspektive macht außerdem die Rolle von BVerfGE 123, 267 – Lissabon als zentraler Referenzpunkt späterer Entscheidungen sichtbar; BVerfGE 89, 155 – Maastricht ist demgegenüber in der Europa-Rechtsprechung des BVerfG deutlich weniger prominent.

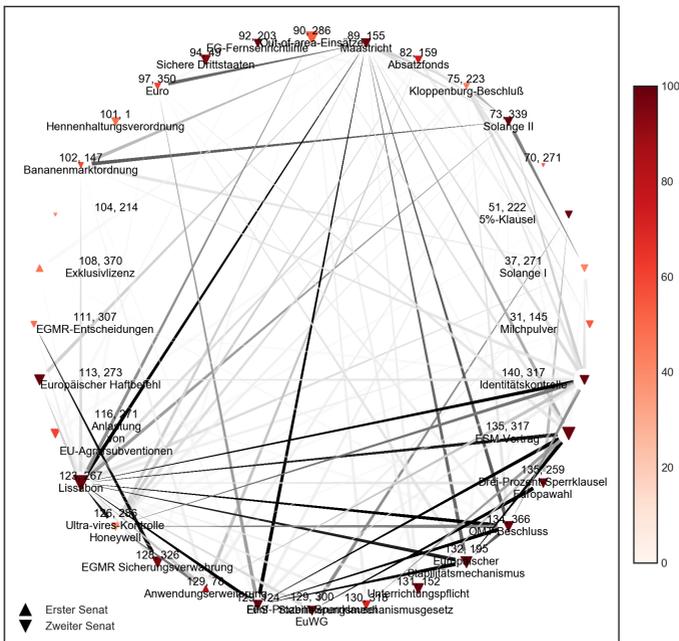


Abbildung 4.30: BVerfGE-Entscheidungen zum Stichwort *Europa* (35 oder mehr Treffer)<sup>161</sup>

als Schwellenwert zu wählen. Damit lässt sich allerdings nicht ausschließen, dass es Situationen gibt, in denen man durch Anwendung der selbstformulierten Regeln keine lesbare Darstellung erhält. Ähnliches gilt bei der Wahl der Intervallgrenzen für die Skalierung der Farbkarte, die hier ebenfalls nicht regelgeleitet erfolgt (oben Fn. 159, S. 302).

<sup>161</sup> Die Entscheidungen ohne Namen sind BVerfGE 70, 271 – {Wahl zum Europäischen Parlament 1984}; und BVerfGE 104, 214 – [NPD-Verbotsverfahren (Aussetzungsantrag)]. Zum Ausdruck, nach dem gesucht wurde, und zur Repräsentation der einzelnen Entscheidungen als Pfeile bereits Fn. 159 (S. 302). Die Anordnung der Dreiecke und die Darstellung der Zitate als Pfeile folgt den Konventionen aus Abbildung 4.21–4.25 (S. 284–288): Die Entscheidungen sind entgegen dem Uhrzeigersinn angeordnet, beginnend bei der ältesten Entscheidung ganz rechts („im Westen“). Zitate verlaufen vom dickeren zum dünneren Ende der Linie. Die Zitierhäufigkeit

Am textbasierten Subgraphen zum Stichwort *Datenschutz* zeigt sich die Evolution des Diskurses zu einem Rechtsproblem – vielleicht sogar die Entwicklung des Rechtsproblems selbst. Das Stichwort *Datenschutz* taucht erstmals in zwei Entscheidungen des Zweiten Senats vom Dezember 1982 auf. Es folgen zwei kürzere Entscheidungen des Ersten Senats von April und Juni 1983 und schließlich die erste Schwerpunktentscheidung des Ersten Senats im Dezember 1983 (BVerfGE 65, 1 – Volkszählung). Anschließend vergehen fast 15 Jahre bis zur nächsten intensiven Diskussion im Juli 1998 (BVerfGE 100, 313 – Telekommunikationsüberwachung). Seit März 2004 (BVerfGE 109, 279 – Großer Lauschangriff) wird in den Entscheidungen des BVerfG immer häufiger über Datenschutz gesprochen, wobei die Entscheidungen mit den meisten Treffern fast alle vom Ersten Senat stammen.

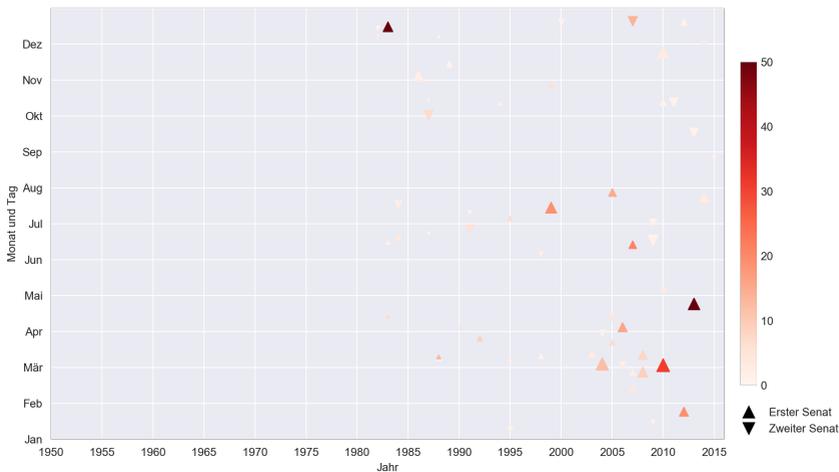


Abbildung 4.31: BVerfGE-Entscheidungen zum Stichwort *Datenschutz*<sup>162</sup>

Die Dominanz des Ersten Senats zeigt sich auch in der thematischen Perspektive (Abb. 4.32, S. 305). Die Entscheidung BVerfGE 119, 331 – [Hartz IV-Arbeitsgemeinschaft] des Zweiten Senats ist im Zitiernetzwerk fast vollständig isoliert. BVerfGE 124, 43 – Beschlagnahme von E-Mails, ebenfalls eine Entscheidung des Zweiten Senats, bezieht sich zwar auf Rechtsprechung des Ersten Senats, wird aber in dessen nachfolgender Rechtsprechung offenbar nicht (bzw. jedenfalls nicht in Form von BVerfGE-Fundstellen) rezipiert. Auffällig ist auch,

wird durch die Linienfarbe signalisiert, die im Intervall  $[-1, 20]$  linear skaliert ist, sodass Pfeile, die für 20 oder mehr Zitate stehen, schwarz und Pfeile, die für ein Zitat stehen, hellgrau erscheinen (die Skalierung beginnt bei  $-1$  statt bei  $0$ , da die einfachen Zitate anderenfalls kaum sichtbar sind).

<sup>162</sup> Gesucht wurde nach [Dd]atensch[üü]tz; zu Einzelheiten der Suche und zur Interpretation der Darstellung bereits oben Fn. 159 (S. 302). Im vorliegenden Fall ist die Farbkarte im Intervall  $[0, 50]$  linear skaliert.

das BVerfGE 118, 168 – Kontostammdaten die Entscheidung BVerfGE 112, 284 – Kontostammdaten nicht zitiert, obwohl beide Entscheidungen zueinander im Verhältnis von Ablehnung der einstweiligen Anordnung und Hauptsacheentscheidung stehen.<sup>163</sup> Dies verwundert besonders, da die Ablehnung der einstweiligen Anordnung mit circa 20 Seiten fast halb so lang ist wie die Hauptsacheentscheidung, die circa 43 Seiten umfasst. Daneben zeigt sich auch in Abbildung 4.32 (S. 305) die bereits in Abbildung 4.24 (S. 287) festgestellte enge Vernetzung zwischen Entscheidungen des Ersten Senats, die Aspekte des Datenschutzes thematisieren.

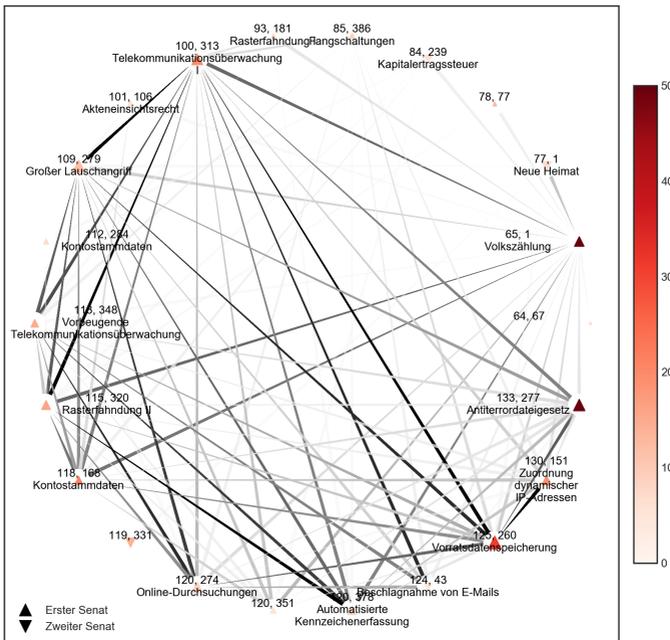


Abbildung 4.32: BVerfGE-Entscheidungen zum Stichwort *Datenschutz* (fünf oder mehr Treffer)<sup>164</sup>

<sup>163</sup> „Kontostammdaten“ ist einer der Entscheidungsnamen, die in der DFR-Sammlung mehrfach vorkommen; dazu bereits oben Fn. 72 (S. 248).

<sup>164</sup> Zum Ausdruck, nach dem gesucht wurde, bereits Fn. 162 (S. 304); zur Darstellungsweise Fn. 161 (S. 303). Die Entscheidungen ohne Namen sind BVerfGE 64, 67 – {Aussetzung Volkszählung}; BVerfGE 78, 77 – {Bekanntmachung der Entmündigung}; BVerfGE 119, 331 – [Hartz IV-Arbeitsgemeinschaft]; und BVerfGE 120, 351 – [Rasterfahndung]. Viele der Treffer in BVerfGE 78, 77 – {Bekanntmachung der Entmündigung} sind darauf zurückzuführen, dass sich

Auf die Frage, wie es das BVerfG mit der Religion hat, kann man mit Blick auf die BVerfGE-Daten antworten: Es kommt darauf an. Zum Stichwort *Religion* im Allgemeinen äußert sich das BVerfG anscheinend seit jeher immer wieder grundsätzlich (Abb. 4.33, S. 306), und zwar sowohl durch den Ersten als auch durch den Zweiten Senat. Dass es dabei oft um das Christentum geht, zeigt die Übereinstimmung vieler „Hotspots“ in Abbildung 4.33 (S. 306) und Abbildung 4.34 (S. 307) (man beachte allerdings die unterschiedliche Skalierung). Zum Stichwort *Judentum* gibt es eine Handvoll Schwerpunktentscheidungen, darunter zwei aus der frühen Nachkriegszeit (Abb. 4.35, S. 307). Der Islam hat das BVerfG in den 1980er Jahren erstmals intensiver beschäftigt, in den 1990er Jahren dagegen kaum, seit der Jahrtausendwende spielt er vor dem BVerfG wieder eine größere Rolle (Abb. 4.36, S. 308).

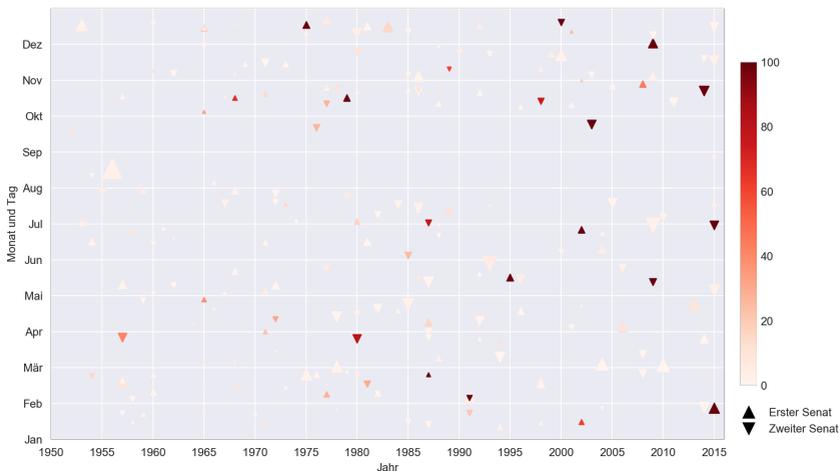
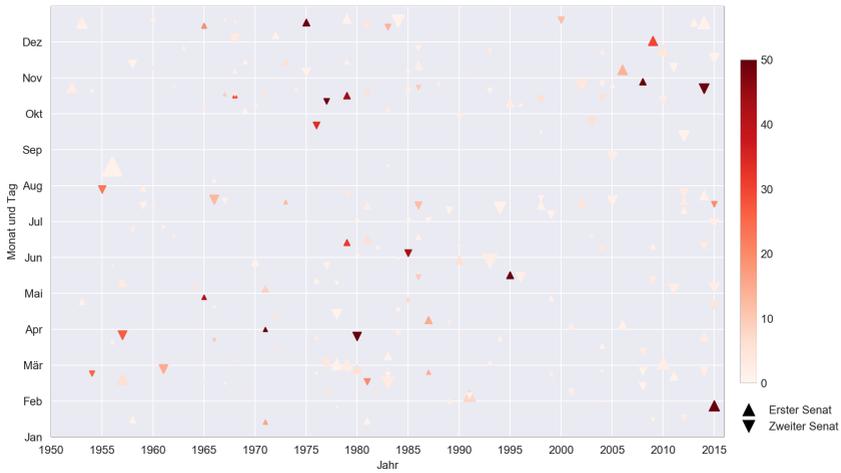
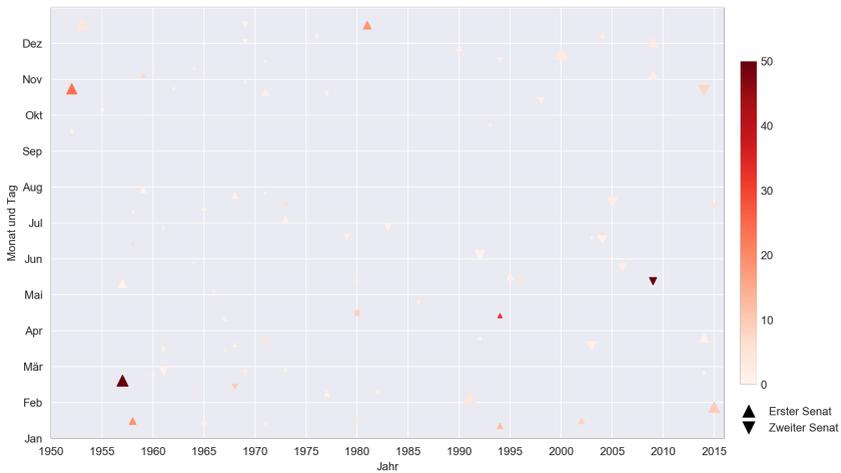


Abbildung 4.33: BVerfGE-Entscheidungen zum Stichwort *Religion*<sup>165</sup>

im Verfahren mehrere Datenschutzbeauftragte äußern, worüber in der Entscheidung berichtet wird.

<sup>165</sup> Gesucht wurde nach [Rr]eligi[oö]; zu Einzelheiten der Suche und zur Interpretation der Darstellung bereits oben Fn. 159 (S. 302). Im vorliegenden Fall ist die Farbkarte im Intervall [0, 100] linear skaliert.

Abbildung 4.34: BVerfGE-Entscheidungen zum Stichwort *Christentum*<sup>166</sup>Abbildung 4.35: BVerfGE-Entscheidungen zum Stichwort *Judentum*<sup>167</sup>

<sup>166</sup> Gesucht wurde nach [Cc]hrist|[Kk]atholi|[Ee]vangel|[Pp]rotestant; zu Einzelheiten der Suche und zur Interpretation der Darstellung bereits oben Fn. 159 (S. 302). Im vorliegenden Fall ist die Farbkarte im Intervall [0, 50] linear skaliert.

<sup>167</sup> Gesucht wurde nach [Jj]ude|[Jj]ud|[Jj]üdi; zu Einzelheiten der Suche und zur Interpretation der Darstellung bereits oben Fn. 159 (S. 302). Im vorliegenden Fall ist die Farbkarte im Intervall [0, 50] linear skaliert.

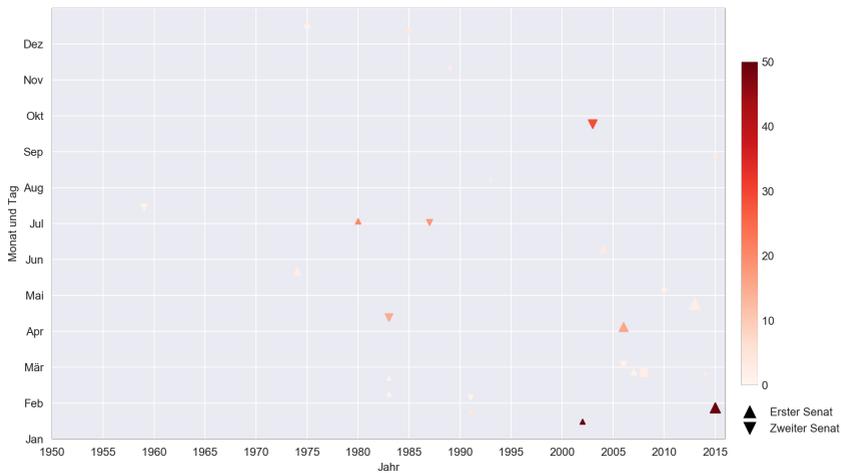


Abbildung 4.36: BVerfGE-Entscheidungen zum Stichwort *Islam*<sup>168</sup>

Die Bilder zu den Stichwörtern *Christentum*, *Judentum* und *Islam* ergänzen sich nicht vollständig zur Darstellung für *Religion*. Das ist darauf zurückzuführen, dass sich das BVerfG immer wieder mit Gruppierungen auseinandersetzen hat, welche den Schutz der Religionsfreiheit beanspruchen, aber nicht zu den Weltreligionen zählen. Wie sich Abbildung 4.37 (S. 309) entnehmen lässt, sind unter den Entscheidungen mit den meisten Treffern zum Stichwort *Religion* unter anderem BVerfGE 83, 341 – Bahá'í, BVerfGE 105, 279 – Osho und zwei Entscheidungen zu den Zeugen Jehovas (BVerfGE 102, 370 – Körperschaftsstatus der Zeugen Jehovas und BVerfGE 139, 321 – Zeugen Jehovas Bremen). Dass einige Entscheidungen eng miteinander verbunden sind, viele Entscheidungen aber andere nur spärlich zitieren, deutet darauf hin, dass die unter dem Dach der Religion thematisierten Probleme verhältnismäßig divers sind.<sup>169</sup>

<sup>168</sup> Gesucht wurde nach [Ii]s1am| [Mm]us1im; zu Einzelheiten der Suche und zur Interpretation der Darstellung bereits oben Fn. 159 (S. 302). Im vorliegenden Fall ist die Farbkarte im Intervall [0, 50] linear skaliert.

<sup>169</sup> Das wird den Kenner der Rechtsprechung wieder wenig überraschen. Ihm bietet Abbildung 4.37 (S. 309) die Möglichkeit, seine Expertise unwissenden Dritten gegenüber zu veranschaulichen.

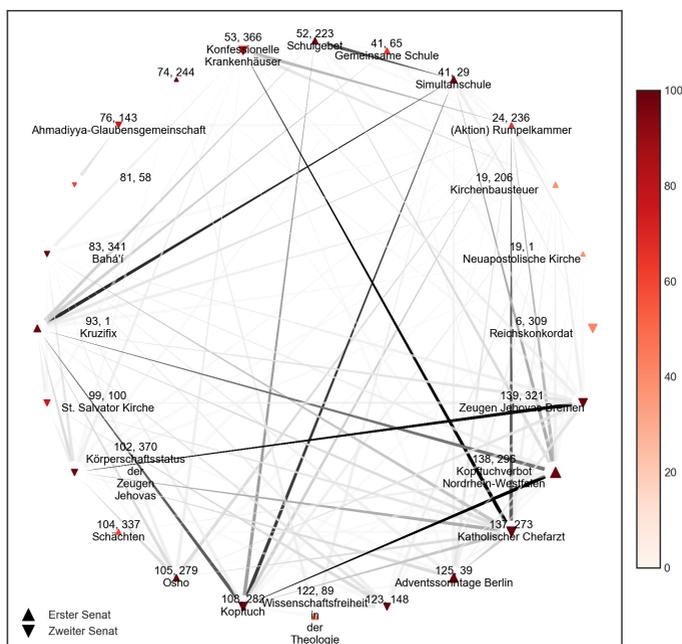


Abbildung 4.37: BVerfGE-Entscheidungen zum Stichwort *Religion* (35 oder mehr Treffer)<sup>170</sup>

Der Eindruck der Problemdiversität in Religionsfragen verstärkt sich noch, wenn man die thematischen Perspektiven zu Christentum, Judentum und Islam betrachtet (Abb. 4.38–4.40, S. 310–312). Einen Schwerpunkt der Diskussionen zum Christentum bildet das Verhältnis von Staat und Kirche (Abb. 4.38, S. 310). Dies gilt insbesondere dort, wo die Kirche an der Wahrnehmung des staatlichen Bildungsauftrags beteiligt ist (z.B. BVerfGE 41, 29 – Simultanschule, BVerfGE 52, 223 – Schulgebet und BVerfGE 93, 1 – Kruzifix), und dort, wo sie als Körperschaft oder Arbeitgeber Sonderrechte für sich in Anspruch nimmt (z.B. BVerfGE 19, 253 – Kirchensteuergesetz, BVerfGE 46, 73 – Stiftungen und BVerfGE 137, 273 – Katholischer Chefarzt).

<sup>170</sup> Zum Ausdruck, nach dem gesucht wurde, bereits Fn. 165 (S. 306); zur Darstellungsweise Fn. 161 (S. 303). Die Entscheidungen ohne Namen sind BVerfGE 74, 244 – [Religionsunterricht]; BVerfGE 81, 58 – [Jeziden I]; und BVerfGE 123, 148 – [Jüdische Gemeinde Brandenburg].

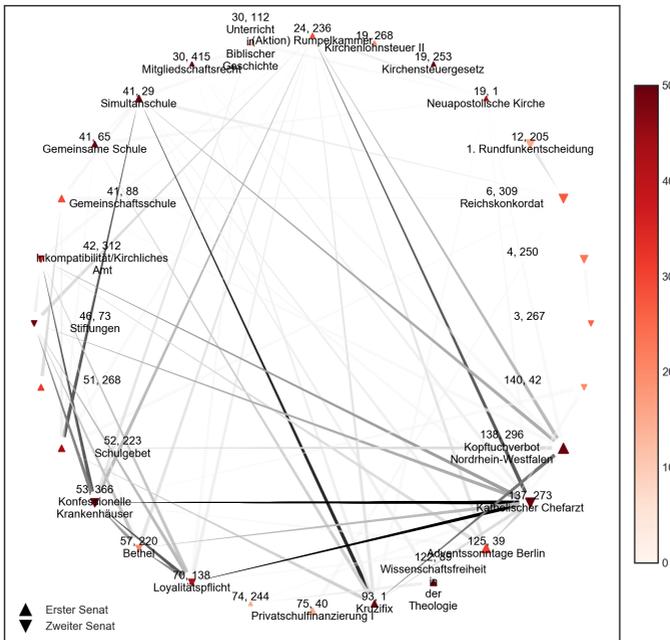


Abbildung 4.38: BVerfGE-Entscheidungen zum Stichwort *Christentum* (15 oder mehr Treffer)<sup>171</sup>

Das Judentum spielt vor allem in Entscheidungen zum Umgang mit dem NS-Unrecht eine Rolle (Abb. 4.39, S. 311), vom Entnazifizierungsprozess (z.B. BVerfGE 6, 132 – Gestapo) über Entschädigungen (z.B. BVerfGE 30, 367 – Bundesentschädigungsgesetz) bis zu den Auswirkungen auf die Meinungsfreiheit (z.B. BVerfGE 90, 241 – Auschwitzlüge). Das Stichwort taucht aber auch in Entscheidungen auf, die primär andere Religionen betreffen (z.B. BVerfGE 137, 273 – Katholischer Chefarzt und BVerfGE 138, 296 – Kopftuchverbot Nordrhein-Westfalen). Die einzige Schwerpunktentscheidung des Zweiten Senats (BVerfGE 123, 148 – [Jüdische Gemeinde Brandenburg]) betrifft die staatliche Neutralitätspflicht bei der Gewährung staatlicher Mittel an Religionsgemeinschaften.

<sup>171</sup> Zum Ausdruck, nach dem gesucht wurde, bereits Fn. 166 (S. 307); zur Darstellungsweise Fn. 161 (S. 303). Die Entscheidungen ohne Namen sind BVerfGE 3, 267 – {Einrichtung von Bekenntnisschulen}; BVerfGE 4, 250 – {Einrichtung von Bekenntnisschulen}; BVerfGE 51, 268 – [Grundschulauflösung]; BVerfGE 74, 244 – [Religionsunterricht]; und BVerfGE 140, 42 – [„Dritter Weg“ im kirchlichen Arbeitsrecht].

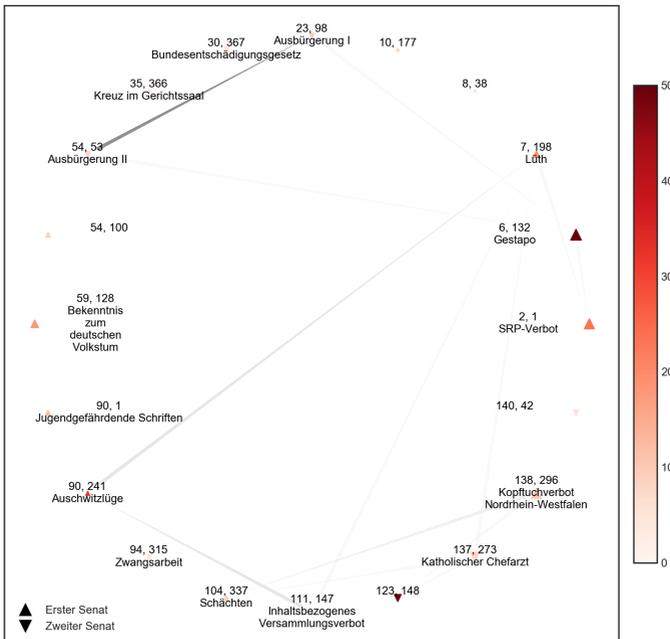


Abbildung 4.39: BVerfGE-Entscheidungen zum Stichwort *Judentum* (fünf oder mehr Treffer)<sup>172</sup>

Das Bild für den Islam, zu dem erst wenige Entscheidungen vorliegen, dominieren drei Problemkontexte (Abb. 4.40, S. 312): *Religionsausübung* (z.B. BVerfGE 104, 337 – Schächten, BVerfGE 108, 282 – Kopftuch und BVerfGE 138, 296 – Kopftuchverbot Nordrhein-Westfalen), *Flucht* (z.B. BVerfGE 54, 341 – Wirtschaftsasyll, BVerfGE 76, 143 – Ahmadiyya-Glaubensgemeinschaft und BVerfGE 83, 216 – Jeziden) und *Terror* (z.B. BVerfGE 115, 320 – Rasterfahndung II, BVerfGE 120, 274 – Online-Durchsuchungen und BVerfGE 133, 277 – Antiterrordateigesetz). Auffällig ist, dass die Rechtsprechung zu den ersten zwei Problemkontexten von beiden Senaten gemeinsam gestaltet wird: BVerfGE 76, 143 – Ahmadiyya-Glaubensgemeinschaft (Zweiter Senat) rezipiert BVerfGE 54, 341 – Wirtschaftsasyll (Erster Senat); BVerfGE 138, 296 – Kopftuchverbot Nordrhein-Westfalen (Erster Senat) bezieht sich häufig auf BVerfGE 108, 282 – Kopftuch (Zweiter Senat). Zum Problemkontext „Terror“ hat sich hingegen bislang allein der Erste Senat geäußert.

<sup>172</sup> Zum Ausdruck, nach dem gesucht wurde, bereits Fn. 167 (S. 307); zur Darstellungsweise Fn. 161 (S. 303). Die Entscheidungen ohne Namen sind BVerfGE 8, 38 – {Härtausgleich (Rechtswegerschöpfung)}; BVerfGE 10, 177 – {Kapitalentschädigung (rechtliches Gehör)}; BVerfGE 54, 100 – {„Verstrickung in ein Unrechtssystem“ (Schuldangemessenheit)}; BVerfGE 123, 148 – [Jüdische Gemeinde Brandenburg]; und BVerfGE 140, 42 – [„Dritter Weg“ im kirchlichen Arbeitsrecht].

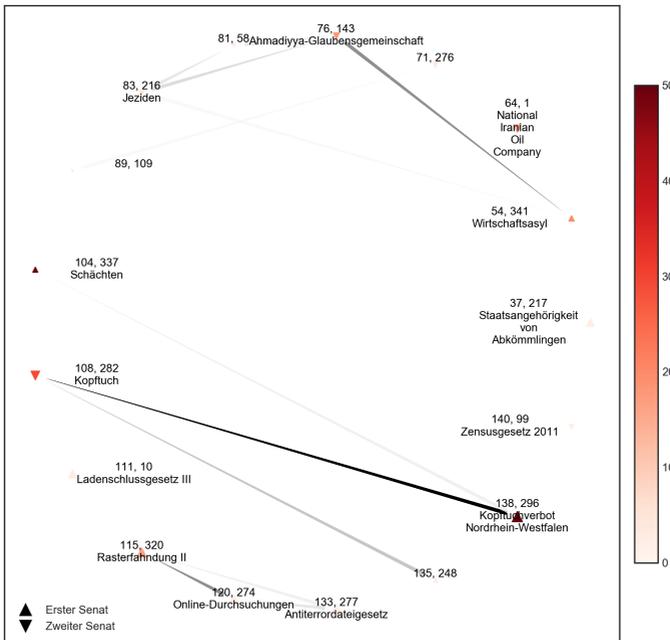


Abbildung 4.40: BVerfGE-Entscheidungen zum Stichwort *Islam* (zwei oder mehr Treffer)<sup>173</sup>

<sup>173</sup> Zum Ausdruck, nach dem gesucht wurde, bereits Fn. 168 (S. 308); zur Darstellungsweise Fn. 161 (S. 303). Die Entscheidungen ohne Namen sind BVerfGE 71, 276 – {Ablehnung Asylklage}; BVerfGE 81, 58 – [Jeziden I]; BVerfGE 89, 109 – {Einreiseverweigerung} und BVerfGE 135, 248 – [Selbstablehnung (Kirchhof)]. BVerfGE 71, 276 – {Ablehnung Asylklage} und BVerfGE 89, 109 – {Einreiseverweigerung} tauchen nur im Bild auf, weil der Schwellenwert hier sehr niedrig gewählt wurde; in der ersten Entscheidung geht es um die verfassungsrechtlichen Grenzen der Ablehnung einer Asylklage als offensichtlich unbegründet (die Treffer in dieser Entscheidung stammen allesamt aus dem Namen einer Widerstandspartei, welcher der Beschwerdeführer anzugehören geltend machte), in der zweiten Entscheidung um die vorläufige Aussetzung der Vollziehung einer Einreiseverweigerung für einen auf dem Luftweg ankommenden Asylbewerber, in deren Rahmen das BVerfG auf „Islamisierungstendenzen“ zu sprechen kommt.

## C. Evaluation

Zu Beginn des Kapitels wurde am Beispiel der Rechtsprechung des BVerfG erläutert, welche Schritte erforderlich sein können, um juristische Daten in eine für die Analyse geeignete Form zu bringen. Die anschließend geschilderten Experimente mit den BVerfGE-Daten vermitteln einen Eindruck davon, wie die Werkzeuge der juristischen Netzwerkforschung zur Untersuchung juristischer Daten eingesetzt werden können. Was folgt daraus für die Rechtswissenschaft? Die kritische Antwort lautet: nicht viel. Denn den zahlreichen Zahlen und Bildern, die präsentiert wurden, lassen sich keine Kausalzusammenhänge entnehmen, und keine einzige normative Frage wurde beantwortet. Bei näherer Betrachtung können aus den dargestellten Bemühungen aber doch einige Erkenntnisse gewonnen werden. Sie betreffen den Untersuchungsinhalt (I.), die Untersuchungsmethoden (II.) und die Untersuchungspraxis (III.).

### *I. Inhaltliche Erkenntnisse*

Bei der Ableitung inhaltlicher Erkenntnisse aus den einzelnen Experimenten ist stets ihre Datengrundlage zu berücksichtigen. Die BVerfGE-Daten enthalten Informationen zu allen Entscheidungen des BVerfG, die in BVerfGE 1–140 veröffentlicht sind. Dieser Ausschnitt der BVerfG-Rechtsprechung ist nicht repräsentativ, nimmt aber für sich in Anspruch, die aus Sicht des Gerichts wichtigsten Entscheidungen zu umfassen.

Auf der Makroebene können zunächst einige juristische Intuitionen bestätigt werden, darunter das „Sommerloch“, die Dominanz einzelner Registerzeichen (allen voran BvR für Verfassungsbeschwerden nach Art. 93 I Nr. 4a und 4b GG) und die zunehmende Länge der einzelnen Entscheidungen. Daneben sind auch einige weniger offensichtliche Trends sichtbar geworden, etwa die Zunahme von Entscheidungen zum Registerzeichen BvE (Verfassungsstreitigkeiten zwischen Bundesorganen nach Art. 93 I Nr. 1 GG) und das hinsichtlich des Rechtsprechungsumfangs leichte Übergewicht des Zweiten Senats seit der Jahrtausendwende. Die Feststellung, dass nicht nur die Entscheidungslängen, sondern auch Innen- und Außengrad (d.h. die Anzahl eingehender und ausgehender Zitate) im BVerfGE-Zitiernetzwerk sehr schief verteilt sind, ist aus juristischer Perspektive wenig überraschend – welche Mechanismen diese Verteilungen hervorbringen, bedarf allerdings der Klärung. Auch das gute Gedächtnis des BVerfG für seine eigene Rechtsprechung, verdeutlicht durch die fortwährende Bezugnahme auf ältere Entscheidungen, verwundert nicht angesichts des Verfahrens, in dem die Entscheidungstexte entstehen. Neu ist allerdings der quantitativ-visuelle Nachweis, dass die jüngere Rechtsprechung vom BVerfG in seinen neuen Entscheidungen intensiver zitiert wird als die ältere Rechtsprechung. Diese Beobachtung stützt die Annahme unterschiedlicher Zwecke von Rechtsprechungszitaten, die teils der inhaltlichen Auseinandersetzung, teils der Verdeutlichung von Rechtsprechungslinien dienen. Sie unterstreicht außerdem, dass sich die Differenzierung zwischen

binärer Zählung (alle zitierten Entscheidungen) und gewichteter Zählung (alle Vorkommen von Zitaten) für die Analyse von Rechtsprechungsziernetzwerken lohnen kann.

Der Mehrwert der Unterscheidung von binärer und gewichteter Zitatanzählung zeigt sich auch auf der Mikroebene. Die für vielzitierte Entscheidungen ermittelten Verhältnisse zwischen Innengrad nach binärer Zählung und Innengrad nach gewichteter Zählung legen hier nahe, dass die Entscheidungen des BVerfG im Zitiernetzwerk unterschiedliche Rollen spielen, nach denen sie sich kategorisieren lassen könnten. Zieht man zur abstrakten Beurteilung von Wichtigkeit bei gewichteter Zählung von Zitaten den temporalen Innengrad heran, so wird deutlich, dass am Ende von Band 140 (d.h. Stand 2015) vor allem Entscheidungen zu den Themenfeldern *Europa*, *Steuern* und *Datenschutz* einen hohen Anteil der nach ihrer Veröffentlichung erfolgten Zitate auf sich vereinigen. Dies könnte als Indiz dafür interpretiert werden, dass sich das BVerfG bei Entscheidungen in diesen Bereichen einem erhöhten Legitimationsdruck ausgesetzt sieht. Die Experimente deuten außerdem daraufhin, dass Entscheidungen mit ähnlichen Themen sich an der parallelen Bewegung ihrer Zitierkurven erkennen lassen, die aus dem temporalen Innengrad bei gewichteter Zählung abgeleitet sind. Legt man auf der Mikroebene keine abstrakte, sondern die vor allem in der Rechtspraxis bedeutsame konkrete, d.h. fallbezogene Definition von Wichtigkeit zugrunde, so können durch Kombination mehrerer Netzwerkmaße Leseempfehlungen für den Fall abgeleitet werden, dass ein Jurist eine bestimmte Entscheidung bereits als relevant identifiziert hat. Eine (auch) netzwerkanalytisch fundierte „More like this“-Funktion für Rechtsprechungsdatenbanken erscheint daher realistisch; welche Daten in die Erzeugung von Leseempfehlungen einfließen sollten und wie diese zur Erstellung einer Rangfolge potenziell relevanter Entscheidungen miteinander zu kombinieren sind, muss sich allerdings im Praxistest zeigen.

Die meisten Einblicke in die Rechtsprechung des BVerfG bietet die Analyse auf der Mesoebene. Hier können mithilfe der Gemeinschaften von  $k$ -Cliquen die *Gleichheit bei Wahlen*, die *kommunale Selbstverwaltung*, der *Datenschutz* und die *Rundfunkordnung* als juristische Problemfelder identifiziert werden: Themenbereiche, zu denen sich das BVerfG immer wieder äußern muss und in denen es (offenbar) besonders systematisch auf einen Kanon früherer Entscheidungen Bezug nimmt. Die algorithmische Gemeinschaftserkennung scheint grundsätzlich in der Lage, die BVerfGE-Entscheidungen nach Rechtsbereichen zu ordnen: die Visualisierung des Gemeinschaftsgraphen liefert Ansätze zu einer Kartographie des Verfassungsrechts; die Aggregation mehrerer Resultate zu Konsensgemeinschaften zeigt, dass die Gemeinschaftsstruktur des BVerfGE-Zitiernetzwerks relativ robust ist. Die umfassende Interpretation der algorithmisch erzielten Resultate, ihre rechtswissenschaftliche Bewertung und die Suche nach dem besten Algorithmus werden allerdings noch einige wissenschaftliche Ressourcen beanspruchen. Das gilt auch für die systematische Verknüpfung von Zitierdaten mit Informationen, die sich aus den Entscheidungstexten entnehmen lassen. Die Erkundung der zeitli-

chen und thematischen Begriffskontexte zu den Stichwörtern *Europa*, *Datenschutz* und *Religion* sowie *Christentum*, *Judentum* und *Islam* anhand ihrer textbasierten Subgraphen zeigt jedoch, dass sich der Aufwand lohnen kann: Eine quantitativ fundierte, visuelle Repräsentation einschlägiger Rechtsprechung erleichtert dem juristischen Beobachter die Auswahl von Entscheidungen zur vertieften Lektüre und vermag seinen Blick auf Muster, Entwicklungen und Anomalien zu lenken, deren rechtswissenschaftliche Aufarbeitung besonderen Erkenntnisgewinn verspricht.

## II. Methodische Erkenntnisse

Die Möglichkeiten zur Herleitung methodischer Erkenntnisse aus den Experimenten mit den BVerfGE-Daten sind in zweierlei Hinsicht eingeschränkt: Erstens wurde nur ein Bruchteil aller methodischen Werkzeuge der juristischen Netzwerkforschung erprobt. Viele beliebte Maße wurden ausgespart (z.B. der PageRank auf der Mikroebene), andere wurden lediglich statisch betrachtet (z.B. die Gemeinschaftserkennung auf der Mesoebene). Das Zitiernetzwerk ist ein Modell der BVerfG-Rechtsprechung und die Zitiergraphen sind Modelle des Zitiernetzwerks, doch Modellierung im Sinne von Kapitel 3 Abschnitt D. (S. 193–223) kommt in den Experimenten nicht vor. Zweitens wurden die ausgewählten Werkzeuge nur an einem sehr konkreten Fall getestet: der Rechtsprechung des BVerfG, die in BVerfGE 1–140 veröffentlicht ist. Damit wurde lediglich gezeigt, wie sich die Anwendung der verwendeten Methoden auf juristische Daten im Einzelfall darstellen kann. Offen bleibt also insbesondere die Frage, inwieweit die Beobachtungen auf andere juristische Daten oder nicht-juristische (Zitier-)Daten verallgemeinerbar sind.

Die Experimente auf der Makroebene deuten daraufhin, dass visuelle und quantitative Erkundungen der als Graphen zu untersuchenden Daten auch abseits originär netzwerkanalytischer Methoden sinnvoll sein können. Diese Erkundungen sollten als Bestandteil der juristischen Netzwerkforschung begriffen werden. Die Frage, welchen abstrakten Verteilungen genau die Gradverteilungen juristischer Netzwerke folgen, ist kaum akkurat zu beantworten und verspricht wenig juristische Erkenntnis. Demgegenüber kann die Betrachtung zweidimensionaler Verteilungen (Beispiel: Vermischung von Entscheidungen nach dem Entscheidungsjahr) bei juristischen Netzwerken, zu denen viele Metadaten vorliegen, durchaus interessante Einblicke liefern.

Die Experimente auf der Mikroebene zeigen die Probleme von Innen- und Außengrad bei zeitabhängigen Zitiernetzwerken, deren zugehörige Graphen annähernd azyklisch sind. Es wird aber auch deutlich, dass sich die festgestellten Schwierigkeiten durch die Einführung zeitsensitiver Maße abmildern lassen, deren Entwicklung über die Zeit anhand von Zitierkurven verfolgt werden kann. Die Experimente legen nahe, dass insbesondere die Zitierkurven auf Basis der gewichteten Zählung wertvolle Informationen dazu enthalten, welche Entscheidungen ähnliche Themen behandeln. Schließlich lenkt die Differenzierung zwischen

abstrakter und konkreter Wichtigkeit den Blick auf die Potenziale netzwerkanalytischer Maße für juristische Empfehlungssysteme. Die Leseempfehlungen, die für BVerfGE-Entscheidungen aus der Kombination mehrerer auf der Mikroebene angesiedelter Netzwerkmaße mithilfe der gewichteten Zitatanzahl konstruiert werden können, lassen vermuten, dass diese Potenziale beträchtlich sind.

Für die Untersuchung juristischer Netzwerke auf der Mesoebene scheinen die hergebrachten Methoden der Netzwerkforschung zum Teil ohne größere Anpassungen nützlich zu sein. Sowohl mit Gemeinschaften von  $k$ -Cliquen als auch mit algorithmischer Gemeinschaftserkennung können hier interessante Einblicke gewonnen werden. Bei nichtdeterministischen Algorithmen zur Gemeinschaftserkennung ermöglicht die Aggregation mehrerer Durchläufe abgesicherte Resultate. Die Frage, welcher Algorithmus zur Strukturierung juristischer Daten am besten geeignet ist, bedarf allerdings nach wie vor der Klärung. Darüber hinaus besteht auf der Mesoebene für die juristische Netzwerkforschung großes Innovationspotenzial – sowohl hinsichtlich der Visualisierung von Ergebnissen bekannter Verfahren als auch bei der Entwicklung neuer Verfahren zur Quantifizierung, die verschiedenartige Informationen aus juristischen Daten fruchtbar machen.

Insgesamt lässt sich festhalten, dass vor allem diejenigen Methoden der juristischen Netzwerkforschung Erkenntnis versprechen, die auf die Besonderheiten ihres Untersuchungsgegenstands Rücksicht nehmen. Anstatt die Praktiken anderer Netzwerkforscher (seien sie nun Natur- oder Sozialwissenschaftler) weitgehend unverändert zu übernehmen, sollte die juristische Netzwerkforschung daher ihre eigenen Werkzeuge entwickeln, um den Besonderheiten juristischer Netzwerke gerecht zu werden. Lohnen könnten sich insbesondere die differenzierte Erhebung von Zitaten (Beispiel: gewichtete Zählung mit Zitierblockerfassung), die Konstruktion zeitsensitiver Maße (Beispiel: temporaler Innengrad) und die Kombination von Zitierinformationen mit anderen Informationen, die sich juristischen Texten entnehmen lassen (Beispiel: Vorkommen bestimmter Stichwörter). Die Methodenentwicklung in der juristischen Netzwerkforschung setzt zwar eine höhere Vertrautheit mit den Grundlagen der Graphentheorie voraus, als sie zur Methodenanwendung erforderlich ist. Da eine solche Vertrautheit auch differenziertere juristische Netzwerkstudien ermöglicht, dürften sich die zu ihrem Erwerb erforderlichen Investitionen allerdings lohnen.

### *III. Praktische Erkenntnisse*

Jenseits der inhaltlichen und methodischen Erkenntnisse, die aus der Datenanalyse im Rahmen der Experimente gewonnen werden können, vermittelt die Arbeit mit den BVerfGE-Daten auch einige praktische Einsichten. Manche dieser Einsichten betreffen speziell die juristische Netzwerkforschung, andere lassen sich auf quantitative juristische Studien im Allgemeinen übertragen. Die Erkenntnisse betreffen vor allem die Sammlung juristischer Daten und die Kommunikation der Ergebnisse ihrer Analyse.

Mit Blick auf die Sammlung juristischer Daten ist zunächst hervorzuheben, dass sie einen wesentlichen Bestandteil quantitativer juristischer Studien darstellt. Die Datensammlung kann aber zahlreiche Wertungsentscheidungen und Komplikationen mit sich bringen, für deren umsichtige Behandlung zum Teil umfassende Kenntnisse der rechtlichen Besonderheiten des untersuchten Bereichs erforderlich sind. Bei der Delegation der Datensammlung an Fachkräfte mit eingeschränkten juristischen Kenntnissen ist daher ebenso Vorsicht geboten wie bei der Nutzung fremder Datensätze, deren Zustandekommen nur lückenhaft dokumentiert ist. In dieser Hinsicht könnte die Entwicklung von *Best Practices* für die Sammlung juristischer Daten und für die Dokumentation der einzelnen Arbeitsschritte quantitative juristische Studien deutlich erleichtern.<sup>174</sup> Offen ist außerdem die Frage, wie sich die Qualität juristischer Daten auf die Ergebnisse juristischer Netzwerkstudien auswirkt.

Die Quellenlage für juristische Daten ist prekär – selbst dort, wo es um Gerichtsentscheidungen geht, die grundsätzlich gemeinfrei sind. Die Ungenauigkeiten in den BVerfGE-Daten, die aus den *juris*-Annotationen resultieren, unterstreichen die Notwendigkeit, juristische Daten für die Forschung mittelfristig direkt von ihren Produzenten (z.B. den Gerichten) zu beziehen, welche die Daten in standardisierten Austauschformaten über das Internet zugänglich machen sollten (z.B. über Programmierschnittstellen). Wie unter anderem das Experiment zur algorithmischen Gemeinschaftserkennung zeigt, erweist sich die im Falle des BVerfG jedenfalls teilweise vorhandene Benennung von Entscheidungen bei der Untersuchung von höchstrichterlicher Rechtsprechung als äußerst hilfreich. Das BVerfG geht mit der Liste der BVerfGE-Entscheidungen auf seiner Website, in der viele jüngere Entscheidungen einen Namen tragen, grundsätzlich einen Schritt in die richtige Richtung; wissenschaftlich wünschenswert wäre die Ergänzung von Namen auch für alle bisher unbenannten Entscheidungen. Wie das Beispiel der BVerfGE-Daten zeigt, können wenige Doppelfundstellen bei der Zusammenstellung von Zitierdaten zur Rechtsprechung erheblichen Mehraufwand verursachen und unerwartete Wertungsentscheidungen erforderlich machen. Für Entscheidungssammlungen, deren Fundstellen nur mit Band und Startseite referenziert werden, wäre es daher wünschenswert, dass jede Entscheidung auf einer neuen Seite beginnt, damit alle Entscheidungen anhand ihrer Kurzfundstelle eindeutig identifiziert werden können.

Speziell für die Sammlung von Zitierdaten kann festgehalten werden, dass die gewichtete Zählung die Analysemöglichkeiten erheblich erweitert, weshalb sie in Studien zu juristischen Zitiernetzwerken nicht länger die Ausnahme, sondern vielmehr die Regel sein sollte. Ähnliches lässt sich für die Erfassung von Informationen zu den Zitierblöcken vermuten, in denen die einzelnen Zitate gefunden werden, wenngleich diese Informationen in den geschilderten Experimenten noch

---

<sup>174</sup> Einen möglichen Ansatzpunkt für solche *Best Practices* bieten die Regeln von *Krishnamurthy u.a.*, *Computer Communications* 34 (2011), S. 43–53, für die allgemeine Netzwerkforschung aufbereitet und ergänzt von *Zweig*, *Network Analysis Literacy*, 2016, S. 353–358.

nicht genutzt wurden. Wie die Visualisierungen der BVerfGE-Daten zeigen, lohnt sich auch die Anreicherung der Zitierdaten um Metadaten; der damit verbundene Mehraufwand ist gering, vervielfacht aber die Untersuchungsmöglichkeiten.

Zur Kommunikation von Analyseergebnissen ist zunächst zu konstatieren, dass es in der Verantwortung des Untersuchenden liegt, den Rezipienten auf die Interpretationseinschränkungen hinzuweisen, die mit dem Zuschnitt des Datensatzes einhergehen, und ihn in die Lage zu versetzen, die Analyse selbst kritisch zu hinterfragen. Hierzu sind sehr viel ausführlichere Erläuterungen zum Zustandekommen des Datensatzes erforderlich, als sie in vielen Studien zur Netzwerkforschung im Recht (und auch in quantitativen juristischen Studien) bislang üblich sind. Um die Nachvollziehbarkeit und Reproduzierbarkeit der geschilderten Resultate zu sichern, sollte (im Rahmen des rechtlich Zulässigen) sowohl die Datengrundlage als auch der zur Analyse verwendete Code öffentlich zugänglich gemacht werden.<sup>175</sup> Hier stoßen klassische juristische Publikationsformate, allen voran die für den Druck optimierte Monographie, an ihre Grenzen.

Dass die Buchform für quantitative juristische Studien im Allgemeinen und für juristische Netzwerkstudien im Besonderen möglicherweise nicht ideal ist, wird in den Experimenten mit den BVerfGE-Daten auch daran deutlich, dass sie sich auf einen Bruchteil des Analysierbaren beschränken (müssen), dessen Auswahl durch die Verfasserin für viele Leser unbefriedigend sein wird – sei es, dass sie an einzelnen Untersuchungsaspekten kein Interesse haben, an manchen Stellen mehr Vertiefung wünschen oder gern andere Fragen an die Daten herangetragen sähen. Obwohl die Selektion von Informationen traditionell zu den Aufgaben der Wissenschaftskommunikation zählt, scheint es gerade bei der Untersuchung juristischer Daten wünschenswert, den Rezipienten zu ermöglichen, diese in interaktiven Formaten quantitativ und visuell zu erkunden, damit sie ihre eigenen Untersuchungsschwerpunkte setzen, ihre eigenen Forschungsfragen entwickeln und diese Fragen durch Arbeit mit den Daten selbst beantworten können. Dass der Forscher, der die Daten zusammengestellt hat, daneben die Ergebnisse seiner Analysen und mögliche Interpretationen anbietet, steht dem nicht entgegen. Ihm kommt im Rahmen quantitativer juristischer Studien regelmäßig eine Doppelrolle zu: Er stellt Daten zur Verfügung und bietet seine Erkenntnisse daraus an – er macht eine Goldmine zugänglich und fördert selbst Gold.

Eine Möglichkeit, ganz unterschiedlichen Interessenten die Chancen der quantitativen juristischen Forschung und insbesondere der juristischen Netzwerkforschung zu eröffnen, besteht langfristig darin, speziell für diese Zwecke entwickelte Anwendungen bereitzustellen (z.B. als Web-Applikationen oder digitale Installationen). Diese Anwendungen können je nach Zielgruppe darauf ausgerichtet sein,

---

<sup>175</sup> Das spricht gegen die Verwendung von Programmen mit einer graphischen Benutzeroberfläche (z.B. Microsoft Excel), da sich die Interaktion des Forschers mit dieser Benutzeroberfläche im Nachhinein kaum zuverlässig rekonstruieren lässt (wie die immer wieder mediales Aufsehen erregenden Fälle von Kalkulationsfehlern in wissenschaftlichen Publikationen aufgrund fahrlässig nicht markierter Excel-Zellen zeigen).

Rechtswissenschaftlern ohne Programmierkenntnisse reproduzierbare Untersuchungen juristischer Daten zu ermöglichen, Studenten an einzelne Rechtsgebiete heranzuführen oder Laien die Welt des Rechts nahezubringen. Sie können sogar dazu dienen, (angehenden) Juristen die Grundlagen der Statistik oder der Informatik anhand von Daten zu vermitteln, die für sie relevant und interessant sind. Welche Kommunikationsformen sich für juristische Daten in welchen Kontexten besonders eignen, können nur weitere Experimente zeigen. Die BVerfGE-Daten mögen hierfür als Rohstoff dienen.

## Kapitel 5

### Fazit

Die vorgelegte Arbeit dient der Einführung in die juristische Netzwerkforschung. Im Folgenden sind die wesentlichen Punkte dieser Einführung zusammengefasst (A.). Abschließend werden einige Herausforderungen der juristischen Netzwerkforschung skizziert, deren Bewältigung über ihre Zukunft in Rechtswissenschaft und Rechtspraxis mitentschieden wird (B.).

#### A. Rückblick

##### *Kapitel 1. Einleitung*

1. *Graphen* sind mathematische Objekte, die sich aus einer Menge von *Punkten* und einer Menge von *Linien* zusammensetzen. Ihrer Untersuchung widmet sich die mathematische *Graphentheorie*. Die Wurzeln der Graphentheorie reichen zurück bis ins 18. Jahrhundert, als *Leonhard Euler* das *Königsberger Brückenproblem* löste.
2. Die *Netzwerkforschung* baut auf der Graphentheorie auf; sie befasst sich mit der Untersuchung von Netzwerken. Ein Netzwerk ist die Zusammenfassung einer Menge von *Objekten* und einer Menge von *Beziehungen* (zwischen diesen Objekten) zu einer Einheit. Beispiele sind Infrastrukturnetzwerke (Bahnhöfe und Gleisabschnitte), Informationsnetzwerke (Webseiten und Hyperlinks) und soziale Netzwerke (Menschen und soziale Beziehungen). Netzwerke sind Konstruktionen in der Realität, Graphen sind ihre Gegenstücke in der Mathematik. Ein Netzwerk lässt sich als Graph darstellen, indem seine Objekte in Punkte und seine Beziehungen in Linien übersetzt werden.
3. Die *juristische Netzwerkforschung* widmet sich der Untersuchung von Netzwerken im Recht. Sie fragt, wie sich Netzwerke im Recht definieren, in Graphen übersetzen und als Graphen analysieren lassen. Und sie ergründet, wie die Ergebnisse solcher Analysen interpretiert und für das Verständnis von Recht fruchtbar gemacht werden können.

## Kapitel 2. Theorie

1. Eine erste Annäherung an die juristische Netzwerkforschung ermöglicht die Literatur zur *Netzwerkforschung im Recht*: Beiträge, die mit Werkzeugen aus der Netzwerkforschung Fragen untersuchen, die das Recht betreffen. Diese Beiträge stammen aus den verschiedensten Disziplinen, darunter viele verschiedene Sozialwissenschaften und Naturwissenschaften, und werden in den unterschiedlichsten Medien publiziert. Sie befassen sich mit juristischen Texten, juristischen Akteuren und den Wechselwirkungen zwischen Recht und Realität, konzentrieren sich aber selten auf die spezifisch juristischen Aspekte ihrer Untersuchungsgegenstände. Dies unterscheidet die Netzwerkforschung im Recht von der juristischen Netzwerkforschung: *Juristische Netzwerkforschung* ist Netzwerkforschung, die das Recht in den Fokus rückt.
2. Die erste Weichenstellung der juristischen Netzwerkforschung ist die Definition ihres Gegenstands: Was sind die Objekte und Beziehungen, deren Untersuchung als Netzwerke die Rechtswissenschaft voranbringen könnte? Bei der Beantwortung dieser Frage hilft eine Taxonomie der Objekte und Beziehungen, deren Analyse juristischen Erkenntnisgewinn verspricht. Unter den Objekten wird zwischen *Rechtssubjekten* (Individuen und Kollektive), *Rechtstexten* (Gerichtsentscheidungen, Gesetzestexte, Verwaltungsentscheidungen, Verträge und Literatur) und *Rechtsbegriffen* differenziert. Beziehungen werden nach ihrem *Grad*, ihrem *Profil* und ihrer *Form* unterschieden; sie lassen sich in verschiedene Arten von Linien übersetzen, die verschiedene Typen von Graphen definieren, für die unterschiedliche Analysemöglichkeiten offenstehen (ungerichtete und gerichtete Graphen, ungewichtete und gewichtete Graphen, schlichte Graphen und Multigraphen, unipartite und bipartite Graphen, Hypergraphen). Da die Netzwerkdefinition im Detail zahlreiche Wertungsentscheidungen erfordert, von denen die Ergebnisse späterer Analysen abhängen, sollte sie in juristischen Netzwerkstudien ausführlich diskutiert werden.
3. Die juristische Netzwerkforschung ist durch ein Bündel von Untersuchungsperspektiven gekennzeichnet, die nach Untersuchungszielen (*Exploration*, *Explikation*, *Extrapolation*) und Untersuchungsebenen (*Mikroebene*, *Mesoebene*, *Makroebene*) unterschieden werden können. Bei der Exploration steht die Aufdeckung von Mustern in Daten, bei der Explikation die Erklärung gefundener Muster und bei der Extrapolation die Vorhersage von Daten anhand bekannter Muster im Vordergrund. Auf jeder Untersuchungsebene stehen andere Fragen zu den Strukturen des betrachteten Netzwerks im Mittelpunkt, die als *Wichtigkeitsfrage* (Mikroebene), *Gruppierungsfrage* (Mesoebene) und *Charakterisierungsfrage* (Makroebene) gekennzeichnet werden.

### Kapitel 3. Methodik

1. Daten sind das Arbeitsmaterial der juristischen Netzwerkforschung. Sie können aus der Realität stammen (*reale Daten*) oder speziell für Untersuchungszwecke erstellt worden sein (*fiktive Daten*). Unter den realen Daten, die für die Rechtswissenschaft interessant sind, lassen sich *juristische Daten im weiteren Sinne* und *juristische Daten im engeren Sinne* unterscheiden. Juristische Netzwerkforschung, die mit realen Daten arbeitet, steht bei der Wahl ihrer Datenquelle(n) vor *inhaltlichen, technischen* und *rechtlichen Schwierigkeiten*; bei der Zusammenstellung der Daten können *Abdeckungs-, Aufzeichnungs- und Aufbereitungsfehler* auftreten. Für eine Einführung in die Methoden der juristischen Netzwerkforschung bietet sich es sich daher an, auf fiktive Daten zurückzugreifen. Hier werden zwei Datensätze aus dem fiktiven Staat namens *Flatland* vorgestellt, von denen einer Rechtsprechungsdaten und der andere Unternehmensdaten enthält. Anhand dieser Datensätze werden einige Grundbegriffe der deskriptiven Statistik eingeführt, die für die juristische Netzwerkforschung bedeutsam sind, allen voran die Unterscheidung von *Merkmalen* nach ihrem *Skalenniveau* (nominal, ordinal oder metrisch).
2. Die methodischen Säulen der juristischen Netzwerkforschung sind Visualisierung, Quantifizierung und Modellierung.
  - (a) *Visualisierung* bezeichnet die Kommunikation von Informationen mittels graphischer Repräsentationen – nicht nur als Prozess, sondern auch als Medium dieses Prozesses. In der juristischen Netzwerkforschung steht die Visualisierung von Daten im Vordergrund, in deren *Visualisierungsprozess* die Zuordnung von Merkmalen (z.B. Datum, Länge und Spruchkörper einer Gerichtsentscheidung) zu *visuellen Variablen* (z.B. Position, Größe und Farbton) eine zentrale Rolle spielt. Zur Visualisierung von Netzwerkdaten als Graphen können *Graphdiagramme* und *Graphmatrizen* eingesetzt werden.
  - (b) *Quantifizierung* ermöglicht intersubjektiv überprüfbare Aussagen zu juristischen Netzwerken, indem sie die ihnen entsprechenden Graphen analysiert. Die Basis der Quantifizierung bilden *Bewegungen auf Graphen*, mit deren Hilfe die wichtigsten Grundbegriffe der Graphentheorie eingeführt werden. Zur Beantwortung von Wichtigkeits-, Gruppierungs- und Charakterisierungsfrage, die den Schwerpunkt der Analysen auf Mikro-, Meso- und Makroebene bilden, können im Rahmen der Quantifizierung viele verschiedene Maße fruchtbar gemacht werden, die in der allgemeinen Netzwerkforschung ihren Ursprung haben. Dabei sind die Besonderheiten juristischer Netzwerke zu berücksichtigen. Speziell für *juristische Ziternetzwerke* kann es sich lohnen, zwischen der *binären Zählung* und der *gewichteten Zählung* von Zitaten sowie dem Auftreten mehrerer Zitate in demselben *Zitierblock* oder (nur) in demselben Dokument zu unterschei-

den. Zu den konzeptuellen Schwerpunkten zählen auf der Mikroebene die *Zentralitäten*, auf der Mesoebene die *Gemeinschaften* und auf der Makroebene die *Verteilungen*. Auf allen Untersuchungsebenen stellt die *Zeit als Zusatzfaktor* die juristische Netzwerkforschung vor die Herausforderung, die *statische Analyse* von Momentaufnahmen einzelner Netzwerke um eine *dynamische Analyse* zu erweitern. Ein erster Schritt dazu, dieser Herausforderung speziell für juristische Zitiernetzwerke gerecht zu werden, ist die Definition des *temporalen Innengrads*, aus dem sich juristische *Zitierkurven* ableiten lassen.

- (c) *Modellierung* bezeichnet das indirekte Studium realer Systeme durch die Konstruktion und Analyse von Modellen. *Modelle* sind interpretierte Strukturen, also Verbindungen einer Struktur mit einer Deutung dieser Struktur. Für die juristische Netzwerkforschung sind vor allem *mathematische Modelle* und *algorithmische Modelle* von Bedeutung, die sich nach ihren primären Gegenständen in Modelle für *Netzwerkstrukturen* (Verdrahtungsmodelle und Wachstumsmodelle) und Modelle für *Netzwerkprozesse* (Modelle für Suchprozesse und Modelle für Diffusionsprozesse) unterscheiden lassen. Modellierung ermöglicht Aussagen über Funktion und Robustheit ganzer Klassen von Netzwerken, die insbesondere im Regulierungsrecht praktische Relevanz entfalten können.
3. Die juristische Netzwerkforschung konzentriert sich bisher auf die Untersuchung von Netzwerken mit einer Art von Objekten und einer Art von Beziehungen; ihre Methoden sind auf die Untersuchung solcher Netzwerke zugeschnitten. Viele juristisch interessante Phänomene lassen sich aber besser als *mehrschichtige Netzwerke* beschreiben, die Objekte und Beziehungen mehrerer verschiedener Arten enthalten. Um dauerhaft zum Erkenntnisgewinn in der Rechtswissenschaft beitragen zu können, muss die juristische Netzwerkforschung Methoden entwickeln, die zur Untersuchung mehrschichtiger Netzwerke geeignet sind.

#### Kapitel 4. Praxis

1. Die Arbeitsweise der juristischen Netzwerkforschung wird anhand einer Fallstudie zur Rechtsprechung des Bundesverfassungsgerichts (BVerfG) illustriert. Bei der Zusammenstellung der hierfür benötigten Daten wird im Rahmen der *Netzwerkdefinition* die Festlegung getroffen, dass alle Entscheidungen des Gerichts, die in BVerfGE 1–140 veröffentlicht sind (*BVerfGE-Entscheidungen*), als Objekte dargestellt und alle Zitate im Format „{Band}, {Startseite}“ als Beziehungen repräsentiert werden sollen, wobei zu Entscheidungen und Zitaten möglichst viele Metadaten erhoben werden sollen. Die *Netzwerkkonstruktion* erfolgt in drei Schritten (Sammlung der Entscheidungen, Extraktion der Zitate und Aufbereitung der Daten), aus denen das *BVerfGE-Korpus* als Sammlung der Volltexte von BVerfGE 1–140, eine *Fundstellenliste* mit Metadaten, eine Zi-

*tatliste* sowie zwei Graphen (*BVerfGE-MultiDiGraph* und *BVerfGE-DiGraph*) hervorgehen. Diese *BVerfGE-Daten* bilden die Grundlage der sich anschließenden Analysen.

2. Die *BVerfGE-Daten* werden in einer Reihe von *Experimenten* untersucht, deren Ziel die *Exploration* ist. Dabei werden die verwendeten Methoden der juristischen Netzwerkforschung nicht nur erprobt, sondern auch gezielt angepasst, um den Besonderheiten von *Rechtsprechungszitiernetzwerken* Rechnung zu tragen. Auf der *Makroebene* werden visuelle und quantitative Zusammenfassungen der *BVerfGE-Daten* erstellt und einige Strukturen des *BVerfGE-Zitiernetzwerks* charakterisiert, bevor auf der *Mikroebene* nach *abstrakt wichtigen* und *konkret wichtigen* *BVerfGE-Entscheidungen* gesucht wird. Die Experimente auf der *Mesoebene* nutzen Methoden der Gemeinschaftserkennung, um die Rechtsprechung des *BVerfG* zu ordnen, und erkunden die Begriffskontexte zu einzelnen Stichwörtern anhand von *textbasierten Subgraphen*.
3. Aus den Experimenten und den dafür erforderlichen Vorbereitungen lassen sich einige inhaltliche, methodische und praktische Erkenntnisse ableiten. Eine wichtige *inhaltliche Erkenntnis* ist, dass die juristische Netzwerkforschung juristische Intuitionen quantitativ und visuell untermauern sowie die Aufmerksamkeit juristischer Beobachter auf Muster und Anomalien lenken kann, deren nähere Untersuchung weiterführende juristische Einsichten erwarten lässt. Als zentrale *methodische Erkenntnis* kann festgehalten werden, dass die Methoden der juristischen Netzwerkforschung vor allem dann Erkenntnisgewinn versprechen, wenn sie speziell auf ihren jeweiligen Untersuchungsgegenstand zugeschnitten sind und dessen juristische Besonderheiten berücksichtigen. Eine wesentliche *praktische Erkenntnis* ist, dass die juristische Netzwerkforschung davon profitieren würde, wenn sie bei der Zusammenstellung ihrer Daten nach *Best Practices* verfahren, dabei frei zugängliche, verlässliche Quellen auf Grundlage standardisierter Austauschformate nutzen und die Ergebnisse ihrer Analysen in interaktiven digitalen Formaten kommunizieren könnte.

## B. Ausblick

Diese Arbeit skizziert, wie juristische Netzwerkforschung aussehen kann. Sie identifiziert einige Bereiche, in denen die juristische Netzwerkforschung zum Erkenntnisgewinn im Recht beitragen kann, erhebt aber nicht den Anspruch, die Frage nach dem Nutzen der juristischen Netzwerkforschung für Rechtswissenschaft und Rechtspraxis zu beantworten. Wie die Antwort auf diese Frage lautet, hängt unter anderem davon ab, inwieweit die juristische Netzwerkforschung drei Herausforderungen bewältigen kann, die jeweils inhaltliche, methodische und praktische Aspekte aufweisen. Sie lauten: *mehr Breite*, *mehr Tiefe* und *mehr Vernetzung*.

### *Mehr Breite*

Ein Großteil der juristischen Netzwerkforschung konzentriert sich auf Rechtsprechungszitiernetzwerke. Diese Arbeit macht davon keine Ausnahme, enthält aber die erste größere Netzwerkstudie zur Rechtsprechung eines deutschen Gerichts. Durch diese inhaltliche Konzentration lässt sich das Erkenntnispotenzial der juristischen Netzwerkforschung insgesamt kaum abschätzen. *Inhaltlich* meint die Forderung nach mehr Breite daher vor allem zweierlei: Mehr Studien zu juristischen Netzwerken, die nicht Rechtsprechungszitiernetzwerke sind, und – speziell aus der Perspektive der deutschen Rechtswissenschaft – mehr Studien zur Rechtsprechung deutscher Gerichte.

Viele juristische Netzwerkstudien begnügen sich bisher damit, einige Standardwerkzeuge der allgemeinen Netzwerkforschung im juristischen Kontext anzuwenden. Das Methodenangebot der allgemeinen Netzwerkforschung wird dabei kaum abgedeckt, auch Anpassungen an den juristischen Untersuchungsgegenstand sind selten. *Methodisch* meint die Forderung nach mehr Breite daher, dass mehr verschiedene Methoden erprobt und mehr *rechtsspezifische* Methoden entwickelt werden sollten. Die Experimente mit den BVerfGE-Daten versuchen, dem ein Stück weit zu entsprechen, schöpfen die Untersuchungsmöglichkeiten aber bei Weitem nicht aus.

Die juristische Netzwerkforschung wird aktuell von einigen wenigen Forschungsgruppen und Forschern vorangetrieben, die mit einigen wenigen Datensätzen arbeiten (das gilt international, aber besonders in Deutschland). Der Erkenntnisfortschritt in der juristischen Netzwerkforschung ist daher von einer kleinen Gruppe von Individuen abhängig; seine Datengrundlage ist schmal. *Praktisch* bedeutet die Forderung nach mehr Breite somit einerseits, dass mehr verschiedene Forscher sich dem weitgehend unbestellten Forschungsfeld annehmen sollten. Sie bedeutet aber auch, dass mehr verschiedene Datensätze erstellt und untersucht werden sollten – nicht nur zu unterschiedlichen Inhalten, sondern auch zu identischen oder ähnlichen Inhalten. So können Fehler in existierenden Datensätzen offengelegt, existierende Resultate bestätigt oder in Zweifel gezogen und neue Analysen ermöglicht werden.

### *Mehr Tiefe*

Viele juristische Netzwerkstudien liefern vor allem Zahlen und (in geringerem Umfang) Bilder, schrecken aber vor der juristischen Interpretation ihrer Resultate zurück. Die Experimente zu den BVerfGE-Daten sind hier keine Ausnahme, da sie vor allem auf die Erprobung einzelner Methoden der juristischen Netzwerkforschung abzielen. *Inhaltlich* verlangt die Forderung nach mehr Tiefe daher – auf die Analyse bezogen – mehr juristische Interpretation und – auf die Forscher gemünzt – mehr Interesse an juristischen Details. Damit geht allerdings zwangsläufig weniger methodische Breite innerhalb einer einzelnen Studie einher, sodass die hier

formulierten Herausforderungen insgesamt nur durch das Zusammenwirken vieler verschiedener Studien adressiert werden können.

Der Großteil der juristischen Netzwerkforschung wendet die einzelnen Methoden der allgemeinen Netzwerkforschung bisher in einer einzigen Form an, ohne mögliche Varianten zu diskutieren. Das trifft auch auf die Experimente zu den BVerfGE-Daten zu, ist es doch etwa im Rahmen der algorithmischen Gemeinschaftserkennung auf der Mesoebene wünschenswert, nicht nur die Ergebnisse verschiedener Läufe desselben (nichtdeterministischen) Algorithmus, sondern auch die Ergebnisse verschiedener Algorithmen miteinander zu vergleichen, und nicht nur die Gemeinschaften zu einem bestimmten Zeitpunkt, sondern auch ihre Entwicklung über die Zeit zu untersuchen. *Methodisch* bedeutet die Forderung nach mehr Tiefe daher, dass künftige Untersuchungen die Potenziale einzelner Methoden in größerem Umfang ausschöpfen und verstärkt die kritische Auseinandersetzung mit den verschiedenen Herangehensweisen zur Lösung desselben Problems suchen sollten.

Juristische Netzwerkstudien werden häufig als relativ kurze Aufsätze publiziert, die mit einer Einführung in die Methoden der Netzwerkforschung beginnen, den verwendeten Datensatz knapp skizzieren und dann auf wenigen Seiten einige Analyseergebnisse zusammenstellen. In diesem Publikationsformat sind inhaltliche und methodische Tiefe kaum zu erreichen, zumal die Erläuterungen zum Datensatz, die den Rezipienten erst in die Lage versetzen, die kommunizierten Resultate einzuordnen, oft sehr kurz ausfallen. Inhaltliche Tiefe ist außerdem nur möglich, wenn inhaltlich reiche Daten erhoben wurden – im Falle von Rechtsprechungszitiernetzwerken etwa Zitate in gewichteter Zählung, möglichst umfassende Metadaten zu den einzelnen Entscheidungen und idealerweise auch die Entscheidungstexte. *Praktisch* verlangt die Forderung nach mehr Tiefe also einerseits mehr Publikationen, die (sowohl aufseiten der Produzenten als auch aufseiten der Rezipienten) Zeit brauchen, und andererseits mehr reichhaltige Datensätze (d.h. für den einzelnen Forscher: mehr Investitionen in die Zusammenstellung reichhaltiger Datensätze).

### *Mehr Vernetzung*

Momentan liegt der Schwerpunkt vieler juristischer Netzwerkstudien auf der Untersuchung eines bestimmten Datensatzes, der einen eng umgrenzten Bereich betrifft. Oft wird zwar darauf hingewiesen, dass in anderen Kontexten ähnliche Resultate erzielt wurden; eine Integration der inhaltlichen Erkenntnisse juristischer Netzwerkstudien, die dazu führt, dass diese systematisch aufeinander aufbauen können, steht allerdings noch aus. Hier könnte sich die Rechtsvergleichung als (intersystemische wie intrasystemische) *Rechtsstrukturvergleichung* verdient machen. Ebenso ist der Brückenschlag zwischen Rechtstheorie und juristischer Netzwerkforschung – Stichwort „Recht als komplexes System“ – bisher unvollendet. *Inhaltlich* bedeutet die Forderung nach mehr Vernetzung daher, dass die Resultate verschiedener juristischer Netzwerkstudien systematisch miteinander

verglichen und empirische Befunde stärker mit theoretisch Erwägungen verknüpft werden sollten.

Die Methoden der juristischen Netzwerkforschung kommen häufig isoliert zum Einsatz. Ihr größtes Erkenntnispotenzial dürften sie jedoch in Kombination mit anderen Methoden entfalten. Das können zum Beispiel Methoden der Informatik sein (etwa aus den Bereichen *Machine Learning* und *Natural Language Processing*), aber auch (rechts-)soziologische oder (rechts-)ökonomische, rechtstheoretische und rechtsdogmatische Methoden. Vor allem die Integration rechtsdogmatischer und netzwerkanalytischer Vorgehensweisen könnte dazu beitragen, die juristische Netzwerkforschung in Deutschland für einen größeren rechtswissenschaftlichen Adressatenkreis interessant zu machen. *Methodisch* verlangt die Forderung nach mehr Vernetzung also, dass die Instrumente der juristischen Netzwerkforschung häufiger gemeinsam mit anderen Methoden eingesetzt werden sollten. Besonders in Deutschland bedeutet dies auch, dass netzwerkanalytische und rechtsdogmatische Methoden zusammen zur Anwendung gelangen sollten.

Wie an dieser Arbeit deutlich wird, greift die juristische Netzwerkforschung auf viele Werkzeuge zurück, die dem Rechtswissenschaftler aus seiner Ausbildung nicht bekannt sind. Ohne ein Verständnis der mathematischen und technischen Hintergründe dieser Werkzeuge ist juristische Netzwerkforschung kaum sinnvoll möglich. *Praktisch* bedeutet die Forderung nach mehr Vernetzung daher, dass mehr Rechtswissenschaftler die Gelegenheit erhalten sollten, in *Multidisziplinarität* als gleichzeitige Beherrschung mehrerer Disziplinen zu investieren. Da solche Investitionen nicht jedem möglich sein werden, lässt sich mehr Vernetzung praktisch auch dadurch erreichen, dass mehr *Interdisziplinarität* gesucht wird. Diese kann etwa darin bestehen, dass Informatiker und Statistiker mit Rechtswissenschaftlern gemeinsam Anwendungen entwickeln, die dem Juristen ohne Spezialkenntnisse die Erforschung juristischer Netzwerke ermöglichen.

\* \* \*

Welche Rolle die juristische Netzwerkforschung im Recht spielen wird, ist nach alledem offen. Jeder Jurist hat die Möglichkeit, diese Rolle mitzugestalten. Die juristische Netzwerkforschung ist für das Recht, was *A Square* für *Flatland* beschreibt:

*A Romance of Many Dimensions.*



## Appendix

## A. Struktur des Online-Appendix

Die Datenarbeit wurde in der Programmiersprache Python durchgeführt. Dabei wurde die Anaconda Distribution verwendet, deren integrierter Environment Manager es ermöglicht, in verschiedenen Environments mit verschiedenen Python-Versionen zu arbeiten. Die Netzwerkanalysen basieren größtenteils auf *networkx-1.11* für *python-3.6*; allein für die Gemeinschaftserkennung wurde primär *igraph-0.7.0* für *python-2.7* genutzt. Der Code wurde in Jupyter Notebooks verfasst, die neben Python-Code auch Texte (in Markdown) und Graphiken enthalten können.

<i>Ordner</i>	<i>Unterordner</i>	<i>Beschreibung</i>
mockdata		Dokumentation zu den in der Arbeit verwendeten fiktiven Daten
	data	Fiktive Daten und daraus abgeleitete Graphen
	notebooks	Verwendeter Code als Jupyter Notebooks
	notebooks.html supplements	Jupyter Notebooks in HTML-Darstellung Durch Quantifizierung abgeleitete Daten zur Unterstützung der Analyse
realdata		Dokumentation zu den in der Arbeit verwendeten realen Daten
	data	BVerfGE-Korpus und unmittelbar abgeleitete Daten (insbesondere Graphen)
	extraction	Ordner zur Durchführung der Zitatextraktion
	notebooks	Zur Analyse verwendeter Code als Jupyter Notebooks
	notebooks.html supplements	Jupyter Notebooks in HTML-Darstellung Durch Quantifizierung abgeleitete Daten zur Unterstützung der Analyse
searchdocs		Dokumente zu den in der Arbeit geschilderten Suchen
setup		Anaconda Environments zum Setup für die Notebooks

## B. Ergebnisse der Stichprobenüberprüfung

<i>Datei</i>	<i>Ges.</i>	<i>Gef.</i>	<i>TP</i>	<i>FP</i>	<i>FN</i>
1953-08-01 BVerfG 1 BvR 281-53 BVerfGE 3-19.txt	6	6	6	0	0
1954-02-26 BVerfG 1 BvR 397-51 BVerfGE 3-366.txt	2	1	1	0	1
1957-10-17 BVerfG 1 BvL 1-57 BVerfGE 7-155.txt	6	5	5	0	1
1958-10-14 BVerfG 1 BvR 510-52 BVerfGE 8-197.txt	1	1	1	0	0
1961-01-24 BVerfG 2 BvR 74-60 BVerfGE 12-81.txt	6	6	6	0	0
1966-03-25 BVerfG 2 BvF 1-65 BVerfGE 20-26.txt	0	0	0	0	0
1969-02-26 BVerfG 1 BvR 619-63 BVerfGE 25-256.txt	14	14	14	0	0
1973-01-30 BVerfG 2 BvH 1-72 BVerfGE 34-216.txt	9	9	9	0	0
1975-10-28 BVerfG 2 BvL 9-74 BVerfGE 40-261.txt	4	4	4	0	0
1976-02-11 BVerfG 2 BvR 849-75 BVerfGE 41-332.txt	12	12	12	0	0
1976-05-11 BVerfG 1 BvR 671-70 BVerfGE 42-143.txt	30	23	23	0	7
1979-05-22 BVerfG 1 BvL 9-75 BVerfGE 51-193.txt	20	18	18	0	2
1979-06-13 BVerfG 1 BvL 97-78 BVerfGE 51-295.txt	7	7	7	0	0
1979-07-25 BVerfG 2 BvR 878-74 BVerfGE 52-131.txt	73	69	69	0	4
1982-03-23 BVerfG 2 BvL 1-81 BVerfGE 60-162.txt	58	45	45	0	13
1982-10-19 BVerfG 2 BvF 1-81 BVerfGE 61-149.txt	19	18	18	0	1
1984-01-10 BVerfG 1 BvL 5-83 BVerfGE 66-84.txt	20	20	20	0	0
1984-02-08 BVerfG 1 BvR 580-83 BVerfGE 66-155.txt	26	26	26	0	0
1987-04-08 BVerfG 1 BvL 8-84 BVerfGE 75-40.txt	25	25	25	0	0
1990-10-17 BVerfG 1 BvR 283-85 BVerfGE 83-1.txt	27	26	26	0	1
1994-04-26 BVerfG 1 BvR 1689-88 BVerfGE 90-255.txt	12	12	12	0	0
1995-05-16 BVerfG 1 BvR 1087-91 BVerfGE 93-1.txt	60	60	60	0	0
2003-10-07 BVerfG 2 BvG 1-02 BVerfGE 109-1.txt	22	22	22	0	0
2004-03-17 BVerfG 1 BvR 1266-00 BVerfGE 110-177.txt	18	18	18	0	0
2004-06-08 BVerfG 2 BvE 1-04 BVerfGE 110-403.txt	23	19	19	0	4
2004-10-12 BVerfG 1 BvR 2130-98 BVerfGE 111-289.txt	31	30	30	0	1
2005-01-11 BVerfG 2 BvR 167-02 BVerfGE 112-164.txt	58	56	56	0	2
2005-04-19 BVerfG 1 BvR 1644-00 BVerfGE 112-332.txt	43	43	43	0	0
2009-10-13 BVerfG 2 BvL 3-05 BVerfGE 124-282.txt	22	22	22	0	0
2010-03-02 BVerfG 1 BvR 256-08 BVerfGE 125-260.txt	178	178	178	0	0
2010-07-21 BVerfG 1 BvL 11-06 BVerfGE 126-369.txt	131	131	131	0	0
2012-09-12 BVerfG 2 BvR 1824-12 BVerfGE 132-287.txt	10	10	10	0	0
2014-03-19 BVerfG 2 BvE 7-12 BVerfGE 136-1.txt	11	11	11	0	0
$\Sigma$	984	947	947	0	37

## C. Registerzeichen des Bundesverfassungsgerichts

<i>Kürzel</i>	<i>Verfahrensart</i>
BvA	Verwirkung von Grundrechten (Art. 18 GG)
BvB	Feststellung der Verfassungswidrigkeit von Parteien (Art. 21 II GG)
BvC	Beschwerden im Wahlprüfungsverfahren (Art. 41 II GG)
BvD	Anklagen gegen den Bundespräsidenten (Art. 61 GG)
BvE	Verfassungsstreitigkeiten zwischen Bundesorganen (Art. 93 I Nr. 1 GG)
BvF	Normenkontrollen auf Antrag von Verfassungsorganen (Art. 93 I Nr. 2 GG)
BvG	Verfassungsstreitigkeiten zwischen Bund und Ländern (Art. 93 I Nr. 3 und Art. 84 IV 2 GG)
BvH	Öffentlichrechtliche Streitigkeiten (Art. 93 I Nr. 4 GG)
BvJ	Richteranklagen (Art. 98 II und V GG)
BvK	Landesverfassungsstreitigkeiten kraft landesrechtlicher Zuweisung (Art. 99 GG)
BvL	Normenkontrollen auf Vorlage der Gerichte (Art. 100 I GG) (Senate und Kammern)
BvM	Völkerrechtsregel als Teil des Bundesrechts (Art. 100 II GG)
BvN	Auslegung des Grundgesetzes auf Vorlage eines Landesverfassungsgerichts (Art. 100 III GG)
BvO	Meinungsverschiedenheiten über das Fortgelten von Recht als Bundesrecht (Art. 126 GG)
BvP	Verfahren in den sonst durch Bundesgesetz zugewiesenen Fällen (Art. 93 III GG) (ab 1971)
BvQ	Einstweilige Anordnungen (§ 32 BVerfGG) und (bis 1970) sonstige Verfahren
BvR	Verfassungsbeschwerden (Art. 93 I Nr. 4a und 4b GG)
BvT	(1) Prüfung der Senatszuständigkeit (jetzt: Verw.AZ 1004/4 E), (2) Bewilligung von Armenrecht (jetzt: BvR), (3) Plenarentscheidungen (§ 16 BVerfGG, jetzt: PBvU) (bis 1960)
BvW	Meinungsverschiedenheiten, ob ein Gesetz den Voraussetzungen, des Art. 72 II, 93 I Nr. 2a GG entspricht (ab 1998)
BvX	Vorlagen des Bundesgerichtshofs nach § 36 II Untersuchungsausschussgesetz (ab 2001)
BvY	Anträge nach Art. 93 II GG (ab 2006)
PBvS	Beendigung des Amtes eines Richters des BVerfG bei Dienstunfähigkeit oder aus sonstigen Gründen (§ 105 BVerfGG)
BvU / PBvU	Plenarentscheidungen (§ 16 BVerfGG)
PBvV	Rechtsgutachten des Plenums (§ 97 BVerfGG) (bis 1952)

## D. Namen der Gemeinschaften in BVerfGE

<i>Nr.</i>	<i>Name</i>	<i>Größe</i>
0	Wahlen	135
1	Steuern	81
12	Art. 100 I GG	78
5	Berufsfreiheit	77
6	Eigentum	76
4	Rückwirkung/Art. 3 I GG	70
16	Einstweilige Anordnung	67
7	Meinungsfreiheit/Pressefreiheit/Allgemeines Persönlichkeitsrecht	66
2	Beamte	65
9	Abgeordnete	57
14	Art. 80 GG	51
8	Abgaben	45
11	Religionsfreiheit	45
10	Strafe	43
17	Rechtsschutz	41
13	Kommunen	40
18	Freie Berufe	40
23	Bestimmtheit	39
3	Datenschutz	38
32	Rechtliches Gehör (Beschwerde)	38
28	Rechtliches Gehör	37
20	Vereinigungsfreiheit	36
26	Art. 6 I GG/Art. 3 I GG	36
30	Gesetzlicher Richter	36
31	Einstweilige Anordnung	36
21	Familie	35
22	Rechtsstaat	35
19	Versorgung	33
29	Zwangsversicherung/Zwangsmitgliedschaft	33
37	Rechtswegerschöpfung/Subsidiarität	32
42	Verfahrensfragen	29
44	Rechtsfortgeltung	29
49	Rechtsfortgeltung	28
50	Gewerbliche Steuern	28

<i>Nr.</i>	<i>Name</i>	<i>Größe</i>
33	Gesetzgebungskompetenz	26
36	Kommunale Abgaben	26
40	Nachkriegszeit (Versorgung)	26
47	Wiedereinsetzung	26
53	Wahlprüfung	26
67	Einstweilige Anordnung	26
27	Hochschulen	25
35	Aufenthalt	25
43	Nichtehelichkeit	24
15	Europa	23
52	Parteienprivileg	23
38	Völkerrecht	22
41	Nachkriegszeit (Kriegsfolgen)	22
25	Verteidigung	21
45	Allgemeine Handlungsfreiheit	20
46	Staatliche Gerichte	20
48	Hochschulzugang	20
51	Hinterbliebene	20
60	Befangenheit Verfassungsrichter	20
24	Rundfunkfreiheit	19
34	Gesetzesvorbehalt/Schule	19
56	Ehefolgen	19
58	Wehrpflicht	19
59	Gerichtliche Kontrolldichte	19
63	Asyl	19
71	Verfahrensfragen	18
80	Verfahrenskosten	18
57	Zivilprozess	17
74	Art. 100 I GG (Zulässigkeit)	17
55	Enteignung	16
61	Direkte Demokratie	16
69	Armenrecht	16
39	Untersuchungsausschüsse	15
65	Präklusion	15
72	Strafbegriff	15
79	Vollstreckung/Willkürverbot	15
81	Haft	15
54	Grundrechtsfähigkeit	14
62	Familienförderung	14
66	Gesetzgebungskompetenz	14

<i>Nr.</i>	<i>Name</i>	<i>Größe</i>
68	Wiedereinsetzung (Urlaub)	14
70	Arbeitnehmerrechte	14
73	Vertragsfreiheit	14
75	Notare	14
83	Rechtliches Gehör	14
85	Art. 3 I GG	14
90	Gesetzlicher Richter	14
64	Rechtsdurchsetzung	13
95	Befangenheit Verfassungsrichter	12
97	Bodenrecht	12
76	Zwangsunterbringung/Zwangsbehandlung	11
78	Äußerungsbefugnisse Exekutive	11
82	Kündigungsschutz	11
84	Devisen	11
92	Untergang von Staaten	11
98	Lohnersatzleistungen	11
112	Antragsrücknahme bei abstrakter Normenkontrolle	11
86	Testierfreiheit	10
88	Informationsfreiheit	10
91	Rechtsverordnungen	10
93	Berlin	10
96	Auftragsverwaltung	10
106	Rechtsbehelfe Parlamentswahlen	10
107	Art. 100 I GG (Erheblichkeit)	10
111	Beschwerdevoraussetzungen Verfassungsbeschwerde	10
77	Kunsthfreiheit	9
89	Verfahren vor dem BVerfG	9
99	Wiedervereinigung	9
105	Beschwerdefrist	9
113	Ablehnung Verfassungsrichter	9
114	Einstweilige Anordnung	9
118	Verfassungskonforme Auslegung	9
121	Rechtliches Gehör/Art. 19 IV GG	9
127	Verfassungsbeschwerde bei Gesetzesänderung	9
101	Art. 3 I GG	8
108	Art. 3 II GG	8
119	Verfassungskonforme Auslegung	8
122	Staatsbürgerschaft/Auslieferung	8
94	Inkompatibilität	7
100	Anpassungsspielraum Gesetzgeber	7

<i>Nr.</i>	<i>Name</i>	<i>Größe</i>
102	Arbeitslosenhilfe	7
103	Namensrecht	7
116	Landesorganstreit	7
123	Art. 3 I GG	7
124	Art. 100 I GG (Vorlagebeschluss)	7
129	Richterausschüsse (BVerfG)	7
130	Landesrecht	7
132	Prozesskostenhilfe (BVerfG)	7
138	Auslagenerstattung	7
142	NPD-Verbotsverfahren	7
144	Zitiergebot	7
149	Zivilprozess	7
154	Rechtswegerschöpfung	7
87	Finanzausgleich	6
117	Amnestie	6
125	Durchsuchungen	6
128	Rechtsweggarantie/Rechtsschutzbedürfnis	6
131	Widerspruch gegen einstweilige Anordnung	6
133	Wiedergutmachungsrecht	6
136	Art. 3 I GG	6
140	Monopole	6
147	Einstweilige Anordnung (Freiheitsentzug)	6
148	Rechtsstaat	6
109	Terrorbekämpfung	5
110	Rentenüberleitung DDR	5
115	Privatschulen	5
120	Gestaltungsspielraum Gesetzgeber	5
135	Transsexuelle	5
139	Aktien	5
145	Mutterschutz	5
151	Besoldung	5
152	Untersuchungshaft	5
153	Prozessvertretung vor dem BVerfG	5
156	Einstweilige Anordnung (Staatshandeln)	5
158	Art. 80 I GG	5
159	Rechtswegbegriff	5
162	Sowjetzonale Urteile	5
169	Entmündigung/Pflegschaft	5
171	Art. 3 I GG	5

## Literatur

- Abbott, E. A.*: Flatland. A Romance of Many Dimensions. London: Seeley & Co., 1884.
- Acemoglu, D./Ozdaglar, A./Tahbaz-Salehi, A.*: Networks, Shocks, and Systemic Risk. In: The Oxford Handbook of the Economics of Networks. Hrsg. von *Bramoullé/Galeotti/Rogers*. Oxford: Oxford University Press, 2016, S. 569–607.
- : Systemic Risk and Stability in Financial Networks. In: *American Economic Review* 105.2 (2015), S. 564–608.
- Aggarwal, C. C.*: Data Mining. The Textbook. Cham: Springer, 2015.
- Agnarsson, G./Greenlaw, R.*: Graph Theory. Modeling, Applications, and Algorithms. Upper Saddle River: Pearson Education, 2007.
- Agnoloni, T./Pagallo, U.*: The Power Laws of the Italian Constitutional Court, and Their Relevance for Legal Scholars. In: Legal Knowledge and Information Systems. Hrsg. von *Rotolo*. Amsterdam: IOS Press, 2015, S. 1–10.
- Ahrens, W.*: Mathematische Unterhaltungen und Spiele. Leipzig: B. G. Teubner, 1901.
- Albert, R./Barabási, A.-L.*: Statistical Mechanics of Complex Networks. In: *Reviews of Modern Physics* 74 (2002), S. 47–97.
- Albert, R./Jeong, H./Barabási, A.-L.*: Error and Attack Tolerance of Complex Networks. In: *Nature* 406 (2000), S. 378–382.
- Alberts, J. E./Fry, B.*: Is Bitcoin a Security? In: *Boston University Journal of Science and Technology Law* 21.1 (2015), S. 1–21.
- Alcácer, J./Gittelman, M./Sampat, B.*: Applicant and Examiner Citations in U.S. Patents: An Overview and Analysis. In: *Research Policy* 38 (2009), S. 415–427.
- Alderson, D. L./Doyle, J. C.*: Contrasting Views of Complexity and Their Implications For Network-Centric Infrastructures. In: *IEEE Transactions on Systems, Man, and Cybernetics* 40.4 (2010), S. 839–852.
- Alexanderson, G. L.*: Euler and Königsberg’s Bridges: A Historical View. In: *Bulletin of the American Mathematical Society* 43.4 (2006), S. 567–573.
- Alschner, W./Skougarevskiy, D.*: Mapping the Universe of International Investment Agreements. In: *Journal of International Economic Law* 19 (2016), S. 561–588.
- Alstott, J./Bullmore, E./Plenz, D.*: powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. In: *PLoS ONE* 9.1 (2014), e85777-1–e85777-11.
- Altfeld, M. F./Spaeth, H. J.*: Measuring Influence on the U.S. Supreme Court. In: *Jurimetrics* 24.3 (1984), S. 236–247.
- Anderson, C. J./Wasserman, S./Crouch, B.*: A p\* Primer: Logit Models for Social Networks. In: *Social Networks* 21 (1999), S. 37–66.
- Artzy-Randrup, Y. u.a.*: Comment on ‚Network Motifs: Simple Building Blocks of Complex Networks‘ and ‚Superfamilies of Evolved and Designed Networks‘. In: *Science* 305 (2004), S. 1107c.
- Athan, T. u.a.*: OASIS LegalRuleML. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law. Hrsg. von *Vlek u.a.* New York: ACM, 2013, S. 3–12.
- Bagrow, J. P./Bollt, E. M.*: Local Method for Detecting Communities. In: *Physical Review E* 72 (2005), S. 046108-1–046108-10.

- Baird, H. S./Tombre, K.*: The Evolution of Document Image Analysis. In: Handbook of Document Image Processing and Recognition. Hrsg. von *Doermann/Tombre*. London: Springer, 2014, S. 63–71.
- Baker, W. E./Faulkner, R. R.*: The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry. In: *American Sociological Review* 58.6 (1993), S. 837–860.
- Balconi, M./Breschi, S./Lissoni, F.*: Networks of Inventors and the Role of Academia: An Exploration of Italian Patent Data. In: *Research Policy* 33.1 (2004), S. 127–145.
- Balkin, J. M.*: The Crystalline Structure of Legal Thought. In: *Rutgers Law Review* 39.1 (1986), S. 1–103.
- Barabási, A.-L.*: Network Science. Cambridge: Cambridge University Press, 2016.
- Barabási, A.-L./Albert, R.*: Emergence of Scaling in Random Networks. In: *Science* 286 (1999), S. 509–512.
- Barber, S. u.a.*: Bitter to Better – How to Make Bitcoin a Better Currency. In: *Financial Cryptography and Data Security*. Hrsg. von *Keromytis*. Berlin und Heidelberg: Springer, 2012, S. 399–414.
- Barrat, A./Barthélemy, M./Vespignani, A.*: Dynamical Processes on Complex Networks. Cambridge: Cambridge University Press, 2008.
- Barzuza, M./Curtis, Q.*: Board Interlocks and Corporate Governance. In: *Delaware Journal of Corporate Law* 39 (2015), S. 669–701.
- : Board Interlocks and Outside Directors' Protection. In: *The Journal of Legal Studies* 46.1 (2017), S. 129–160.
- Battiston, S./Caldarelli, G.*: Financial Networks. In: *Networks of Networks: The Last Frontier of Complexity*. Hrsg. von *D'Agostino/Scala*. Cham: Springer, 2014, S. 311–321.
- Battiston, S./Caldarelli, G./D'Errico, M.*: The Financial System as a Nexus of Interconnected Networks. In: *Interconnected Networks*. Hrsg. von *Garas*. Cham: Springer, 2016, S. 195–229.
- Battiston, S./Nicosia, V./Latora, V.*: Structural Measures for Multiplex Networks. In: *Physical Review E* 89 (2014), S. 032804-1–032804-14.
- Battiston, S. u.a.*: DebtRank: Too Central to Fail? Financial Networks, the FED and Systemic Risk. In: *Scientific Reports* 2 (2012), S. 541-1–541-6.
- Battiston, S. u.a.*: Liaisons Dangereuses: Increasing Connectivity, Risk Sharing, and Systemic Risk. In: *Journal of Economic Dynamics & Control* 36 (2012), S. 1121–1141.
- Baumann, N./Stiller, S.*: Network Models. In: *Network Analysis. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 341–372.
- Bavelas, A.*: Communication Patterns in Task-Oriented Groups. In: *The Journal of the Acoustical Society of America* 22.6 (1950), S. 725–730.
- Baxter, G. J. u.a.*: Cycles and Clustering in Multiplex Networks. In: *Physical Review E* 94 (2016), S. 062308-1–062308-8.
- Beauchamp, M. A.*: An Improved Index of Centrality. In: *Behavioral Science* 10.2 (1965), S. 161–163.
- Bell, E. T.*: Exponential Numbers. In: *The American Mathematical Monthly* 41.7 (1934), S. 411–419.
- Bent, J. R.*: Hidden Priors: Toward a Unifying Theory of Systemic Disparate Treatment Law. In: *Denver University Law Review* 91.4 (2014), S. 807–851.
- : P-Values, Priors, and Procedure in Antidiscrimination Law. In: *Buffalo Law Review* 63 (2015), S. 85–145.
- Berge, C.*: Graphes et Hypergraphes. Paris: Dunot, 1970.

- Berg-Kirkpatrick, T./Durrett, G./Klein, D.*: Unsupervised Transcription of Historical Documents. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (2013), S. 207–217.
- Bergmans, B.*: Visualisierungen in Rechtslehre und Rechtswissenschaft. Ein Beitrag zur Rechtsvisualisierung. Berlin: Logos, 2009.
- Berlusconi, G.*: Social Network Analysis and Crime Prevention. In: Crime Prevention in the 21st Century. Hrsg. von *Leclerc/Savona*. Cham: Springer, 2017, S. 129–141.
- Bertin, J.*: Sémiologie Graphique. 4. Aufl. Paris: EHESS, 2005.
- Beutelspacher, A.*: Lineare Algebra. 8. Aufl. Wiesbaden: Springer Spektrum, 2014.
- Bianconi, G./Barabási, A.-L.*: Competition and Multiscaling in Evolving Networks. In: Europhysics Letters 54.4 (2001), S. 436–442.
- Bianconi, G.*: Statistical Mechanics of Multiplex Networks: Entropy and Overlap. In: Physical Review E 87 (2013), S. 062806-1–062806-15.
- Bizjak, J./Lemmon, M./Whitby, R.*: Option Backdating and Board Interlocks. In: The Review of Financial Studies 22.11 (2009), S. 4821–4847.
- Black, R. C./Spriggs, J. F.*: The Citation and Depreciation of U.S. Supreme Court Precedent. In: Journal of Empirical Legal Studies 10.2 (2013), S. 325–358.
- Blumoff, T. Y.*: On the Nature of the Action-Omission Network. In: Georgia State University Law Review 24 (2008), S. 1003–1042.
- Boccaletti, S. u.a.*: Complex Networks: Structure and Dynamics. In: Physics Reports 424 (2006), S. 175–308.
- Boccaletti, S. u.a.*: The Structure and Dynamics of Multilayer Networks. In: Physics Reports 544 (2014), S. 1–122.
- Boehme-Neßler, V.*: BilderRecht. Berlin und Heidelberg: Springer, 2009.
- Bollobás, B.*: Degree Sequences of Random Graphs. In: Discrete Mathematics 33 (1981), S. 1–19.
- : Random Graphs. 2. Aufl. Cambridge: Cambridge University Press, 2001.
- Bommarito, M. J./Katz, D. M.*: A Mathematical Approach to the Study of the United States Code. In: Physica A 389 (2010), S. 4195–4200.
- Bommarito, M. J./Katz, D. M./Isaacs-See, J.*: An Empirical Survey of the Population of U.S. Tax Court Written Decisions. In: Virginia Tax Review 30 (2011), S. 523–557.
- Bommarito, M. J./Katz, D. M./Zelner, J.*: On the Stability of Community Detection Algorithms on Longitudinal Citation Data. In: Procedia Social and Behavioral Sciences (2010), S. 26–37.
- Bommarito, M. J. u.a.*: Distance Measures for Dynamic Citation Networks. In: Physica A 389 (2010), S. 4201–4208.
- Bonabeau, E.*: Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. In: PNAS 99 (suppl. 3) (2002), S. 7280–7287.
- Bonacich, P.*: A Technique for Analyzing Overlapping Memberships. In: Sociological Methodology 4 (1972), S. 176–185.
- : Factoring and Weighting Approaches to Status Scores and Clique Identification. In: Journal of Mathematical Sociology 2 (1972), S. 113–120.
- : Power and Centrality: A Family of Measures. In: American Journal of Sociology 92.5 (1987), S. 1170–1182.
- Bonacich, P./Holdren, A. C./Johnston, M.*: Hyper-Edges and Multidimensional Centrality. In: Social Networks 26 (2004), S. 189–203.
- Borgatti, S. P.*: Centrality and Network Flow. In: Social Networks 27 (2005), S. 55–71.
- Borgatti, S. P./Everett, M. G.*: A Graph-Theoretic Perspective on Centrality. In: Social Networks 28 (2006), S. 466–484.
- Börner, K.*: Atlas of Knowledge. Anyone Can Map. Cambridge: MIT Press, 2015.

- Börner, K.*: Atlas of Science. Visualizing What We Know. Cambridge: MIT Press, 2010.
- Bormann, L./Daniel, H.-D.*: What do Citation Counts Measure? A Review of Studies on Citing Behavior. In: *Journal of Documentation* 64.1 (2008), S. 45–80.
- Boulet, R./Barros-Platiau, A. F./Mazzeza, P.*: 35 Years of Multilateral Environmental Agreements Ratifications: a Network Analysis. In: *Artificial Intelligence and Law* (2016), S. 133–148.
- Boulet, R./Mazzeza, P./Bourcier, D.*: A Network Approach to the French System of Legal Codes – Part I: Analysis of a Dense Network. In: *Artificial Intelligence and Law* 19 (2011), S. 333–355.
- : Network Analysis of the French Environmental Code. In: *AI Approaches to the Complexity of Legal Systems*. Hrsg. von *Casanovas u.a.* Berlin und Heidelberg: Springer, 2010, S. 39–53.
- Bourcier, D./Mazzeza, P.*: Codification, Law Article and Graphs. In: *Legal Knowledge and Information Systems*. Hrsg. von *Lodder/Mommers*. Amsterdam: IOS Press, 2007, S. 29–38.
- : Toward Measures of Complexity in Legal Systems. In: *ICAIL Proceedings* (2007), S. 211–215.
- Bramoullé, Y./Galeotti, A./Rogers, B. W.*, Hrsg.: *The Oxford Handbook of the Economics of Networks*. Oxford: Oxford University Press, 2016.
- Brandeis, L. D.*: *Other People's Money. And How the Bankers Use It*. New York: Frederick A. Stokes Company, 1914.
- Brandes, U.*: Graphentheorie. In: *Handbuch Netzwerkforschung*. Hrsg. von *Stegbauer/Häußling*. Wiesbaden: VS Verlag, 2010, S. 345–353.
- Brandes, U./Erlebach, T.*: Fundamentals. In: *Network Analysis. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 7–15.
- Brandes, U./Kosub, S./Nick, B.*: Was messen Zentralitätsindizes? In: *Die Integration von Theorie und Methode in der Netzwerkforschung*. Hrsg. von *Hennig/Stegbauer*. Wiesbaden: Springer VS, 2012, S. 33–52.
- Brandes, U. u.a.*: On Modularity Clustering. In: *IEEE Transactions on Knowledge and Data Engineering* 20.2 (2008), S. 172–188.
- Brauer, F.*: Compartmental Models in Epidemiology. In: *Mathematical Epidemiology*. Hrsg. von *Brauer/van den Driessche/Wu*. Berlin und Heidelberg: Springer, 2008, S. 19–79.
- Braun, J.*: *Leitbilder im Recht*. Tübingen: Mohr Siebeck, 2015.
- Brewer, C. A.*: Color Use Guidelines for Mapping and Visualization. In: *Visualization in Modern Cartography*. Hrsg. von *MacEachren/Taylor*. Oxford: Elsevier, 1994, S. 123–147.
- Bright, D. A./Hughes, C. E./Chalmers, J.*: Illuminating Dark Networks: a Social Network Analysis of an Australian Drug Trafficking Syndicate. In: *Crime, Law & Social Change* 57.2 (2012), S. 151–176.
- Brin, S./Page, L.*: The Anatomy of a Large-Scale Hypertextual Web Search Engine. In: *Computer Networks and ISDN Systems* 30 (1998), S. 107–117.
- Brinkmeier, M./Schank, T.*: Network Statistics. In: *Network Analysis. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 293–317.
- Broder, A. u.a.*: Graph Structure in the Web. In: *Computer Networks* 33 (2000), S. 309–320.
- Brown, J. L.*: The Spread of Aggressive Corporate Tax Reporting. A Detailed Examination of the Corporate-Owned Life Insurance Shelter. In: *The Accounting Review* 86.1 (2011), S. 23–57.
- Bruck, P. u.a.*: Recognition of Emerging Technology Trends: Class-Selective Study of Citations in the U.S. Patent Citation Network. In: *Scientometrics* 107.3 (2016), S. 1465–1475.
- Brughmans, T./Collar, A./Coward, F.*, Hrsg.: *The Connected Past. Challenges to Network Studies in Archaeology and History*. Oxford: Oxford University Press, 2016.
- Buch-Hansen, H.*: Interlocking Directorates and Collusion: An Empirical Analysis. In: *International Sociology* 29.3 (2014), S. 249–267.

- Bumke, C.*: Rechtsdogmatik. Eine Disziplin und ihre Arbeitsweise. Zugleich eine Studie über das rechtsdogmatische Arbeiten Friedrich Carl von Savignys. Tübingen: Mohr Siebeck, 2017.
- Burt, R. S.*: Models of Network Structure. In: *Annual Review of Sociology* 6 (1980), S. 79–141.
- : *Structural Holes. The Social Structure of Competition*. Cambridge: Harvard University Press, 1995.
- Burton, R.*: *The Anatomy of Melancholy*. Oxford: Henry Cripps, 1638 [1621].
- Butts, C. T.*: Network Inference, Error, and Informant (In)Accuracy: a Bayesian Approach. In: *Social Networks* 25 (2003), S. 103–140.
- : Revisiting the Foundations of Network Analysis. In: *Science* 325 (2009), S. 414–416.
- Cabrales, A./Gale, D./Gottardi, P.*: Financial Contagion in Networks. In: *The Oxford Handbook of the Economics of Networks*. Hrsg. von *Bramoullé/Galeotti/Rogers*. Oxford: Oxford University Press, 2016, S. 543–568.
- Cabrelli, D./Siems, M. M.*: Convergence, Legal Origins, and Transplants in Comparative Corporate Law: A Case-Based and Quantitative Analysis. In: *The American Journal of Comparative Law* 63 (2015), S. 109–153.
- Caimo, A./Friel, N.*: Bayesian Inference for Exponential Random Graph Models. In: *Social Networks* 33 (2011), S. 41–55.
- Caldarelli, G. u.a.*: Scale-Free Networks from Varying Vertex Intrinsic Fitness. In: *Physical Review Letters* (2002), S. 258702-1–258702-4.
- Caldeira, G. A.*: Legal Precedent: Structures of Communication between State Supreme Courts. In: *Social Networks* 10 (1988), S. 29–55.
- : The Transmission of Legal Precedent: A Study of State Supreme Courts. In: *The American Political Science Review* 79.1 (1985), S. 178–194.
- Calderoni, F.*: The Structure of Drug Trafficking Mafias: the 'Ndrangheta and Cocaine. In: *Crime, Law & Social Change* 58 (2012), S. 321–349.
- Callaway, D. S. u.a.*: Network Robustness and Fragility: Percolation on Random Graphs. In: *Physical Review Letters* 85.25 (2000), S. 5468–5471.
- Campgiotto, R./Guillaume, J.-L./Saïfi, M.*: The Power of Consensus: Random Graphs Have No Communities. In: *IEEE/ACM ASONAM* (2013), S. 272–276.
- Cao, H./Natarajan, P.*: Machine-Printed Character Recognition. In: *Handbook of Document Image Processing and Recognition*. Hrsg. von *Doermann/Tombre*. London: Springer, 2014, S. 331–358.
- Carley, K. M.*: Destabilization of Covert Networks. In: *Computational & Mathematical Organization Theory* 12 (2006), S. 51–66.
- Carley, K. M./Lee, J.-S./Krackhardt, D.*: Destabilizing Networks. In: *Connections* 24.3 (2002), S. 79–92.
- Carlson, J. M./Doyle, J.*: Complexity and Robustness. In: *PNAS* 99 (2002), S. 2538–2545.
- Carlson, J. M./Doyle, J. C.*: Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems. In: *Physical Review E* 60.2 (1999), S. 1412–1427.
- Carrington, P. J.*: Crime and Social Network Analysis. In: *The SAGE Handbook of Social Network Analysis*. Hrsg. von *Scott/Carrington*. London: SAGE Publications, 2011, S. 236–255.
- Carroll, W. K./Sapinski, J. P.*: Corporate Elites and Interorganizational Networks. In: *The SAGE Handbook of Social Network Analysis*. Hrsg. von *Scott/Carrington*. London: SAGE Publications, 2011, S. 180–195.
- Chandler, S. J.*: The Network Structure of Supreme Court Jurisprudence. In: *The Mathematica Journal* 10.3 (2007), S. 501–526.
- Charlotin, D.*: The Place of Investment Awards and WTO Decisions in International Law: A Citation Analysis. In: *Journal of International Economic Law* 20 (2017), S. 279–299.

- Chen, C.-C. u.a.: BibPro: A Citation Parser Based on Sequence Alignment. In: IEEE Transactions on Knowledge and Data Engineering 24.2 (2012), S. 236–250.
- Choi, J./Hwang, Y.-S.: Patent Keyword Network Analysis for Improving Technology Development Efficiency. In: Technological Forecasting and Social Change 83 (2014), S. 170–182.
- Christensen, M. L./Olsen, H. P./Tarissan, F.: Identification of Case Content with Quantitative Network Analysis: An Example from the ECtHR. In: Legal Knowledge and Information Systems. Hrsg. von Bex/Villata. Amsterdam: IOS Press, 2016, S. 53–62.
- Clark, T. S./Lauderdale, B. E./Katz, J. N.: The Genealogy of Law. In: Political Analysis 20.3 (2012), S. 329–350.
- Clauset, A./Moore, C./Newman, M. E. J.: Structural Inference of Hierarchies in Networks. In: Statistical Network Analysis: Models, Issues, and New Directions. Hrsg. von Airoidi u.a. Berlin und Heidelberg: Springer, 2007, S. 1–13.
- Clauset, A./Shalizi, C. R./Newman, M. E. J.: Power-Law Distributions in Empirical Data. In: SIAM Review 51.4 (2009), S. 661–703.
- Cleveland, W. S./Harris, C. S./McGill, R.: Experiments on Quantitative Judgments of Graphs and Maps. In: The Bell System Technical Journal 62.6 (1983), S. 1659–1674.
- Cleveland, W. S.: The Elements of Graphing Data. Revised Edition. Summit: Hobart Press, 1994.
- : Visualizing Data. Summit: Hobart Press, 1993.
- Clough, J. R./Evans, T. S.: What is the Dimension of Citation Space? In: Physica A 448 (2016), S. 235–247.
- Clough, J. R. u.a.: Transitive Reduction of Citation Networks. In: Journal of Complex Networks 3.2 (2015), S. 189–203.
- Cohen, R./Havlin, S.: Complex Networks. Structure, Robustness and Function. Cambridge: Cambridge University Press, 2010.
- Cohen, R./Havlin, S./Ben-Avraham, D.: Efficient Immunization Strategies for Computer Networks and Populations. In: Physical Review Letters 91.24 (2003), S. 247901-1–247901-4.
- Cormen, T. H. u.a.: Introduction to Algorithms. 3. Aufl. Cambridge: MIT Press, 2009.
- Coupette, C./Fleckner, A. M.: Quantitative Rechtswissenschaft. In: JZ 73.8 (2018), S. 379–389.
- Cross, F. B./Smith, T. A./Tomarchio, A.: The Reagan Revolution in the Network of Law. In: Emory Law Journal 57 (2008), S. 1227–1257.
- Cross, F. B. u.a.: Citations in the U.S. Supreme Court: An Empirical Study of their Use and Significance. In: University of Illinois Law Review 2 (2010), S. 489–575.
- Cunningham, L. A.: Language, Deals, and Standards: The Future of XML Contracts. In: Washington University Law Review 84 (2006), S. 313–373.
- Davis, G. F.: Agents Without Principles? The Spread of the Poison Pill Through the Intercorporate Network. In: Administrative Science Quarterly 36.4 (1991), S. 583–613.
- Davis, G. F./Mizruchi, M. S.: The Money Center Cannot Hold: Commercial Banks in the U.S. System of Corporate Governance. In: Administrative Science Quarterly 44.2 (1999), S. 215–239.
- Davis, R. H.: Social Network Analysis: An Aid in Conspiracy Investigations. In: FBI Law Enforcement Bulletin 50 (1981), S. 11–19.
- De Domenico, M./Porter, M. A./Arenas, A.: MuxViz: a Tool for Multilayer Analysis and Visualization of Networks. In: Journal of Complex Networks 3.2 (2015), S. 159–176.
- De Domenico, M. u.a.: Mathematical Formulation of Multilayer Networks. In: Physical Review X 3 (2013), S. 041022-1–041022-15.
- De Domenico, M. u.a.: Navigability of Interconnected Networks under Random Failures. In: PNAS 111.23 (2014), S. 8351–8356.

- Derényi, I./Palla, G./Vicsek, T.*: Clique Percolation in Random Networks. In: *Physical Review Letters* 94 (2005), S. 160202-1–160202-4.
- Derlén, M./Lindholm, J.*: Characteristics of Precedent: The Case Law of the European Court of Justice in Three Dimensions. In: *German Law Journal* 16.5 (2015), S. 1073–1098.
- : Goodbye van Gend en Loos, Hello Bosman? Using Network Analysis to Measure the Importance of Individual CJEU Judgements. In: *European Law Journal* 20.5 (2014), S. 667–687.
- : Is it Good Law? Network Analysis and the CJEU’s Internal Market Jurisprudence. In: *Journal of International Economic Law* 20 (2017), S. 257–277.
- : Peek-A-Boo, It’s a Case Law System! Comparing the European Court of Justice and the United States Supreme Court from a Network Perspective. In: *German Law Journal* 18.3 (2017), S. 647–686.
- : The Court of Justice and the Ankara Agreement: Exploring the Empirical Approach. In: *Europarättslig Tidskrift* 3 (2012), S. 462–481.
- Derlén, M. u.a.*: Coherence Out of Chaos: Mapping European Union Law by Running Randomly Through the Maze of CJEU Case Law. In: *Europarättslig Tidskrift* (2013), S. 517–535.
- Desmarais, B. A. u.a.*: Measuring Legislative Collaboration: The Senate Press Events Network. In: *Social Networks* 40 (2015), S. 43–54.
- Dezső, Z./Barabási, A.-L.*: Halting Viruses in Scale-Free Networks. In: *Physical Review E* 65 (2002), 055103(R)-1–055103(R)-4.
- Di Lorenzo, V.*: Legislative Chaos: An Exploratory Study. In: *Yale Law & Policy Review* 12.2 (1994), S. 425–485.
- : Legislative Heart and Phase Transitions: An Exploratory Study of Congress and Minority Interests. In: *William & Mary Law Review* 38.5 (1997), S. 1729–1815.
- Dibadj, R.*: Networks of Fairness Review in Corporate Law. In: *San Diego Law Review* 45 (2008), S. 1–32.
- : Networks of Heightened Scrutiny in Corporate Law. In: *San Diego Law Review* 46.1 (2009), S. 1–26.
- Dooley, P. C.*: The Interlocking Directorate. In: *American Economic Review* 59.3 (1969), S. 314–323.
- Dorogovtsev, S. N./Mendes, J. F. F.*: Evolution of Networks. In: *Advances in Physics* 51.4 (2002), S. 1079–1187.
- Dorogovtsev, S. N./Mendes, J. F. F./Samukhin, A. N.*: Structure of Growing Networks with Preferential Linking. In: *Physical Review Letters* 85.21 (2000), S. 4633–4636.
- Doyle, J. C. u.a.*: The „Robust Yet Fragile“ Nature of the Internet. In: *PNAS* 102.41 (2005), S. 14497–14502.
- Duijn, P. A. C./Kashirin, V./Sloot, P. M. A.*: The Relative Ineffectiveness of Criminal Network Disruption. In: *Scientific Reports* 4 (2014), S. 4238-1–4238-15.
- Düring, M./Eumann, U.*: Historische Netzwerkforschung. Ein neuer Ansatz in den Geschichtswissenschaften. In: *Geschichte und Gesellschaft* 39.3 (2013), S. 369–390.
- Easley, D./Kleinberg, J.*: *Networks, Crowds, and Markets. Reasoning about a Highly Connected World.* New York: Cambridge University Press, 2010.
- Ehrlich, E.*: *Die Rechtsfähigkeit.* Berlin: Puttkammer & Mühlbrecht, 1909.
- Elliott, M./Golub, B./Jackson, M. O.*: Financial Networks and Contagion. In: *American Economic Review* 104.10 (2014), S. 3115–3153.
- Engst, B. G. u.a.*: Zum Einfluss der Parteinähe auf das Abstimmungsverhalten der Bundesverfassungsrichter – eine quantitative Untersuchung. In: *JZ* 72.17 (2017), S. 816–826.
- Eom, Y.-H./Fortunato, S.*: Characterizing and Modeling Citation Dynamics. In: *PLoS ONE* 6.9 (2011), e24926-1–e24926-7.

- Epstein, J. M.*: Agent-Based Computational Models and Generative Social Science. In: *Complexity* 4.5 (1999), S. 41–60.
- Érdi, P. u.a.*: Prediction of Emerging Technologies Based on Analysis of the US Patent Citation Network. In: *Scientometrics* 95.1 (2013), S. 225–242.
- Erdős, P./Rényi, A.*: On Random Graphs I. In: *Publicationes Mathematicae* 6 (1959), S. 290–297.
- Eser, A./Sternberg-Lieben, D.*: StGB § 223. In: *Schönke/Schröder Strafgesetzbuch*. Hrsg. von Eser. 29. Aufl. München: C.H. Beck, 2014.
- Estrada, E.*: *The Structure of Complex Networks. Theory and Applications*. Oxford: Oxford University Press, 2012.
- Estrada, E./Knight, P. A.*: *A First Course in Network Theory*. Oxford: Oxford University Press, 2015.
- Estrada, E. u.a.*, Hrsg.: *Network Science. Complexity in Nature and Technology*. London: Springer, 2010.
- Euler, L.*: *Solutio problematis ad geometriam situs pertinentis*. In: *Commentarii Academiae Scientiarum Petropolitanae* 8 (1741), S. 128–140.
- Everett, M. G.*: Role Similarity and Complexity in Social Networks. In: *Social Networks* 7 (1985), S. 353–359.
- Everett, M. G./Borgatti, S.*: Calculating Role Similarities: An Algorithm that Helps Determine the Orbits of a Graph. In: *Social Networks* 10 (1988), S. 77–91.
- Everton, S. F.*: *Disrupting Dark Networks*. Cambridge: Cambridge University Press, 2012.
- Fahrmeir, L. u.a.*: *Statistik. Der Weg zur Datenanalyse*. 8. Aufl. Berlin: Springer Spektrum, 2016.
- Falce, V.*: Interlocking Directorates: An Italian Antitrust Dilemma. In: *Journal of Competition Law & Economics* 9.2 (2013), S. 457–472.
- Feldman, R./Sanger, J.*: *The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data*. Cambridge: Cambridge University Press, 2007.
- Feldman, Y./Nadler, J.*: The Law and Norms of File Sharing. In: *San Diego Law Review* 43 (2006), S. 577–618.
- Fennema, M./Schiff, H.*: Analysing Interlocking Directorates: Theory and Methods. In: *Social Networks* 1 (1978), S. 297–332.
- Finkelstein, M. O./Levin, B.*: *Statistics for Lawyers*. 3. Aufl. Springer: New York, 2015.
- Fischer, E. N.*: Serving More than One Master: A Social Network Analysis of Section 8 of the Clayton Act. In: *The Journal of Corporation Law* 41.1 (2015), S. 313–341.
- Fiss, P. C./Kennedy, M. T./Davis, G. F.*: How Golden Parachutes Unfolded: Diffusion and Variation of a Controversial Practice. In: *Organization Science* 23.4 (2012), S. 1077–1099.
- Fohlin, C.*: The Rise of Interlocking Directorates in Imperial Germany. In: *The Economic History Review* 52.2 (1999), S. 307–333.
- Fontana, R./Nuvolari, A./Verspagen, B.*: Mapping Technological Trajectories as Patent Citation Networks. An Application to Data Communication Standards. In: *Economics of Innovation and New Technology* 18.4 (2009), S. 311–336.
- Fortunato, S.*: Community Detection in Graphs. In: *Physics Reports* 486 (2010), S. 75–174.
- Foster, J. G. u.a.*: Edge Direction and the Structure of Networks. In: *PNAS* 107.24 (2010), S. 10815–10820.
- Fowler, J. H.*: Connecting the Congress: A Study of Legislative Cosponsorship Networks. In: *Political Analysis* 14.4 (2006), S. 454–465.
- : Legislative Cosponsorship Networks in the US House and Senate. In: *Social Networks* 28 (2006), S. 454–465.
- Fowler, J. H./Jeon, S.*: The Authority of Supreme Court Precedent. In: *Social Networks* 30.1 (2008), S. 16–30.

- Fowler, J. H. u.a.*: Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. In: *Political Analysis* 15.3 (2007), S. 324–346.
- Frankenreiter, J.*: Network Analysis and the Use of Precedent in the Case Law of the CJEU - A Reply to Derlén and Lindholm. In: *German Law Journal* 18.3 (2017), S. 687–693.
- Freeman, L. C.*: A Set of Measures of Centrality Based on Betweenness. In: *Sociometry* 40.1 (1977), S. 35–41.
- : Centrality in Social Networks. Conceptual Clarification. In: *Social Networks* 1 (1979), S. 215–239.
- : *The Development of Social Network Analysis*. Vancouver: Booksurge Publishing, 2004.
- : The Development of Social Network Analysis – With an Emphasis on Recent Events. In: *The SAGE Handbook of Social Network Analysis*. Hrsg. von *Scott/Carrington*. London: SAGE Publications, 2011, S. 26–39.
- Frölich, K.*: Probleme der Rechtskartographie. In: *Vierteljahresschrift für Sozial- und Wirtschaftsgeschichte* 27.1 (1934), S. 40–64.
- Fruchterman, T. M. J./Reingold, E. M.*: Graph Drawing by Force-Directed Placement. In: *Software-Practice and Experience* 21.11 (1991), S. 1129–1164.
- Gaertler, M.*: Clustering. In: *Network Analysis. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 178–215.
- Gai, P./Haldane, A./Kapadia, S.*: Complexity, Concentration and Contagion. In: *Journal of Monetary Economics* 58 (2011), S. 453–470.
- Gai, P./Kapadia, S.*: Contagion in Financial Networks. In: *Proceedings of the Royal Society A* (2010), S. 1–23.
- Galbiati, M./Soramäki, K.*: Clearing Networks. In: *Journal of Economic Behavior & Organization* 83 (2012), S. 609–626.
- Garfield, E.*: Breaking the Subject Index Barrier — A Citation Index for Chemical Patents. In: *Journal of the Patent Office Society* 39.8 (1957), S. 583–595.
- Garoupa, N./Ogus, A.*: A Strategic Interpretation of Legal Transplants. In: *The Journal of Legal Studies* 35.2 (2006), S. 339–363.
- Gelter, M./Siems, M. M.*: Citations to Foreign Courts - Illegitimate and Superfluous, or Unavoidable - Evidence from Europe. In: *The American Journal of Comparative Law* 62.1 (2014), S. 35–86.
- : Language, Legal Origins, and Culture before the Courts: Cross-Citations between Supreme Courts in Europe. In: *Supreme Court Economic Review* 21 (2013), S. 215–270.
- : Networks, Dialogue or One-Way Traffic: An Empirical Analysis of Cross-Citations between Ten of Europe’s Highest Courts. In: *Utrecht Law Review* 8.2 (2005), S. 88–99.
- Gerber, B. M.*: Enabling Interlock Benefits While Preventing Anticompetitive Harm: Toward an Optimal Definition of Competitors Under Section 8 of the Clayton Act. In: *Yale Journal on Regulation* 24.1 (2007), S. 107–137.
- Geu, T. E.*: The Tao of Jurisprudence: Chaos, Brain Science, Synchronicity, and the Law. In: *Tennessee Law Review* 61 (1994), S. 933–990.
- Gewirtzman, D.*: Lower Court Constitutionalism: Circuit Court Discretion in a Complex Adaptive System. In: *American University Law Review* 61 (2012), S. 457–522.
- Gilbert, E. N.*: Random Graphs. In: *The Annals of Mathematical Statistics* 30.4 (1959), S. 1141–1144.
- Girvan, M./Newman, M. E. J.*: Community Structure in Social and Biological Networks. In: *PNAS* 99.12 (2002), S. 7821–7826.
- Gómez, S. u.a.*: Diffusion Dynamics on Multiplex Networks. In: *Physical Review Letters* 110 (2013), S. 028701-1–028701-5.

- Goodman, L. A.*: Snowball Sampling. In: *The Annals of Mathematical Statistics* 32.1 (1961), S. 148–170.
- Governatori, G.*: Representing Business Contracts in RuleML. In: *International Journal of Cooperative Information Systems* 14 (2005), S. 181–216.
- Granovetter, M. S.*: The Strength of Weak Ties. In: *American Journal of Sociology* 78.6 (1973), S. 1360–1380.
- Guclu, H. u.a.*: State-Level Legal Preparedness for Nuclear and Radiological Emergencies in the U.S.: A Network Analysis of State Laws and Regulations. In: *Public Health Reports* 129 (2014), S. 154–165.
- Guimerà, R./Amaral, L. A. N.*: Cartography of Complex Networks: Modules and Universal Roles. In: *Journal of Statistical Mechanics: Theory and Experiment* (2005), P02001-1–P02001-13.
- Guimerà, R./Sales-Pardo, M.*: Missing and Spurious Interactions and the Reconstruction of Complex Networks. In: *PNAS* 106.52 (2009), S. 22073–22078.
- Gulati, R./Westphal, J. D.*: Cooperative or Controlling? The Effects of CEO-Board Relations and the Content of Interlocks on the Formation of Joint Ventures. In: *Administrative Science Quarterly* 44 (1999), S. 473–506.
- Haldane, A. G./May, R. M.*: Systemic Risk in Banking Ecosystems. In: *Nature* 469 (2011), S. 351–355.
- Hall, B. H./Jaffe, A. B./Trajtenberg, M.*: The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. Working Paper No. 8498. National Bureau of Economic Research, 2001.
- Hamann, H.*: Die Fußnote, das unbekannte Wesen – Potential und Grenzen juristischer Zitationsanalyse. In: *RW* 4 (2014), S. 501–534.
- Harary, F.*: *Graph Theory*. Reading: Addison-Wesley, 1969.
- Hart, H. C.*: Re: „Citation System for Patent Office“. In: *Journal of the Patent Office Society* 31.9 (1949), S. 714.
- Hastings, M. B.*: Community Detection as an Inference Problem. In: *Physical Review E* 74 (2006), S. 035102-1–035102-4.
- Haunschild, P. R.*: Interorganizational Imitation: The Impact of Interlocks on Corporate Acquisition Activity. In: *Administrative Science Quarterly* 38.4 (1993), S. 564–592.
- Hautsch, N./Schaumburg, J./Schienle, M.*: Financial Network Systemic Risk Contributions. In: *Review of Finance* 19 (2015), S. 685–738.
- Haynie, D. L./Doogan, N. J./Soller, B.*: Gender, Friendship Networks, and Delinquency: A Dynamic Network Approach. In: *Criminology* 52.4 (2014), S. 688–722.
- Heath, A. u.a.*: CCPs and Network Stability in OTC Derivatives Markets. In: *Journal of Financial Stability* 27 (2016), S. 217–233.
- Heer, J./Agrawala, M.*: Multi-Scale Banking to 45°. In: *IEEE Transactions on Visualization and Computer Graphics* 12.5 (2006), S. 701–708.
- Heinz, J. P.*: Lawyers’ Professional and Political Networks Compared: Core and Periphery. In: *Arizona Law Review* 53 (2011), S. 455–492.
- Heinz, J. P./Laumann, E. O.*: The Changing Character of Lawyers’ Work: Chicago in 1975 and 1995. In: *Law & Society Review* 32.4 (1998), S. 751–775.
- Heinz, J. P./Manikas, P. M.*: Networks among Elites in a Local Criminal Justice System. In: *Law & Society Review* 26.4 (1992), S. 831–862.
- Heinz, J. P./Southworth, A./Paik, A.*: Lawyers for Conservative Causes: Clients, Ideology, and Social Distance. In: *Law & Society Review* 37.1 (2003), S. 5–50.
- Heinz, J. P. u.a.*: Inner Circles or Hollow Cores? Elite Networks in National Policy Systems. In: *The Journal of Politics* 52.2 (1990), S. 356–390.

- Heinz, J. P. u.a.: The Constituencies of Elite Urban Lawyers. In: *Law & Society Review* 31.3 (1997), S. 441–472.
- Heinz, J. P. u.a.: *The Hollow Core. Private Interests in National Policy Making*. Cambridge: Harvard University Press, 1993.
- Heinz, J. P. u.a.: *Urban Lawyers*. Chicago: The University of Chicago Press, 2005.
- Heinz, J./Laumann, E.: *Chicago Lawyers. The Social Structure of the Bar*. Revised Edition. St. Evanston: Northwestern University Press, 1994.
- Heinze, T.: Die Struktur der Personalverflechtung großer deutscher Aktiengesellschaften zwischen 1989 und 2001. In: *Zeitschrift für Soziologie* 31.5 (2002), S. 391–410.
- Helbing, D., Hrsg.: *Social Self-Organization. Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Berlin und Heidelberg: Springer, 2012.
- Holland, P. W./Laskey, K. B./Leinhardt, S.: Stochastic Blockmodels: First Steps. In: *Social Networks* 5 (1983), S. 109–137.
- Holland, P. W./Leinhardt, S.: A Method for Detecting Structure in Sociometric Data. In: *American Journal of Sociology* 76.3 (1970), S. 492–513.
- Holme, P.: Efficient Local Strategies for Vaccination and Network Attack. In: *Europhysics Letters* 68.6 (2004), S. 908–914.
- Holme, P./Saramäki, J.: Temporal Networks. In: *Physics Reports* 519 (2012), S. 97–125.
- Holten, D./van Wijk, J. J.: A User Study on Visualizing Directed Edges in Graphs. In: *Proceedings of SIGCHI* (2009), S. 2299–2308.
- Hood, W. W./Wilson, C. S.: The Literature of Bibliometrics, Scientometrics, and Informetrics. In: *Scientometrics* 52.2 (2001), S. 291–314.
- Hook, P. A.: The Aggregate Harmony Metric and a Statistical and Visual Contextualization of the Rehnquist Court: 50 Years of Data. In: *Constitutional Commentary* 24 (2007), S. 221–264.
- Hopkins, B./Wilson, R.: The Truth about Königsberg. In: *The College Mathematics Journal* 35.3 (2004), S. 198–207.
- Hric, D./Darst, R. K./Fortunato, S.: Community Detection in Networks: Structural Communities Versus Ground Truth. In: *Physical Review E* 90 (2014), S. 062805-1–062805-19.
- Hu, Y./Lei, S.: Visualizing Large Graphs. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 7 (2015), S. 107–136.
- Iori, G. u.a.: A Network Analysis of the Italian Overnight Money Market. In: *Journal of Economic Dynamics & Control* 32 (2008), S. 259–278.
- Itzkowitz, S./Alon, U.: Subgraphs and Network Motifs in Geometric Networks. In: *Physical Review E* 71 (2005), S. 026117-1–026117-9.
- Jackson, M. O.: *Social and Economic Networks*. Princeton: Princeton University Press, 2008.
- Jaffe, A. B./de Rassenfosse, G.: Patent Citation Data in Social Science Research: Overview and Best Practices. In: *Journal of the Association for Information Science and Technology* 68.6 (2017), S. 1360–1374.
- Jänich, K.: *Lineare Algebra*. 11. Aufl. Heidelberg: Springer, 2013.
- Jeidels, O.: *Das Verhältnis der deutschen Großbanken zur Industrie mit besonderer Berücksichtigung der Eisenindustrie*. Leipzig: Duncker & Humblot, 1905.
- Jones, G. T.: Dynamical Jurisprudence: Law as a Complex System. In: *Georgia State University Law Review* 24.4 (2008), S. 873–883.
- Jurafsky, D./Martin, J. H.: *Speech and Language Processing*. 2. Aufl. Upper Saddle River: Pearson Education, 2009.
- Kamada, T./Kawai, S.: An Algorithm for Drawing General Undirected Graphs. In: *Information Processing Letters* 31 (1989), S. 7–15.
- Karp, R. M.: Reducibility Among Combinatorial Problems. In: *Complexity of Computer Computations*. Hrsg. von Miller/Thatcher. New York: Plenum Press, 1972, S. 85–103.

- Karrer, B./Newman, M. E. J.*: Random Graph Models for Directed Acyclic Networks. In: *Physical Review E* 80 (2009), S. 046110-1–046110-12.
- Katz, D. M.*: Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry. In: *Emory Law Journal* 62 (2013), S. 909–966.
- Katz, D. M./Bommarito, M. J.*: Measuring the Complexity of the Law: the United States Code. In: *Artificial Intelligence and Law* 22 (2014), S. 337–374.
- Katz, D. M./Stafford, D. K.*: Hustle and Flow: A Social Network Analysis of the American Federal Judiciary. In: *Ohio State Law Journal* 71 (2010), S. 457–509.
- Katz, D. M./Stafford, D. K./Provins, E.*: Social Architecture, Judicial Peer Effects and the ‘Evolution’ of the Law: Toward a Positive Theory of Judicial Social Structure. In: *Georgia State University Law Review* 24.4 (2008), S. 977–1001.
- Katz, D. M. u.a.*: Reproduction of Hierarchy? A Social Network Analysis of the American Law Professoriate. In: *Journal of Legal Education* 61.1 (2011), S. 76–103.
- Katz, L.*: A New Status Index Derived From Sociometric Analysis. In: *Psychometrika* 18.1 (1953), S. 39–43.
- Kauermann, G./Küchenhoff, H.*: *Stichproben. Methoden und praktische Umsetzung* in R. Berlin: Springer, 2011.
- Keeling, M. J./Eames, K. T. D.*: Networks and Epidemic Models. In: *Journal of the Royal Society Interface* 2 (2005), S. 295–307.
- Kenett, D. Y./Havlin, S.*: Network Science: A Useful Tool in Economics and Finance. In: *Mind & Society* 14.2 (2015), S. 155–167.
- Kessler, M. M.*: Bibliographic Coupling Between Scientific Papers. In: *American Documentation* 14.1 (1963), S. 10–25.
- Kim, E./Cho, Y./Kim, W.*: Dynamic Patterns of Technological Convergence in Printed Electronics Technologies: Patent Citation Network. In: *Scientometrics* 98.2 (2014), S. 975–998.
- Kim, H./Song, J.*: Social Network Analysis of Patent Infringement Lawsuits. In: *Technological Forecasting and Social Change* 80.5 (2013), S. 944–955.
- Kim, R. E.*: The Emergent Network Structure of the Multilateral Environmental Agreement System. In: *Global Environmental Change* 23.5 (2013), S. 980–991.
- Kiss, I. Z./Miller, J. C./Simon, P. L.*: *Mathematics of Epidemics on Networks. From Exact to Approximate Models*. Cham: Springer, 2017.
- Kivelä, M. u.a.*: Multilayer Networks. In: *Journal of Complex Networks* 2 (2014), S. 203–271.
- Kleinberg, J. M.*: Authoritative Sources in a Hyperlinked Environment. In: *Journal of the ACM* 46.5 (1999), S. 604–632.
- : Navigation in a Small World. In: *Nature* 406 (2000), S. 845.
- Kleinberg, J. M. u.a.*: *The Web as a Graph: Measurements, Models, and Methods*. In: *Computing and Combinatorics*. Hrsg. von *Asano u.a.* Berlin und Heidelberg: Springer, 1999, S. 1–17.
- Knappett, C.*: *Network Analysis in Archaeology. New Approaches to Regional Interaction*. Oxford: Oxford University Press, 2013.
- Knoke, D./Yang, S.*: *Social Network Analysis*. 2. Aufl. London: SAGE Publications, 2008.
- Kogut, B./Walker, G.*: The Small World of Germany and the Durability of National Networks. In: *American Sociological Review* 66.3 (2001), S. 317–335.
- König, D.*: *Theorie der endlichen und unendlichen Graphen. Kombinatorische Topologie der Streckenkomplexe*. Leipzig: Akademische Verlagsgesellschaft m.b.H., 1936.
- Koschützki, D. u.a.*: Centrality Indices. In: *Network Analysis: Methodological Foundations. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 16–61.

- Koskinen, J. H./Snijders, T. A.*: Bayesian Inference for Dynamic Social Network Data. In: *Journal of Statistical Planning and Inference* 137 (2007), S. 3930–3938.
- Koskinen, J. H. u.a.*: Bayesian Analysis for Partially Observed Network Data, Missing Ties, Attributes and Actors. In: *Social Networks* 35 (2013), S. 514–527.
- Krapivsky, P. L./Redner, S.*: Network Growth by Copying. In: *Physical Review E* 71 (2005), S. 036118-1–036118-7.
- : Organization of Growing Random Networks. In: *Physical Review E* 63 (2001), S. 066123-1–066123-14.
- Kreager, D. A. u.a.*: Toward a Criminology of Inmate Networks. In: *Justice Quarterly* 33.6 (2016), S. 1000–1028.
- Krebs, V. E.*: Mapping Networks of Terrorist Cells. In: *Connections* 24.3 (2002), S. 43–52.
- Krischke, A./Röpcke, H.*: *Graphen und Netzwerktheorie. Grundlagen – Methoden – Anwendungen.* München: Carl Hanser Verlag, 2015.
- Krishnamurthy, B./Wang, J.*: Topology Modeling via Cluster Graphs. In: *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement* (2001), S. 19–23.
- Krishnamurthy, B. u.a.*: A Socratic Method for Validation of Measurement-Based Networking Research. In: *Computer Communications* 34.1 (2011), S. 43–53.
- Krohn, M. D.*: The Web of Conformity: A Network Approach to the Explanation of Delinquent Behavior. In: *Social Problems* 33.6 (1986), S81–S93.
- Krumke, S. O./Noltemeier, H.*: *Graphentheoretische Konzepte und Algorithmen.* 3. Aufl. Wiesbaden: Springer Vieweg, 2012.
- Kumar, M./Bhatia, R./Rattan, D.*: A Survey of Web Crawlers for Information Retrieval. In: *WIRES Data Mining and Knowledge Discovery* 7 (2017), S. 1–45.
- Lancichinetti, A./Fortunato, S.*: Community Detection Algorithms: A Comparative Analysis. In: *Physical Review E* 80 (2009), S. 056117-1–056117-11.
- : Consensus Clustering in Complex Networks. In: *Scientific Reports* 2 (2012), S. 336-1–336-7.
- Lancichinetti, A./Fortunato, S./Radicchi, F.*: Benchmark Graphs for Testing Community Detection Algorithms. In: *Physical Review E* 78 (2008), S. 046110-1–046110-5.
- Lancichinetti, A./Radicchi, F./Ramasco, J. J.*: Statistical Significance of Communities in Networks. In: *Physical Review E* 81 (2010), S. 046110-1–046110-9.
- Lancichinetti, A. u.a.*: Finding Statistically Significant Communities in Networks. In: *PLoS ONE* 6.4 (2011), e18961-1–e18961-18.
- Landes, W. M./Lessig, L./Solimine, M. E.*: Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges. In: *The Journal of Legal Studies* 27.2 (1998), S. 271–332.
- Langer, M.*: Revolution in Latin American Criminal Procedure: Diffusion of Legal Ideas from the Periphery. In: *The American Journal of Comparative Law* 55.4 (2007), S. 617–676.
- Larose, D. T./Larose, C. D.*: *Data Mining and Predictive Analytics.* 2. Aufl. Hoboken: Wiley, 2015.
- Laumann, E. O./Heinz, J. P.*: Specialization and Prestige in the Legal Profession: The Structure of Deference. In: *American Bar Foundation Research Journal* 2.1 (1977), S. 155–216.
- : The Organization of Lawyers' Work: Size, Intensity, and Co-Practice of the Fields of Law. In: *American Bar Foundation Research Journal* 4.2 (1979), S. 217–246.
- Laumann, E. O. u.a.*: Washington Lawyers and Others: The Structure of Washington Representation. In: *Stanford Law Review* 37.2 (1985), S. 465–502.
- Lazega, E.*: Analyse de Réseaux d'une Organisation Collégiale: Les Avocats d'Affaires. In: *Revue Française de Sociologie* 33.4 (1992), S. 559–589.
- : Concurrence, Coopération et Flux de Conseils dans un Cabinet Américain d'Avocats d'Affaires: Les Échanges d'Idées entre Collègues. In: *Revue Suisse de Sociologie* 21 (1995), S. 61–85.

- Lazega, E.*: Mapping Judicial Dialogue across National Borders: An Exploratory Network Study of Learning from Lobbying among European Intellectual Property Judges. In: *Utrecht Law Review* 8.2 (2012), S. 115–128.
- Lazega, E./Lemerrier, C./Mounier, L.*: A Spinning Top Model of Formal Organization and Informal Behavior: Dynamics of Advice Networks among Judges in a Commercial Court. In: *European Management Review* 3 (2006), S. 113–122.
- Lazega, E./Mounier, L.*: Interlocking Judges: On Joint Exogenous and Self-Governance of Markets. In: *The Governance of Relations in Markets and Organizations*. Hrsg. von *Buskens/Raub/Snijders*. San Francisco: JAI Press, 2003, S. 267–296.
- Lazega, E./Pattison, P. E.*: Multiplexity, Generalized Exchange and Cooperation in Organizations: a Case Study. In: *Social Networks* 21 (1999), S. 67–90.
- Lazega, E./Quitane, E./Casenaz, S.*: Collegial Oligarchy and Networks of Normative Alignments in Transnational Institution Building. In: *Social Networks* 48 (2017), S. 10–22.
- Lazega, E./van Duijn, M.*: Position in Formal Structure, Personal Characteristics and Choices of Advisors in a Law Firm: a Logistic Regression Model for Dyadic Network Data. In: *Social Networks* 19 (1997), S. 375–397.
- Lazega, E. u.a.*: Norms, Status and the Dynamics of Advice Networks: A Case Study. In: *Social Networks* 34 (2012), S. 323–332.
- Lazer, D.*: Networks in Political Science: Back to the Future. In: *PS: Political Science & Politics* 44.1 (2011), S. 61–68.
- Leicht, E. A. u.a.*: Large-Scale Structure of Time Evolving Citation Networks. In: *The European Physical Journal B* 59 (2007), S. 75–83.
- Lerner, J.*: Role Assignments. In: *Network Analysis. Methodological Foundations*. Hrsg. von *Brandes/Erlebach*. Berlin, Heidelberg und New York: Springer, 2005, S. 216–252.
- Li, W. u.a.*: Law is Code: A Software Engineering Approach to Analyzing the United States Code. In: *Journal of Business and Technology Law* 10 (2015), S. 297–374.
- Liebowitz, S. J.*: File Sharing: Creative Destruction or Just Plain Destruction? In: *Journal of Law and Economics* 49 (2006), S. 1–28.
- Liiv, I./Vedeshin, A./Täks, E.*: Visualization and Structure Analysis of Legislative Acts: A Case Study of the Law of Obligations. In: *ICAIL Proceedings* (2007), S. 189–190.
- Lorrain, F./White, H. C.*: Structural Equivalence of Individuals in Social Networks. In: *The Journal of Mathematical Sociology* 1.1 (1971), S. 49–80.
- Lucas, É.*: *Récréations Mathématiques*. Paris: Gauthier-Villars, 1891.
- Luce, R. D./Perry, A. D.*: A Method of Matrix Analysis of Group Structure. In: *Psychometrika* 14.1 (1949), S. 95–116.
- Lupu, Y./Fowler, J. H.*: Strategic Citations to Precedent in the U.S. Supreme Court. In: *The Journal of Legal Studies* 42.1 (2013), S. 151–186.
- Lupu, Y./Voeten, E.*: Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. In: *British Journal of Political Science* 42.2 (2012), S. 413–439.
- Lusher, D./Koskinen, J./Robins, G.*, Hrsg.: *Exponential Random Graph Models for Social Networks*. Cambridge: Cambridge University Press, 2012.
- Mandelbrot, B. B.*: How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. In: *Science* 156 (1967), S. 636–638.
- : *Les Objets Fractals. Forme, Hasard et Dimensions*. Paris: Flammarion, 1975.
- : *The Fractal Geometry of Nature*. New York: W. H. Freeman und Company, 1982.
- Marin, A./Wellman, B.*: Social Network Analysis: An Introduction. In: *The SAGE Handbook of Social Network Analysis*. Hrsg. von *Scott/Carrington*. London: SAGE Publications, 2011, S. 11–25.

- Markose, S. M.*: Systemic Risk Analytics: A Data-Driven Multi-Agent Financial Network (MAFN) Approach. In: *Journal of Banking Regulation* 14.3/4 (2013), S. 285–305.
- Markose, S./Giansante, S./Shaghghi, A. R.*: ‘Too Interconnected to Fail’ Financial Network of US CDS Market: Topological Fragility and Systemic Risk. In: *Journal of Economic Behavior & Organization* 83 (2012), S. 627–646.
- Marquardt, M. C. G.*: UrhG § 5. In: *Praxiskommentar zum Urheberrecht*. Hrsg. von *Wandtke/Bullinger*. 4. Aufl. München: C.H. Beck, 2014.
- Martinelli, A./Nomaler, Ö.*: Measuring Knowledge Persistence: A Genetic Approach to Patent Citation Networks. In: *Journal of Evolutionary Economics* 24.3 (2014), S. 623–652.
- Marx, S. M.*: Citation Networks in the Law. In: *Jurimetrics Journal* 10.4 (1970), S. 121–137.
- McIllwain, J. S.*: Organized Crime: A Social Network Approach. In: *Crime, Law & Social Change* 32 (1999), S. 301–323.
- McPherson, M./Smith-Lovin, L./Cook, J. M.*: Birds of a Feather: Homophily in Social Networks. In: *Annual Review of Sociology* 27 (2001), S. 415–444.
- Merryman, J. H.*: The Authority of Authority: What the California Supreme Court Cited in 1950. In: *Stanford Law Review* 6 (1954), S. 613–673.
- Merton, R. K.*: The Matthew Effect in Science. In: *Science* 159 (1968), S. 56–63.
- Meusel, R. u.a.*: The Graph Structure in the Web – Analyzed on Different Aggregation Levels. In: *The Journal of Web Science* 1 (2015), S. 33–47.
- Milgram, S.*: The Small-World Problem. In: *Psychology Today* 1.1 (1967), S. 61–67.
- Miller, J. H./Page, S. E.*: *Complex Adaptive Systems. An Introduction to Computational Models of Social Life*. Princeton: Princeton University Press, 2007.
- Milo, R. u.a.*: Response to Comment on „Network Motifs: Simple Building Blocks of Complex Networks“ and „Superfamilies of Evolved and Designed Networks“. In: *Science* 305 (2004), S. 1107d.
- Milo, R. u.a.*: Network Motifs: Simple Building Blocks of Complex Networks. In: *Science* 298 (2002), S. 824–827.
- Mirshahvalad, A. u.a.*: Significant Communities in Large Sparse Networks. In: *PLoS ONE* 7.3 (2012), e33721-1–e33721-7.
- Mitchell, M.*: *Complexity. A Guided Tour*. Oxford: Oxford University Press, 2011.
- Mittag, H.-J.*: *Statistik. Eine Einführung mit interaktiven Elementen*. 3. Aufl. Berlin: Springer, 2014.
- Mizruchi, M. S.*: What Do Interlocks Do? An Analysis, Critique, and Assessment of Research on Interlocking Directorates. In: *Annual Review of Sociology* 22 (1996), S. 271–298.
- Moldoveanu, M. C./Baum, J. A. C.*: *Epinets. The Epistemic Structure and Dynamics of Social Networks*. Stanford: Stanford University Press, 2014.
- Molloy, M./Reed, B.*: A Critical Point for Random Graphs with a Given Degree Sequence. In: *Random Structures and Algorithms* 6 (1995), S. 161–179.
- : The Size of the Giant Component of a Random Graph with a Given Degree Sequence. In: *Combinatorics, Probability and Computing* 7.3 (1998), S. 295–305.
- Monasson, R.*: Diffusion, Localization and Dispersion Relations on „Small-World“ Lattices. In: *The European Physical Journal B* 12.4 (1999), S. 555–567.
- Moore, D. S./McCabe, G. P./Craig, B. A.*: *Introduction to the Practice of Statistics*. 9. Aufl. New York: W. H. Freeman und Company, 2017.
- Moreno, J. L.*: Emotions Mapped by New Geography. In: *The New York Times* (3. Apr. 1933), S. 17.
- : *Who Shall Survive? A New Approach to the Problem of Human Interrelations*. Washington, D.C.: Nervous und Mental Disease Publishing Co., 1934.

- Morrissey, R. M.*: Archives of Connection. Whole Network Analysis and Social History. In: *Historical Methods* 48.2 (2015), S. 67–79.
- Morselli, C.*: *Inside Criminal Networks*. New York: Springer, 2009.
- Mucha, P. J. u.a.*: Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. In: *Science* 328 (2010), S. 876–878.
- Munzner, T.*: *Visualization Analysis & Design*. Boca Raton: CRC Press, 2015.
- Murphy, K. P.*: *Machine Learning: A Probabilistic Perspective*. Cambridge: MIT Press, 2012.
- Nakamura, H. u.a.*: The Effect of Patent Family Information in Patent Citation Network Analysis: A Comparative Case Study in the Drivetrain Domain. In: *Scientometrics* 104.2 (2015), S. 437–452.
- Neale, T.*: Citation Analysis of Canadian Case Law. In: *Journal of Open Access to Law* 1 (2013), S. 1–60.
- Nelson, R. L. u.a.*: Lawyers and the Structure of Influence in Washington. In: *Law & Society Review* 22.2 (1988), S. 237–300.
- Newman, M. E. J.*: Ego-Centered Networks and the Ripple Effect. In: *Social Networks* 25 (2003), S. 83–95.
- : Fast Algorithm for Detecting Community Structure in Networks. In: *Physical Review E* 69 (2004), S. 066133-1–066133-5.
- : *Networks. An Introduction*. Oxford: Oxford University Press, 2010.
- : Power Laws, Pareto Distributions and Zipf’s Law. In: *Contemporary Physics* 46.5 (2005), S. 323–351.
- : The Structure and Function of Complex Networks. In: *SIAM Review* 45.2 (2003), S. 167–256.
- Newman, M. E. J./Girvan, M.*: Finding and Evaluating Community Structure in Networks. In: *Physical Review E* 69 (2004), S. 026113-1–026113-15.
- : Mixing Patterns and Community Structure in Networks. In: *Statistical Mechanics of Complex Networks*. Hrsg. von *Pastor-Satorras/Rubil/Diaz-Guilera*. Berlin und Heidelberg: Springer, 2003, S. 66–87.
- Newman, M. E. J./Strogatz, S. H./Watts, D. J.*: Random Graphs with Arbitrary Degree Distributions and their Applications. In: *Physical Review E* (2001), S. 026118-1–026118-17.
- Newman, M. E. J./Watts, D. J.*: Renormalization Group Analysis of the Small-World Network Model. In: *Physics Letters A* 263 (1999), S. 341–346.
- : Scaling and Percolation in the Small-World Network Model. In: *Physical Review E* 60 (1999), S. 7332–7342.
- Nicosia, V. u.a.*: Collective Phenomena Emerging from the Interactions between Dynamical Processes in Multiplex Networks. In: *Physical Review Letters* 118 (2017), S. 138302-1–138302-6.
- Nicosia, V. u.a.*: Growing Multiplex Networks. In: *Physical Review Letters* 111 (2013), S. 058701-1–058701-5.
- Nobile, N./Suen, C. Y.*: Text Segmentation for Document Recognition. In: *Handbook of Document Image Processing and Recognition*. Hrsg. von *Doermann/Tombre*. London: Springer, 2014, S. 257–290.
- Nolte, J.*: *The Philosophy, Theory and Methods of J. L. Moreno. The Man Who Tried to Become God*. New York: Routledge, 2014.
- Northway, M. L.*: A Method for Depicting Social Relationships Obtained by Sociometric Testing. In: *Sociometry* 3.2 (1940), S. 144–150.
- : A Plan for Sociometric Studies in a Longitudinal Programme of Research in Child Development. In: *Sociometry* 17.3 (1954), S. 272–281.

- Nußberger, A.*: Wer zitiert wen? - Zur Funktion von Zitaten bei der Herausbildung gemeineuropäischen Verfassungsrechts. In: *JZ* 61.15/16 (2006), S. 763–770.
- Ogden, P.*: "Mastering the Lawless Science of Our Law": A Story of Legal Citation Indexes. In: *Law Library Journal* 85 (1993), S. 1–48.
- Olsen, H. P./Küçüksu, A.*: Finding Hidden Patterns in ECtHR's Case Law: On How Citation Network Analysis Can Improve Our Knowledge of ECtHR's Article 14 Practice. In: *International Journal of Discrimination and the Law* 17.1 (2017), S. 4–22.
- Paik, A./Heinz, J. P./Southworth, A.*: Political Lawyers: The Structure of a National Network. In: *Law & Social Inquiry* 36.4 (2011), S. 892–918.
- Paik, A./Southworth, A./Heinz, J. P.*: Lawyers of the Right: Networks and Organization. In: *Law & Social Inquiry* 32.4 (2007), S. 883–917.
- Palla, G./Barabási, A.-L./Vicsek, T.*: Quantifying Social Group Evolution. In: *Nature Letters* 446 (2007), S. 664–667.
- Palla, G. u.a.*: Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society. In: *Nature* 435 (2005), S. 814–818.
- Pam, M.*: Interlocking Directorates, the Problem and its Solution. In: *Harvard Law Review* 26.6 (1913), S. 467–492.
- Panagis, Y./Šadl, U.*: The Force of EU Case Law: A Multi-Dimensional Study of Case Citations. In: *Legal Knowledge and Information Systems*. Hrsg. von *Rotolo*. Amsterdam: IOS Press, 2015, S. 71–80.
- Papachristos, A. V.*: Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. In: *American Journal of Sociology* 115.1 (2009), S. 74–128.
- : The Network Structure of Crime. In: *Sociology Compass* 8.4 (2014), S. 347–357.
- Papachristos, A. V./Hureau, D. M./Braga, A. A.*: The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence. In: *American Sociological Review* 78.3 (2013), S. 417–447.
- Papachristos, A. V./Meares, T. L./Fagan, J.*: Why Do Criminals Obey the Law? The Influence of Legitimacy and Social Networks on Active Gun Offenders. In: *The Journal of Criminal Law & Criminology* 102.2 (2012), S. 397–440.
- Pastor-Satorras, R./Vespignani, A.*: Immunization of Complex Networks. In: *Physical Review E* 65 (2002), S. 036104-1–036104-8.
- Pattison, P./Wasserman, S.*: Logit Models and Logistic Regressions for Social Networks: II. Multivariate Relations. In: *British Journal of Mathematical and Statistical Psychology* 52 (1999), S. 169–193.
- Pecora, N./Spelta, A.*: Shareholding Relationships in the Euro Area Banking Market: A Network Perspective. In: *Physica A* 434 (2015), S. 1–12.
- Pelc, K. J.*: The Politics of Precedent in International Law: A Social Network Application. In: *American Political Science Review* 108.3 (2014), S. 547–564.
- Peltonen, T. A./Scheicher, M./Vuilleme, G.*: The Network Structure of the CDS Market and its Determinants. In: *Journal of Financial Stability* 13 (2014), S. 118–133.
- Petersen, V.*: Interlocking Directorates in the European Union: An Argument for Their Restriction. In: *European Business Law Review* 27.6 (2016), S. 821–864.
- Peterson, N. R.*: Distance Between Two Random  $k$ -Out Digraphs, With and Without Preferential Attachment. In: *Advances in Applied Probability* 47 (2015), S. 858–879.
- Pfeffer, J.*: Visualisierung sozialer Netzwerke. In: *Netzwerkanalyse und Netzwerktheorie. Ein neues Paradigma in den Sozialwissenschaften*. Hrsg. von *Stegbauer*. 2. Aufl. Wiesbaden: VS Verlag, 2010, S. 227–238.

- Piškorec, M./Sluban, B./Šmuc, T.*: MultiNets: Web-Based Multilayer Network Visualization. In: Machine Learning and Knowledge Discovery in Databases. Hrsg. von *Bifet u.a.* Cham: Springer, 2015, S. 298–302.
- Polak, S.*: Algorithms for the Network Analysis of Bilateral Tax Treaties. Master Thesis. Universiteit van Amsterdam, 23. Dez. 2014.
- Poledna, S./Thurner, S.*: Elimination of Systemic Risk in Financial Networks by Means of a Systemic Risk Transaction Tax. In: Quantitative Finance 16.10 (2016), S. 1599–1613.
- Pollner, P./Palla, G./Vicsek, T.*: Preferential Attachment of Communities: The Same Principle, but a Higher Level. In: Europhysics Letters 73.3 (2006), S. 478–484.
- Porter, M. A./Onnela, J.-P./Mucha, P. J.*: Communities in Networks. In: Notices of the American Mathematical Society 56.9 (2009), S. 1082–1097.
- Porter, M. A. u.a.*: A Network Analysis of Committees in the U.S. House of Representatives. In: PNAS 102.20 (2005), S. 7057–7062.
- Porter, M. A. u.a.*: Community Structure in the United States House of Representatives. In: Physica A 386 (2007), S. 414–438.
- Post, D. G./Eisen, M. B.*: How Long is the Coastline of the Law? Thoughts on the Fractal Nature of Legal Systems. In: The Journal of Legal Studies 29 (2000), S. 545–584.
- Post, D. G./Johnson, C. A.*: Chaos Prevailing on Every Continent: Towards a New Theory of Decentralized Decision-Making in Complex Systems. In: Chicago-Kent Law Review (1998), S. 1055–1099.
- Price, D. d. S.*: A General Theory of Bibliometric and Other Cumulative Advantage Processes. In: Journal of the American Society for Information Science 27 (1976), S. 292–306.
- Price, D. J. d. S.*: Networks of Scientific Papers. In: Science 149.3683 (1965), S. 510–515.
- Puig, S.*: Social Capital in the Arbitration Market. In: The European Journal of International Law 25.2 (2014), S. 387–424.
- Pujo Committee*: Report of the Committee Appointed Pursuant to House Resolutions 428 and 504 to Investigate the Concentration of Control of Money and Credit. House Report No. 1593. 62nd Congress, 3rd Session, 28. Feb. 1913.
- Radicchi, F./Fortunato, S./Vespignani, A.*: Citation Networks. In: Models of Science Dynamics. Hrsg. von *Scharnhorst/Börner/van den Besselaar*. Berlin und Heidelberg: Springer, 2012, S. 233–257.
- Radke, H./Hilgert, P./Mardorf, D.*: Die Beschaffung von juristischen Datenbanken als Vergabeproblem. In: NVwZ (2008), S. 1070–1073.
- Raghavan, U. N./Albert, R./Kumara, S.*: Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks. In: Physical Review E 76 (2007), S. 036106-1–036106-11.
- Reichardt, J./Bornholdt, S.*: Detecting Fuzzy Community Structures in Complex Networks with a Potts Model. In: Physical Review Letters 93.21 (2004), S. 218701-1–218701-4.
- : Statistical Mechanics of Community Detection. In: Physical Review E 74 (2006), S. 016110-1–016110-14.
- Reynolds, G. H.*: Chaos and the Court. In: Columbia Law Review 91.1 (1991), S. 110–117.
- Rhodes, C. J./Keefe, E. M. J.*: Social Network Topology: a Bayesian Approach. In: Journal of the Operational Research Society 58 (2007), S. 1605–1611.
- Rich, M. L.*: Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment. In: University of Pennsylvania Law Review 164 (2016), S. 871–929.
- Richter, H.*: Zur Weiterverwendung von Informationen der öffentlichen Hand: BVerwG klärt erstmals grundsätzliche Anwendungsvoraussetzungen des IWG. In: NVwZ (2016), S. 1143–1145.

- Riedl, C. u.a.: Detecting Figures and Part Labels in Patents: Competition-Based Development of Graphics Recognition Algorithms. In: *International Journal on Document Analysis and Recognition* 19 (2016), S. 155–172.
- Robins, G. u.a.: An Introduction to Exponential Random Graph (p\*) Models for Social Networks. In: *Social Networks* 29 (2007), S. 173–191.
- Robins, G. u.a.: Recent Developments in Exponential Random Graph (p\*) Models for Social Networks. In: *Social Networks* 29 (2007), S. 192–215.
- Roe, M. J.: Chaos and Evolution in Law and Economics. In: *Harvard Law Review* 109 (1996), S. 641–668.
- Rogers, E.: *Diffusion of Innovations*. 5. Aufl. New York: Free Press, 2003.
- Röhl, K. F.: Bilder in gedruckten Rechtsbüchern. In: *Die Sprache des Rechts*. Hrsg. von Lerch. Bd. 3. Berlin und New York: Walter de Gruyter, 2005, S. 267–348.
- Rossi, L./Magnani, M.: Towards Effective Visual Analytics on Multiplex and Multilayer Networks. In: *Chaos, Solitons & Fractals* 72 (2015), S. 68–76.
- Rosvall, M./Axelsson, D./Bergstrom, C. T.: The Map Equation. In: *The European Physical Journal Special Topics* 178 (2009), S. 13–23.
- Rosvall, M./Bergstrom, C. T.: An Information-Theoretic Framework for Resolving Community Structure in Complex Networks. In: *PNAS* 104.18 (2007), S. 7327–7331.
- : Maps of Random Walks on Complex Networks Reveal Community Structure. In: *PNAS* 105.4 (2008), S. 1118–1123.
- Rouchier, J./Lazega, E./Mounier, L.: Articulation of Hierarchy and Networks as an Evolving Social Structure. In: *Agent-Based Approaches in Economic and Social Complex Systems IV*. Hrsg. von Terano u.a. Tokyo: Springer, 2007, S. 97–104.
- Rouchier, J./Tubaro, P.: Can Opinion Be Stable in an Open Network with Hierarchy? An Agent-Based Model of the Commercial Court of Paris. In: *Procedia Social and Behavioral Sciences* 10 (2011), S. 123–131.
- Rouse, W. W.: *Mathematical Recreations and Problems of Past and Present Times*. New York: Macmillan, 1892.
- Ruhl, J. B.: Complexity Theory as a Paradigm for the Dynamical Law-and-Society System: A Wake-Up Call for Legal Reductionism and the Modern Administrative State. In: *Duke Law Journal* 45.5 (1996), S. 849–928.
- : Law’s Complexity: A Primer. In: *Georgia State University Law Review* 24.4 (2008), S. 885–911.
- : Managing Systemic Risk in Legal Systems. In: *Indiana Law Journal* 89 (2014), S. 559–603.
- : The Fitness of Law: Using Complexity Theory to Describe the Evolution of Law and Society and Its Practical Meaning for Democracy. In: *Vanderbilt Law Review* 49 (1996), S. 1406–1490.
- : Thinking of Environmental Law as a Complex Adaptive System: How to Clean up the Environment by Making a Mess of Environmental Law. In: *Houston Law Review* 34.4 (1997), S. 933–1002.
- : Thinking of Mediation as a Complex Adaptive System. In: *BYU Law Review* (1997), S. 777–801.
- Ruhl, J. B./Katz, D. M.: Measuring, Monitoring and Managing Legal Complexity. In: *Iowa Law Review* 101 (2015), S. 191–244.
- Ruhl, J. B./Katz, D. M./Bommarito, M. J.: Harnessing Legal Complexity. In: *Science* 355 (2017), S. 1377–1378.
- Ruhl, J. B./Ruhl, H. J.: The Arrow of the Law in Modern Administrative States: Using Complexity Theory to Reveal the Diminishing Returns and Increasing Risks the Burgeoning of Law Poses to Society. In: *U.C. Davis Law Review* 30 (1997), S. 405–482.

- Russell, S. J./Norvig, P.*: Artificial Intelligence. A Modern Approach. 3. Aufl. Upper Saddle River: Pearson Education, 2011.
- Saban, D./Bonomo, F./Stier-Moses, N. E.*: Analysis and Models of Bilateral Investment Treaties Using a Social Networks Approach. In: *Physica A* 389 (2010), S. 3661–3673.
- Sachs, H./Stiebitz, M./Wilson, R. J.*: Euler's Königsberg Letters. In: *Journal of Graph Theory* 12.1 (1988), S. 133–139.
- Šadl, U.*: The Role of *Effet Utile* in Preserving the Continuity and Authority of European Union Law: Evidence from the Citation Web of the Pre-Accession Case Law of the Court of Justice of the EU. In: *European Journal of Legal Studies* 8.1 (2015), S. 18–45.
- Šadl, U./Olsen, H. P.*: Can Quantitative Methods Complement Doctrinal Legal Studies? Using Citation Network and Corpus Linguistic Analysis to Understand International Courts. In: *Leiden Journal of International Law* 30 (2017), S. 327–349.
- Salathé, M./Jones, J. H.*: Dynamics and Control of Diseases in Networks with Community Structure. In: *PLoS Computational Biology* 6.4 (2010), e1000736-1–e1000736-11.
- Salehi, M. u.a.*: Spreading Processes in Multilayer Networks. In: *IEEE Transactions on Network Science and Engineering* 2.2 (2015), S. 65–83.
- Schadt, H.*: Die Darstellungen der Arbores Consanguinitatis und der Arbores Affinitatis. Bildschemata in juristischen Handschriften. Tübingen: Wasmuth, 1982.
- Schaks, N./Kahmert, S.*: Die Einführung einer Impfpflicht zur Bekämpfung der Masern. Eine zulässige staatliche Handlungsoption. In: *MedR* 33 (2015), S. 860–866.
- Schaper, M.*: The Structure and Organization of EU Law in the Field of Direct Taxes. Amsterdam: IBFD, 2013.
- Schelling, T. C.*: Models of Segregation. In: *The American Economic Review* 59.2 (1969), S. 488–493.
- Schildknecht, W.*: Designing for Robustness: Overcoming Systemic Risk in the Political Branches. In: *California Law Review* 103 (2015), S. 433–466.
- Schöningh, U.*: Theoretische Informatik – kurz gefasst. 5. Aufl. Heidelberg: Spektrum Akademischer Verlag, 2008.
- Schwartz, D. M./Rouselle, T.*: Using Social Network Analysis to Target Criminal Networks. In: *Trends in Organized Crime* 12 (2009), S. 188–207.
- Scott, R. E.*: Chaos Theory and the Justice Paradox. In: *William & Mary Law Review* 35 (1993), S. 329–351.
- Seidel, A. H.*: Citation System for Patent Office. In: *Journal of the Patent Office Society* 31.7 (1949), S. 554.
- Seifi, M. u.a.*: Stable Community Cores in Complex Networks. In: *Complex Networks*. Hrsg. von *Menezes/Evsukoff/González*. Berlin und Heidelberg: Springer, 2013, S. 87–98.
- Shamir, R./Sharan, R./Tsur, D.*: Cluster Graph Modification Problems. In: *Discrete Applied Mathematics* 144 (2004), S. 173–182.
- Shapiro, F. R.*: Origins of Bibliometrics, Citation Indexing, and Citation Analysis: The Neglected Legal Literature. In: *Journal of the American Society for Information Science* 43.5 (1992), S. 337–339.
- Shen-Orr, S. S. u.a.*: Network Motifs in the Transcriptional Regulation Network of *Escherichia Coli*. In: *Nature Genetics* 31 (2002), S. 64–68.
- Shomade, S. A./Hartley, R. E.*: The Application of Network Analysis to the Study of Trial Courts. In: *Justice System Journal* 31.2 (2010), S. 144–163.
- Shur-Ofry, M.*: IP and the Lens of Complexity. In: *The Intellectual Property Law Review* 54.1 (2013), S. 55–102.
- Siems, M. M.*: Citation Patterns of the German Federal Supreme Court and of the Court of Appeal of England and Wales. In: *King's Law Journal* 21.1 (2010), S. 152–171.

- : The Web of Creditor and Shareholder Protection in 25 Countries: A Comparative Legal Network Analysis. In: *Arizona Journal of International & Comparative Law* 27.3 (2010), S. 747–784.
- : Varieties of Legal Systems: Towards a New Global Taxonomy. In: *Journal of Institutional Economics* 12.3 (2016), S. 579–602.
- Silverman, G. M.*: Rise of the Machines: Justice Information Systems and the Question of Public Access to Court Records Over the Internet. In: *Washington Law Review* (2004), S. 175–221.
- Simkin, M. V./Roychowdhury, V. P.*: Stochastic Modeling of Citation Slips. In: *Scientometrics* 62.3 (2005), S. 367–384.
- Simmons, E. S.*: ‚Black Sheep‘ in the Patent Family. In: *World Patent Information* 31 (2009), S. 11–18.
- Small, H.*: Co-Citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. In: *Journal of the American Society for Information Science* 24.4 (1973), S. 265–269.
- Smith, A. R.*: Color Gamut Transform Pairs. In: *SIGGRAPH 78 Conference Proceedings* (1978), S. 12–19.
- Smith, T. A.*: The Web of Law. In: *San Diego Law Review* 44 (2007), S. 309–354.
- Solomonoff, R./Rapoport, A.*: Connectivity of Random Nets. In: *Bulletin of Mathematical Biophysics* 13 (1951), S. 107–117.
- Sonquist, J. A./Koenig, T.*: Interlocking Directorates in the Top U.S. Corporations: A Graph Theory Approach. In: *Insurgent Sociologist* 5 (1975), S. 196–229.
- Southworth, A./Paik, A./Heinz, J. P.*: Lawyers in National Policymaking. In: *The Paradox of Professionalism: Lawyers and the Possibility of Justice*. Hrsg. von *Cummings*. Cambridge: Cambridge University Press, 2011, S. 220–242.
- Spaeth, H. J./Altfeld, M. F.*: Influence Relationships within the Supreme Court: A Comparison of the Warren and Burger Courts. In: *The Western Political Quarterly* 38.1 (1985), S. 70–83.
- Spamann, H.*: Contemporary Legal Transplants: Legal Families and the Diffusion of (Corporate) Law. In: *BYU Law Review* (2009), S. 1813–1877.
- Sparrow, M. K.*: The Application of Network Analysis to Criminal Intelligence. In: *Social Networks* 13 (1991), S. 251–274.
- Stauffer, D.*: Small World. In: *Handbuch Netzwerkforschung*. Hrsg. von *Stegbauer/Häußling*. Wiesbaden: VS Verlag, 2010, S. 219–225.
- Stearns, L. B./Mizruchi, M. S.*: Board Composition and Corporate Financing: The Impact of Financial Institution Representation on Borrowing. In: *The Academy of Management Journal* 36.3 (1993), S. 603–618.
- Stegbauer, C.*: „Die Invasion der Physiker“: Naturwissenschaft und Soziologie in der Netzwerkanalyse. In: *Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006*. Hrsg. von *Deutsche Gesellschaft für Soziologie*. Frankfurt am Main: Campus Verlag, 2008, S. 1060–1077.
- : Reziprozität. In: *Handbuch Netzwerkforschung*. Hrsg. von *Stegbauer/Häußling*. Wiesbaden: VS Verlag, 2010, S. 113–122.
- Stegbauer, C./Häußling, R.*: Einleitung in das Handbuch Netzwerkforschung. In: *Handbuch Netzwerkforschung*. Hrsg. von *Stegbauer/Häußling*. Wiesbaden: VS Verlag, 2010, S. 13–16.
- : Einleitung: Selbstverständnis der Netzwerkforschung. In: *Handbuch Netzwerkforschung*. Hrsg. von *Stegbauer/Häußling*. Wiesbaden: VS Verlag, 2010, S. 57–60.
- Sternitzke, C./Bartkowski, A./Schramm, R.*: Visualizing Patent Statistics by Means of Social Network Analysis Tools. In: *World Patent Information* 30.2 (2008), S. 115–131.
- Stevens, S. S.*: On the Theory of Scales of Measurement. In: *Science* 103 (1946), S. 677–680.

- Stevens, S. S.*: Psychophysics. Introduction to Its Perceptual, Neural, and Social Prospects. New York: Wiley, 1975.
- Stevens, S. S./Guirao, M.*: Subjective Scaling of Length and Area and the Matching of Length to Loudness and Brightness. In: *Journal of Experimental Psychology* 66.2 (1963), S. 177–186.
- Stirling, J.*: *Methodus Differentialis: Sive Tractatus de Summatione et Interpolatione Serierum Infinitarum*. London: Gul. Bowyer, 1730.
- Strahilevitz, L. J.*: Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks. In: *Virginia Law Review* 89.3 (2003), S. 505–595.
- Strandburg, K. J. u.a.*: Law and the Science of Networks: An Overview and an Application to the ‘Patent Explosion’. In: *Berkeley Technology Law Journal* 21.4 (2006), S. 1293–1362.
- : Patent Citation Networks Revisited: Signs of a Twenty-First Century Change? In: *North Carolina Law Review* 87 (2009), S. 1657–1698.
- Strang, S.*: Network Analysis in Criminal Intelligence. In: *Networks and Network Analysis for Defence and Security*. Hrsg. von *Masys*. Cham: Springer, 2014, S. 1–26.
- Strogatz, S. H.*: Exploring Complex Networks. In: *Nature* 410 (2001), S. 268–276.
- Summer, M.*: Financial Contagion and Network Analysis. In: *Annual Review of Financial Economics* 5 (2013), S. 277–297.
- Sweeney, P. M. u.a.*: Network Analysis of Manually-Encoded State Laws and Prospects for Automation. In: *Network Analysis in Law*. Hrsg. von *Winkels/Lettieri/Faro*. Napoli: Edizioni Scietifiche Italiane, 2014, S. 53–77.
- Takes, F. W./Heemskerk, E. M.*: Centrality in the Global Network of Corporate Control. In: *Social Network Analysis and Mining* 6 (2016), S. 97–1–97-18.
- Tam Cho, W. K./Fowler, J. H.*: Legislative Success in a Small World: Network Analysis and the Dynamics of Congressional Legislation. In: *The Journal of Politics* 72.1 (2010), S. 124–135.
- Tarissan, F./Nollez-Goldbach, R.*: Analysing the First Case of the International Criminal Court from a Network-Science Perspective. In: *Journal of Complex Networks* 4.4 (2016), S. 616–634.
- : Temporal Properties of Legal Decision Networks: A Case Study from the International Criminal Court. In: *Legal Knowledge and Information Systems*. Hrsg. von *Rotolo*. Amsterdam: IOS Press, 2015, S. 111–120.
- : The Network of the International Criminal Court Decisions as a Complex System. In: *Emergence, Complexity and Computation*. Hrsg. von *Sanayei/Zelinka/Rosslers*. Berlin und Heidelberg: Springer, 2013, S. 255–264.
- Tarissan, F./Panagis, Y./Šadl, U.*: Selecting the Cases that Defined Europe: Complementary Metrics for a Network Analysis. In: *IEEE/ACM ASONAM* (2016), S. 661–668.
- Tayebi, M. A./Glässer, U.*: *Social Network Analysis in Predictive Policing. Concepts, Models and Methods*. Cham: Springer, 2016.
- Telea, A. C.*: *Data Visualization. Principles and Practice*. 2. Aufl. Boca Raton: Taylor & Francis, 2015.
- Tita, G. E./Radil, S. M.*: Spatializing the Social Networks of Gangs to Explore Patterns of Violence. In: *Journal of Quantitative Criminology* 27.4 (2011), S. 521–545.
- Torgerson, W. S.*: Multidimensional Scaling: I. Theory and Method. In: *Psychometrika* 17.4 (1952), S. 401–419.
- Torrance, A. W./West, J. D.*: All Patents Great and Small: A Big Data Network Approach to Valuation. In: *Virginia Journal of Law & Technology* 20.3 (2017), S. 466–504.
- Toutenburg, H./Heumann, C.*: *Deskriptive Statistik*. 7. Aufl. Heidelberg: Springer, 2009.
- Tufte, E. R.*: *The Visual Display of Quantitative Information*. 2. Aufl. Cheshire: Graphics Press, 2001.

- Tukey, J. W.: Some Graphic and Semigraphic Displays. In: Statistical Papers in Honor of George W. Snedecor. Hrsg. von Bancroft. Ames: Iowa State University Press, 1972, S. 293–316.
- Twining, W.: Mapping Law. The MacDermott Lecture. In: Northern Ireland Legal Quarterly 50.1 (1999), S. 12–49.
- Upper, C.: Simulation Methods to Assess the Danger of Contagion in Interbank Markets. In: Journal of Financial Stability 7 (2011), S. 111–125.
- Uzzi, B./Lancaster, R.: Embeddedness and Price Formation in the Corporate Law Market. In: American Sociological Review 69 (2004), S. 319–344.
- van Veen, K./Kratzer, J.: National and International Interlocking Directorates within Europe: Corporate Networks within and among Fifteen European Countries. In: Economy & Society 40.1 (2011), S. 1–25.
- van't Riet, M./Lejour, A.: Ranking the Stars. Network Analysis of Bilateral Tax Treaties. In: CPB Discussion Paper 290 (2014), S. 1–68.
- Vicsek, T. u.a.: Novel Type of Phase Transition in a System of Self-Driven Particles. In: Physical Review Letters 75.6 (1995), S. 1226–1229.
- Vincent, W.: Legal XML and Standards for the Legal Industry. In: S.M.U. Law Review 53 (2000), S. 1395–1429.
- Vitanov, N. K./Ausloos, M. R.: Knowledge Epidemics and Population Dynamics Models for Describing Idea Diffusion. In: Models of Science Dynamics. Hrsg. von Schamhorst/Börner/van den Besselaar. Berlin und Heidelberg: Springer, 2012, S. 69–125.
- Vogel, F./Hamann, H./Gauer, I.: Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies. In: Law & Social Inquiry (2017), DOI: 10.1111/lasi.12305.
- Vogel, F./Pötters, S./Christensen, R.: Richterrecht der Arbeit – empirisch untersucht. Möglichkeiten und Grenzen computergestützter Textanalyse am Beispiel des Arbeitnehmerbegriffs. Berlin: Duncker & Humblot, 2015.
- Volkman, L.: Fundamente der Graphentheorie. Wien: Springer, 1996.
- Vörös, A./Snijders, T. A.: Cluster Analysis of Multiplex Networks: Defining Composite Network Measures. In: Social Networks 49 (2017), S. 93–112.
- Wagner-Döbler, R.: Präjudizien in deutschen, englischen und US-amerikanischen Gerichtsentscheidungen: Ein quantitativer Vergleich. In: RabelsZ 59 (1995), S. 113–127.
- : The Frequency Distribution of Legal Decision Citations in the German Jurisdiction. In: Scientometrics 29.1 (1994), S. 15–26.
- Wagner-Döbler, R./Philipps, L.: Präjudizien in der Rechtsprechung: Statistische Untersuchungen anhand der Zitierpraxis deutscher Gerichte. In: Rechtslehre 23 (1992), S. 228–241.
- Walker, J. L.: The Diffusion of Innovations Among the American States. In: The American Political Science Review 63.3 (1969), S. 880–899.
- Ward, M./Grinstein, G./Keim, D.: Interactive Data Visualization. Foundations, Techniques, and Applications. 2. Aufl. Boca Raton: Taylor & Francis, 2015.
- Ward, M. D./Stovel, K./Sacks, A.: Network Analysis and Political Science. In: Annual Review of Political Science 14 (2011), S. 246–264.
- Ware, C.: Information Visualization. Perception for Design. 3. Aufl. Waltham: Elsevier, 2013.
- Wasserman, S./Faust, K.: Social Network Analysis. Methods and Applications. Cambridge: Cambridge University Press, 1994.
- Wasserman, S./Pattison, P.: Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and  $p^*$ . In: Psychometrika 61.3 (1996), S. 401–425.
- Watson, A.: Legal Transplants. 2. Aufl. Athens: The University of Georgia Press, 1993.
- Watts, D. J.: Six Degrees. The Science of a Connected Age. New York: W. W. Norton & Company, 2003.
- : The “New” Science of Networks. In: Annual Review of Sociology 30 (2004), S. 243–270.

- Watts, D. J./Strogatz, S. H.: Collective Dynamics of ‘Small-World’ Networks. In: *Nature* 393 (1998), S. 440–442.
- Wei, Y. u.a.: Credit Scoring with Social Network Data. In: *Marketing Science* 35.2 (2016), S. 234–258.
- Weisberg, M.: *Simulation and Similarity. Using Models to Understand the World*. Oxford: Oxford University Press, 2013.
- Whalen, R.: Bad Law Before It Goes Bad: Citation Networks and the Life Cycle of Overruled Supreme Court Precedent. In: *Network Analysis in Law*. Hrsg. von *Winkels/Lettieri/Faro*. Napoli: Edizioni Scietifiche Italiane, 2014, S. 1–21.
- : Legal Networks: The Promises and Challenges of Legal Network Analysis. In: *Michigan State Law Review* (2016), S. 539–565.
- : Modeling Annual Supreme Court Influence: The Role of Citation Practices and Judicial Tenure in Determining Precedent Network Growth. In: *Complex Networks*. Hrsg. von *Menezes/Evsukoff/González*. Berlin und Heidelberg: Springer, 2013, S. 169–176.
- White, D. R./Reitz, K. P.: Graph and Semigroup Homomorphisms on Networks of Relations. In: *Social Networks* 5 (1983), S. 193–234.
- Widener, M./Weiner, M. S.: *Law’s Picture Books*. The Yale Law Library Collection. Clark: Talbot Publishing, 2017.
- Wiebe, A.: UrhG § 5. In: *Recht der elektronischen Medien*. Hrsg. von *Spindler/Schuster*. 3. Aufl. München: C.H. Beck, 2015.
- : UrhG § 87b. In: *Recht der elektronischen Medien*. Hrsg. von *Spindler/Schuster*. 3. Aufl. München: C.H. Beck, 2015.
- Wilson, J. D. u.a.: A Testing Based Extraction Algorithm for Identifying Significant Communities in Networks. In: *Annals of Applied Statistics* 8.3 (2014), S. 1853–1891.
- Wilson, R. J.: An Eulerian Trail Through Königsberg. In: *Journal of Graph Theory* 10.3 (1986), S. 265–275.
- Windolf, P.: Coordination and Control in Corporate Networks: United States and Germany in Comparison, 1896–1938. In: *European Sociological Review* 25.4 (2009), S. 443–457.
- Winkels, R./Boer, A./Plantevin, I.: Creating Context Networks in Dutch Legislation. In: *Legal Knowledge and Information Systems*. Hrsg. von *Ashley*. Amsterdam: IOS Press, 2014, S. 155–164.
- Winkels, R./de Ruyter, J.: Survival of the Fittest: Network Analysis of Dutch Supreme Court Cases. In: *AI Approaches to the Complexity of Legal Systems*. Hrsg. von *Palmirani u.a.* Berlin und Heidelberg: Springer, 2012, S. 106–115.
- Winkels, R./de Ruyter, J./Kroese, H.: Determining Authority of Dutch Case Law. In: *Frontiers in Artificial Intelligence and Applications*. Hrsg. von *Atkinson*. Amsterdam: IOS Press, 2011, S. 103–112.
- Winship, C.: Thoughts about Roles and Relations: An Old Document Revisited. In: *Social Networks* 10 (1988), S. 209–231.
- Winship, C./Mandel, M.: Roles and Positions: A Critique and Extension of the Blockmodeling Approach. In: *Sociological Methodology* 14 (1983), S. 314–344.
- Wolcott, C. L.: Citation in Examiners’ Letters of References Introduced by Applicants. In: *Journal of the Patent Office Society* 29.3 (1947), S. 214.
- Yan, E./Ding, Y.: Scholarly Network Similarities: How Bibliographic Coupling Networks, Citation Networks, Cocitation Networks, Topical Networks, Coauthorship Networks, and Coword Networks Relate to Each Other. In: *Journal of the American Association for Information Science and Technology* 63.7 (2012), S. 1313–1326.
- Yang, G.-C. u.a.: Using the Comprehensive Patent Citation Network (CPC) to Evaluate Patent Value. In: *Scientometrics* 105.3 (2015), S. 1319–1346.

- Zhang, Y. u.a.: Community Structure in Congressional Cosponsorship Networks. In: *Physica A* 387 (2008), S. 1705–1712.
- Ziegler, R.: Netzwerkanalyse: Metapher, Methode oder strukturelles Forschungsprogramm für die Sozialwissenschaften? In: *Zeitschrift für Klinische Psychologie* 16.4 (1987), S. 339–352.
- Zollner, M.: *Komplexität und Recht*. Berlin: Duncker & Humblot, 2014.
- Zuck, R.: Die Wissenschaftlichen Mitarbeiter des Bundesverfassungsgerichts. In: *Handbuch Bundesverfassungsgericht im politischen System*. Hrsg. von van Ooyen/Möllers. 2. Aufl. Wiesbaden: Springer VS, 2015, S. 443–457.
- Zweig, K. A.: Are Word-Adjacency Networks Networks? In: *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*. Hrsg. von Mehler u.a. Heidelberg: Springer, 2016, S. 153–163.
- : Good versus Optimal: Why Network Analytic Methods Need More Systematic Evaluation. In: *Central European Journal of Computer Science* 1 (2011), S. 137–153.
- : *Network Analysis Literacy*. Berlin: Springer, 2016.

# Glossar

Für viele der hier zusammengestellten Begriffe gibt es mehrere Definitionsmöglichkeiten. Die angegebenen Definitionen entsprechen der Begriffsverwendung in dieser Arbeit.

$\Sigma$	Summenzeichen; $\sum_{i=1}^n i = 1 + 2 + \dots + n$ .
$\Pi$	Produktzeichen; $\prod_{i=1}^n i = 1 \cdot 2 \cdot \dots \cdot n$ .
$n!$	Fakultät, d.h. Anzahl der Permutationen von $n$ Elementen: $n! = \prod_{i=1}^n i = n \cdot (n-1) \cdot \dots \cdot 1$ mit $0! = 1$ per Definition.
$\binom{n}{k}$	Binomialkoeffizient, d.h. Anzahl der paarweise verschiedenen $k$ -elementigen Teilmengen einer Menge mit $n$ Elementen: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ .
$\neg$	Symbol für die logische Negation (Nicht).
$\wedge$	Symbol für die logische Konjunktion (Und).
$\vee$	Symbol für die logische Disjunktion (Oder).
$\Rightarrow$	Symbol für die logische Implikation: $a \Rightarrow b$ (aus $a$ folgt $b$ ).
$\Leftrightarrow$	Symbol für die logische Äquivalenz: $a \Leftrightarrow b$ (aus $a$ folgt $b$ und aus $b$ folgt $a$ ).
$\forall$	Allquantor (eine Eigenschaft muss für alle Elemente einer bestimmten Menge gelten).
$\exists$	Existenzquantor (eine Eigenschaft muss für mindestens ein Element einer bestimmten Menge gelten).
$\in$	Symbol für die Mitgliedschaft in einer Menge.
$\cap$	Symbol für den Schnitt zweier Mengen.
$\cup$	Symbol für die Vereinigung zweier Mengen.
$\emptyset$	Leere Menge.
<b>Adjazenzmatrix</b>	Matrix, die Informationen darüber enthält, welche Punkte in einem Graphen auf welche Weise miteinander verbunden sind.
<b>Äquivalenzklasse</b>	Maximale Menge zueinander äquivalenter Elemente; Menge in einer Zerlegung, die durch eine Äquivalenzrelation definiert wird.
<b>Äquivalenzrelation</b>	Relation, die reflexiv, symmetrisch und transitiv ist; zerlegt die Menge, auf der sie definiert ist, in Äquivalenzklassen.
<b>Algorithmus</b>	Eindeutig spezifizierte Prozedur zur Lösung eines Problems; Abfolge von Rechenschritten, die eine Eingabe in eine Ausgabe transformiert.
<b>API</b>	Programmierschnittstelle, die einer Anwendung ermöglicht, die Funktionalität einer anderen Anwendung zu nutzen; kurz für <i>Application Programming Interface</i> .
<b>Arithmetisches Mittel</b>	Für ein metrisch skaliertes Merkmal die Summe der

<b>Baum</b>	Beobachtungswerte, geteilt durch ihre Anzahl: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .
<b>Beobachtung</b>	Verbundener Graph ohne Zyklen.
<b>Beziehung</b>	Untersuchungseinheit, deren Merkmalsausprägungen im Rahmen einer statistischen Untersuchung erhoben wurden.
<b>Bibliographische Kopplung</b>	Element eines Netzwerks, das Objekte miteinander verbindet; im Englischen <i>relationship</i> .
<b>Boxplot</b>	Beziehung zwischen zwei Dokumenten, die beide (mindestens) ein anderes Dokument referenzieren.
<b>CCDF</b>	Rechteckbasierte Darstellung der Lage von unterem Quartil, Median, oberem Quartil, Randwerten und Extremwerten einer Verteilung.
<b>CDF</b>	Funktion, die für eine Verteilung die Wahrscheinlichkeit anzeigt, dass ein Wert $X$ größer als ein bestimmter Wert $x$ ist: $p(X > x)$ , zum Teil auch $p(X \geq x)$ ; kurz für <i>complementary cumulative distribution function</i> .
<b>Clique</b>	Funktion, die für eine Verteilung die Wahrscheinlichkeit anzeigt, dass ein Wert $X$ kleiner oder gleich einem bestimmten Wert $x$ ist: $p(X \leq x)$ ; kurz für <i>cumulative distribution function</i> .
<b>Co-Referenz</b>	Vollständig verbundener Subgraph.
<b>Code</b>	Beziehung zwischen zwei Dokumenten, die von (mindestens) einem anderen Dokument referenziert werden.
<b>Dendrogramm</b>	Menge von Befehlen zur Ausführung durch einen Computer (Computerprogramm).
<b>Digraph</b>	Baumdarstellung des Ergebnisses eines hierarchischen Clusterings.
<b>Distanz</b>	Gerichteter Graph.
<b>Dreieck</b>	Die Länge des kürzesten Weges zwischen zwei Punkten in einem Graphen.
<b>Egonetzwerk</b>	Ein Zyklus der Länge 3 in einem Graphen.
<b>Eigenvektor</b>	Für ein Objekt aus einem bestimmten Netzwerk das Subnetzwerk, das auf das Objekt selbst ( <i>ego</i> ) und alle Objekte innerhalb einer bestimmten Maximalentfernung ( <i>alteri</i> ) eingeschränkt ist; bei einem Egonetzwerk mit Radius 1 das Subnetzwerk aus dem Objekt selbst und seinen Nachbarn.
	In Bezug auf eine Abbildung ein vom Nullvektor verschiedener Vektor, der durch die Abbildung nur skaliert wird, nicht aber seine Richtung ändert.
	<b>linker</b> Zeilenvektor $u$ , der für eine Matrix $\mathbf{M}$ und eine Konstante $\kappa$ die Gleichung $u \cdot \mathbf{M} = \kappa \cdot u$ erfüllt.
	<b>rechter</b> Spaltenvektor $v$ , der für eine Matrix $\mathbf{M}$ und eine Konstante $\lambda$ die Gleichung $\mathbf{M} \cdot v = \lambda \cdot v$ erfüllt.
<b>Eigenwert</b>	In Bezug auf einen Eigenvektor $u$ bzw. $v$ der Faktor, um den dieser durch die zu ihm gehörige Abbildung skaliert wird; Konstante $\kappa$ bzw. $\lambda$ , die für eine Matrix $\mathbf{M}$ die Gleichung $u \cdot \mathbf{M} = \kappa \cdot u$ bzw. $\mathbf{M} \cdot v = \lambda \cdot v$ erfüllt.

	<b>dominanter</b>	Betragsmäßig größter Eigenwert. Mitglied einer Menge; Oberbegriff für Objekte und Beziehungen in einem Netzwerk.
<b>Element</b>		Harmonisches Mittel von Genauigkeit und Trefferquote; $F = \frac{2 \cdot \text{Genauigkeit} \cdot \text{Trefferquote}}{\text{Genauigkeit} + \text{Trefferquote}}$ ; Synonym: $F_1$ .
<b>F-Maß</b>		Für eine Suchaufgabe ein gesuchtes Ergebnis, das nicht gefunden wurde; kurz: FN.
<i>false negative</i>		Für eine Suchaufgabe ein gefundenes Ergebnis, das nicht gesucht wurde; kurz: FP.
<i>false positive</i>		Vereinigung von $k$ -Cliques, die durch $k - 1$ Punkte miteinander verbunden sind.
<b>Gemeinschaft von <math>k</math>-Cliques</b>		Für eine Suchaufgabe der Anteil der <i>true positives</i> an der Menge aller <i>gefundenen</i> Ergebnisse: $\frac{ TP }{ \text{gefunden} }$ .
<b>Genauigkeit</b>		Möglichkeit der Durchschnittsberechnung in der Zeitreihenanalyse; hier: zu einem Zeitpunkt $t$ das arithmetische Mittel der Werte im Zeitfenster $[t - k, t + k]$ : $y_t^* = \frac{1}{2k+1} \cdot \sum_{i=-k}^k y_{t+i}$ .
<b>Gleitender Durchschnitt</b>		Anzahl der Linien, die an einen Punkt im Graphen grenzen. Zusammenfassung einer Menge von Punkten $V$ und einer Menge von Linien oder Pfeilen $E$ zu einer Einheit, die unter anderem der mathematischen Repräsentation eines Netzwerks dient: $G = (V, E)$ .
<b>Grad</b>	<b>azyklisch</b>	Ohne Zyklen.
<b>Graph</b>	<b>bipartit</b>	Die Punktmenge lässt sich so in zwei disjunkte Mengen einteilen, dass alle Linien (Pfeile) im Graphen nur zwischen diesen Mengen verlaufen.
	<b>gerichtet</b>	Mit Pfeilen.
	<b>gewichtet</b>	Linien oder Pfeilen können von 1 ver- schiedene Gewichte zugeordnet werden.
	<b>schlicht</b>	Ohne Pfeile, Mehrfachkanten, Kantengewichte oder Schleifen.
	<b>signiert</b>	Linien oder Pfeilen können Vorzeichen zugeordnet werden.
	<b>verbunden</b>	Alle Punkte sind voneinander erreichbar.
	<b>vollständig</b>	Alle möglichen Linien sind vorhanden; d.h. jeder Punkt ist mit jedem anderen Punkt verbunden.
<b>Graphenensemble</b>		Menge der Graphen, die durch ein bestimmtes Modell erzeugt werden können, das stochastische Elemente enthält.
<b>Grundgesamtheit</b>		Menge aller Untersuchungseinheiten.
<b>Hamming-Distanz</b>		Für zwei Untersuchungseinheiten der Anteil ihrer Merkmale, hinsichtlich derer sie nicht übereinstimmen.
<b>Harmonisches Mittel</b>		Wenn alle Werte mit gleichem Gewicht in das Mittel eingehen sollen, die Anzahl der Beobachtungswerte, geteilt durch die Summe ihrer Kehrwerte: $\bar{x}_H = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$ .
<b>Hierarchisches Clustering</b>		Methode zur Datenanalyse, die eine Menge von Datenpunkten so in disjunkte Teilmengen einteilt, dass Elemente in derselben Teilmenge einander relativ

	ähnlich sind, und dabei eine Hierarchie von Zerlegungen mit unterschiedlich strengen Ähnlichkeitsanforderungen produziert.
<b>Histogramm</b>	Visuelle Darstellung der Häufigkeitsverteilung eines metrischen Merkmals, in der die Ausprägungen auf der x- und die Häufigkeiten auf der y-Achse abgetragen sind.
<b>Hypergraph</b>	Graph mit Hyperkanten.
<b>Hyperkante</b>	Linie, die mehr als zwei Endpunkte hat.
<b>Häufigkeitsverteilung</b>	Funktion, die für eine Menge von Beobachtungen angibt, wie häufig die im Einzelnen beobachteten Werte vorkommen.
	<b>absolute</b> Häufigkeiten sind als Anzahlen angegeben.
	<b>kumulative</b> Häufigkeiten werden aufsummiert.
	<b>relative</b> Häufigkeiten sind als Anteile angegeben.
<b>Implementierung</b>	Umsetzung eines Algorithmus in einem Computerprogramm.
<b>Instanz</b>	Konkrete Realisierung eines Modells.
<b>Interquartilsabstand</b>	Differenz zwischen oberem und unterem Quartil.
<b>k-Clique</b>	Clique mit $k$ Punkten.
<b>Kardinalität</b>	Die Anzahl der Elemente in einer Menge.
<b>Kartesisches Produkt</b>	Für zwei Mengen $A$ und $B$ die Menge aller Paare $(a, b)$ , sodass das erste Element in $A$ und das zweite Element in $B$ ist (verallgemeinerbar auf mehr als zwei Mengen): $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$ .
<b>Komponente</b>	Maximaler verbundener Subgraph.
<b>Kontingenztafel</b>	Tabellarische Darstellung der gemeinsamen Merkmalsausprägungen zweier diskreter Merkmale (zweidimensionale Verteilung); im Englischen <i>contingency table</i> .
<b>Korrelationskoeffizient</b>	Maß für den Zusammenhang zweier metrischer Merkmale, das die Differenzen zwischen den Beobachtungen zweier Merkmale und deren arithmetischen Mitteln zueinander in Beziehung setzt (in der Version von <i>Bravais</i> und <i>Pearson</i> ): $r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$ Vorgabe für die Anordnung der Punkte und Linien (Pfeile) eines Graphendiagramms im zwei- oder dreidimensionalen Raum.
<b>Layout</b>	
<b>Linie</b>	Element eines Graphen, das eine symmetrische Beziehung darstellt; Synonym: Kante; im Englischen <i>edge</i> , <i>line</i> , <i>link</i> , <i>arc</i> .
<b>Markup</b>	Annotationen in einem Dokument, die syntaktisch vom Text unterscheidbar sind.
<b>Matrix</b>	Tabelle, mit der man rechnen kann; eine $n \times m$ -Matrix $\mathbf{A}$ hat $n$ Zeilen, $m$ Spalten und $n \cdot m$ Felder, die mit doppelten Indizes angegeben werden, z.B. $a_{ij}$ für das Feld in Reihe $i$ , Spalte $j$ .
<b>maximal</b>	Ein mathematisches Objekt ist in Bezug auf eine bestimmte Eigenschaft maximal, wenn man es nicht erweitern kann, ohne dass es diese Eigenschaft verliert.
<b>Median</b>	Für eine ungerade Zahl von Beobachtungen der Wert, der in der Mitte der nach der Größe geordneten Beobachtungswerte liegt; für eine gerade Zahl von Beobachtungen bei einem metrischen Merkmal das arithmetische Mittel der zwei Werte, die in der Mitte der nach der Größe geordneten Beobachtungswerte liegen; allgemein das 0,5-Quantil.

<b>Mehrfachkante</b>	Linie (Pfeil), die (der) in einem Graphen mehrfach vorkommt; Graphen mit Mehrfachkanten heißen Multigraphen; Synonym: Multikante.
<b>Menge</b>	Mathematische Struktur, die Elemente ohne Rücksicht auf eine Reihenfolge zusammenfasst und keine Duplikate enthalten kann.
	<b>abzählbar</b> Menge, deren Elemente sich mit den natürlichen Zahlen indexiert aufzählen lassen.
	<b>disjunkt</b> Mehrere Mengen sind disjunkt, wenn es kein Element gibt, das ihnen allen angehört, d.h. wenn ihre Schnittmenge leer ist.
	<b>überabzählbar</b> Menge, deren Elemente sich <i>nicht</i> mit den natürlichen Zahlen indexiert aufzählen lassen.
<b>Merkmal</b>	Eigenschaft einer Untersuchungseinheit.
	<b>binär</b> Merkmal, das nur zwei verschiedene Ausprägungen annehmen kann.
	<b>diskret</b> Merkmal, das nur abzählbar viele Ausprägungen annehmen kann.
	<b>metrisch</b> Merkmal, dessen Ausprägungen sich ordnen lassen, wobei auch die Abstände zwischen den Ausprägungen interpretierbar sind.
	<b>nominal</b> Merkmal, dessen Ausprägungen sich nur auf (Un-)Gleichheit prüfen lassen, ohne dass die Bildung einer Reihenfolge möglich wäre.
	<b>ordinal</b> Merkmal, dessen Ausprägungen in eine Reihenfolge gebracht werden können, ohne dass sich die Abstände zwischen den Ausprägungen interpretieren ließen.
	<b>qualitativ</b> Merkmal, bei dem die artmäßige Charakterisierung im Vordergrund steht.
	<b>quantitativ</b> Merkmal, bei dem die zahlenmäßige Charakterisierung im Vordergrund steht.
	<b>quasi-stetig</b> Diskretes Merkmal, das so viele verschiedene Ausprägungen annehmen kann, dass es sich als stetig behandeln lässt.
	<b>stetig</b> Merkmal, das überabzählbar viele Ausprägungen annehmen kann.
<b>Merkmalsausprägung</b>	Konkreter Wert eines Merkmals für eine Untersuchungseinheit.
<b>Metadaten</b>	Daten, die weitere Informationen über andere Daten enthalten, z.B. Informationen zu Autor und Entstehungsjahr eines Textdokuments.
<b>Modus</b>	Hinsichtlich einer Menge von Beobachtungswerten der Wert, der am häufigsten beobachtet wird.

<b>Motiv</b>	Kleiner Subgraph, der in einem Graphen statistisch auffällig häufig vorkommt.
<b>Multidimensionale Skalierung</b>	Methode zur Visualisierung von Daten, die Ähnlichkeiten zwischen Datenpunkten in Distanzen für einen zwei- oder dreidimensionalen Raum übersetzt.
<b>Multigraph</b>	Graph mit Mehrfachkanten.
<b>Multimenge</b>	Menge, in der Duplikate vorkommen können.
<b>Nachbar</b>	Hinsichtlich eines bestimmten Punktes ein anderer Punkt, der zu diesem adjazent ist.
<b>Nachbarschaft</b>	Hinsichtlich eines bestimmten Punktes die Menge aller Punkte, die zu ihm adjazent sind.
<b>Netzwerk</b>	Zusammenfassung einer Menge von Objekten und einer Menge von Beziehungen zu einer Einheit; abstrakte Repräsentation eines Phänomens in der Realität, die mathematisch als Graph dargestellt werden kann.
	<b>mehrschichtig</b> Netzwerk mit mehreren Mengen von Objekten und mehreren Mengen von Beziehungen.
	<b>multiplex</b> Netzwerk mit einer Menge von Objekten und mehreren Mengen von Beziehungen.
<b>Objekt</b>	Element eines Netzwerks, das eine abgrenzbare Einheit in der Realität kennzeichnet, die mit anderen Einheiten in Beziehung stehen kann; im Englischen <i>entity</i> .
<b>OCR</b>	Texterkennungsverfahren, kurz für <i>Optical Character Recognition</i> .
<b>Paar</b>	Tupel mit zwei Elementen.
<b>Parameter</b>	Stellschraube, an der man ein Modell einstellen kann; Variablen, deren Werte für eine bestimmte Instanz eines Modells konstant sind.
<b>PDF</b>	Funktion, die für eine Verteilung die Wahrscheinlichkeit eines bestimmten Wertes $X$ anzeigt: $p(X)$ ; kurz für <i>probability density function</i> .
<b>Permutation</b>	Reihenfolge für Elemente in einer Menge.
<b>Pfeil</b>	Element eines Graphen, das eine asymmetrische Beziehung darstellt; Synonym: gerichtete Kante; im Englischen <i>directed line, directed link, directed edge</i> .
<b>Projektion</b>	Abbildung von einer Menge in eine ihrer Teilmengen, die auch mehrfach angewendet werden kann, ohne dass sich das Ergebnis verändert; hier vor allem die Abbildung eines bipartiten Graphen auf einen seiner Teile (d.h. eine seiner <i>partes</i> ).
<b>Punkt</b>	Element eines Graphen, das ein Objekt darstellt; Synonym: Knoten; im Englischen <i>point, node, vertex</i> .
	<b>adjazent</b> Zwei Punkte sind adjazent, wenn sie über eine Linie (bzw. einen Pfeil) direkt miteinander verbunden sind.
	<b>erreichbar</b> Ein Punkt ist von einem anderen Punkt aus erreichbar, wenn es einen Weg zwischen den Punkten gibt.

<b>Quantil</b>	Für einen Anteil $p$ derjenige von $n$ Beobachtungswerten, für den höchstens $n \cdot p$ Werte kleiner oder gleich und $n \cdot (1 - p)$ Werte größer oder gleich sind.
<b>Quartil</b>	0,25-Quantil (unteres Quartil), 0,5-Quantil (Median) oder 0,75-Quantil (oberes Quartil).
<b>regulärer Ausdruck</b>	Zeichenkette, mit der sich Mengen anderer Zeichenketten abstrakt repräsentieren lassen.
<b>Relation</b>	Teilmenge des Kartesischen Produkts zweier Mengen $A$ und $B$ (binäre Relation) zur Darstellung von Beziehungen (verallgemeinerbar auf mehr als zwei Mengen): $R \subseteq A \times B$ und $aRb$ für $(a, b) \in R$ ; falls $A = B$ spricht man von einer Relation <i>auf</i> der Menge $A$ . <b>antisymmetrisch</b> Relation auf einer Menge $A$ , für die gilt: $\forall a, b \in A$ mit $a \neq b$ : $aRb \Rightarrow \neg bRa$ . <b>reflexiv</b> Relation auf einer Menge $A$ , für die gilt: $\forall a \in A$ : $aRa$ . <b>symmetrisch</b> Relation auf einer Menge $A$ , für die gilt: $\forall a, b \in A$ : $aRb \Rightarrow bRa$ . <b>transitiv</b> Relation auf einer Menge $A$ , für die gilt: $\forall a, b, c \in A$ : $aRb \wedge bRc \Rightarrow aRc$ .
<b>Rohdaten</b>	Daten in der Form, in der sie unmittelbar nach ihrer Erhebung vorliegen.
<b>Schleife</b>	Linie oder Pfeil, für den Ausgangs- und Endpunkt identisch sind; Zyklus der Länge 1; im Englischen <i>loop</i> .
<b>seed</b>	Zahl, mit der ein Zufallszahlengenerator initialisiert wird, damit dieser bei jedem Aufruf die gleichen Zahlen generiert.
<b>Soziogramm</b>	Graphische Darstellung eines sozialen Netzwerks, insbesondere in der Tradition von <i>Moreno</i> .
<b>Standardabweichung</b>	Positive Wurzel aus der Varianz: $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ .
<b>Stichprobe</b>	Teil der Grundgesamtheit, der in einer statistischen Erhebung betrachtet wird.
<b>Streudiagramm</b>	Graphische Darstellung der gemeinsamen Verteilung zweier Merkmale (zweidimensionale Verteilung) als Punkte in einem Koordinatensystem mit x-Achse und y-Achse; im Englischen <i>scatter plot</i> .
<b>Subgraph</b>	Teil eines Graphen.
<b>Subnetzwerk</b>	Teil eines Netzwerks.
<b>Teilmenge</b>	Hinsichtlich einer bestimmten Ausgangsmenge eine andere Menge, die nur Elemente der Ausgangsmenge enthält.
<b>Text Mining</b>	Computergestützte Informationsextraktion aus Texten.
<b>Trefferquote</b>	Für eine Suchaufgabe der Anteil der <i>true positives</i> an der Menge aller <i>gesuchten</i> Ergebnisse: $\frac{ TP }{ \text{gesucht} }$ .
<b>Tripel</b>	Subgraph mit drei Punkten.
<b>true positive</b>	Für eine Suchaufgabe ein gefundenes Ergebnis, das gesucht wurde; kurz: TP.
<b>Tupel</b>	Zusammenfassung von Werten so, dass diese eine Reihenfolge haben, sodass z.B. $(a, b) \neq (b, a)$ .
<b>Überdeckung</b>	Einteilung einer Menge (z.B. der Punktmenge eines Graphen) in Teilmengen so, dass jedes Element mindestens einer

<b>Untersuchungseinheit</b>	Menge angehört; im Englischen <i>cover</i> .
<b>URL</b>	Objekt, auf das sich eine statistische Untersuchung bezieht. Bezeichner zur eindeutigen Identifikation und Lokalisierung von Ressourcen, insbesondere in Computernetzwerken; kurz für <i>Uniform Resource Locator</i> .
<b>Varianz</b>	Mittlere quadratische Abweichung vom arithmetischen Mittel: $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ . <b>empirische</b> Varianz, die für eine Stichprobe ermittelt wird (Division durch $n - 1$ statt durch $n$ ); Synonym: Stichprobenvarianz; $\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ .
<b>Wahrscheinlichkeitsverteilung</b>	Funktion, die jeder möglichen Ausprägung einer Zufallsvariablen ihre Wahrscheinlichkeit zuordnet.
<b>Webseite</b>	Einzelnes Dokument, das über das <i>World Wide Web</i> abrufbar und anhand seiner URL eindeutig identifizierbar ist; im Englischen <i>web page</i> .
<b>Website</b>	Menge von Webseiten, die unter einer Domäne abrufbar sind; im Englischen <i>website</i> .
<b>Weg</b>	Folge von Linien oder Pfeilen zwischen zwei Punkten; Synonym z.T.: Pfad; im Englischen <i>walk</i> oder <i>path</i> . <b>kürzester</b> Weg zwischen zwei Punkten mit minimaler Länge; Synonym z.T.: Pfad; im Englischen <i>path</i> , <i>shortest path</i> oder <i>geodesic</i> . <b>Länge</b> Anzahl der Linien- bzw. Pfeilüberquerungen auf einem Weg.
<b>XML</b>	Markup-Sprache, die Kodierungsregeln für ein Dateiformat festlegt, das für Menschen und Maschinen lesbar ist; kurz für <i>eXtensible Markup Language</i> .
<b>Zerlegung</b>	Einteilung einer Menge (z.B. der Punktmenge eines Graphen) in disjunkte Teilmengen; im Englischen <i>partition</i> .
<b>Zufallslauf</b>	Hier: Bewegung über die Linien bzw. Pfeile eines Graphen, bei welcher der nächste Schritt auf Basis einer Wahrscheinlichkeitsverteilung aus den möglichen Schritten ausgewählt wird; im Englischen <i>random walk</i> .
<b>Zufallsvariable</b>	Merkmal, dessen Ausprägungen als Ergebnis eines Zufallsvorgangs aufgefasst werden.
<b>Zyklus</b>	Weg in einem Graphen, für den Startpunkt und Endpunkt identisch sind.



# Sachregister

- Abdeckungsfehler, 85
- Abstraktion, 2, 9
- abzählbar, 96
- Adjazenzmatrix, 118, 125
- Ähnlichkeit
  - als Beziehung, 50
- Äquivalenz
  - Äquivalenzklasse, 161
  - Äquivalenzrelation, 161
  - automorphe, 160
  - reguläre, 160
  - stochastische, 161
  - strukturelle, 160
- Algorithmus
  - Gemeinschaftserkennung, 143, 151, 289
  - *greedy*, 158
  - Layout-Algorithmus, 110
  - Such-Algorithmus, 214
- Ansteckungseffekt, 24
- Anwälte
  - Netzwerkstudien, 18
- Application Programming Interface (API)*,  
*siehe* Programmierschnittstelle
- Arithmetisches Mittel, 95
- Artificial Intelligence (AI)*, 69
- Aspect Ratio*, 109
- Aufbereitungsfehler, 87
- Aufzeichnungsfehler, 86
- Ausgangsgrad, *siehe* Außengrad
- Ausgangskomponente, 168
- Außengrad, 131, 185
  - Außengradverteilung, 175, 264
  - temporaler, 192
- Balkendiagramm, 99
- Baum (Graphentheorie), 8, 205
- Bellzahl, 152
- Beobachterperspektive, 129
- Beobachtung (Statistik), 93
- Berufsfreiheit
  - BVerfGE-Entscheidungen, 292
- Beziehung (Netzwerkforschung), 5
  - antisymmetrische, 49
  - Beziehungsdichte, 145
  - Beziehungsform, 51
  - Beziehungsgrad, 38
  - Beziehungsprofil, 39
  - symmetrische, 49
- Bibliographische Kopplung, 47
- Bibliometrie, 36
- Binomialkoeffizient, 148, 197, 245
- Binomialverteilung, 198
- Boxplot*, 261
- Breitensuche (BFS), 214, 215
- Bürgerliches Gesetzbuch (BGB)
  - Baumdarstellung, 7
- Bundesverfassungsgericht
  - Geschäftsordnung (BVerfGG), 231
- Burst*-Fehler, 246
- BVerfGE, 230
  - Daten, 250
  - Entscheidungssammlung, 231
  - Fundstelle, 233
    - Fundstellenliste, 238
    - Kurzfundstelle, 233
    - Langfundstelle, 233
  - Korpus, 239
  - Metadaten, 238
  - Überblick, 254, 256, 257
  - Zitate, 232
  - Zitatliste, 243
  - Zitiergraph, 243
- Cascading Failure*, 24
- Charakterisierungsfrage, 73, 165, 252
- Christentum
  - BVerfGE-Entscheidungen, 307, 310
- Civil Law*, 33, 63, 211
- Clique, 148
  - *k*-Clique, 148, 282
- Cluster, *siehe* Gemeinschaft (Netzwerkforschung)

- Clustering-Koeffizient, 171
- Co-Evolution, 216
- Co-Okkurrenz, 48
- Co-Produktion, 44
- Co-Reaktion, 46
- Co-Referenz, 46
- Co-Sponsorship*, 19
- Co-Status, 48
- Co-Zitation, 190
- Code, 79, 88, 318
- Common Law*, 32, 63, 211
- Complementary Cumulative Distribution Function* (CCDF), 178, 265
- Corporate Governance*, 23
- Corpus Iuris Canonici*, 6
- Court Opinion*, 17
- CourtListener*, 81
- Crawler*, 212
- Curve Fitting*, 176
- Cut-Off*, 118
  
- Data Mining*, 50, 69
- Data Science*, 69
- Daten, 74
  - fiktive, 88
  - juristische, 74, 228
  - Modellierung, 90
  - Quellen, 75, 317
  - reale, 74
  - relationale, 109
  - Rohdaten, 103
  - Sammlung, 75, 317
  - semistrukturierte, 81
  - strukturierte, 81, 103
  - unstrukturierte, 81
- Datenbank
  - Datenbankrichtlinie, 83
  - Hersteller, 83
  - juristische, 75, 78, 83, 128
  - Urheberrecht, 82
- Datenschutz
  - BVerfGE-Entscheidungen, 280, 304, 305
- Definitionsproblem, 129
- Dendrogramm, 154
- Deutschsprachiges Fallrecht (DFR), 248
- Dichtekurve, 176
- Diffusion, 217
  - als Beziehung, 41
  - Diffusionsprozess, 216
- Dissensmatrix, 296
- Document Type Definition*, 82
- Dreieck (Graphentheorie), 171
- Durchschnitt, 95
  - gleitender, 188
- Egonetzwerk, 60
- Eigentum
  - BVerfGE-Entscheidungen, 292
- Eigenvektor, 136
  - Eigenvektorzentralität, 135
- Eigenwert, 136
- Einbettung, 145
- Eingangsgrad, *siehe* Innengrad
- Eingangskomponente, 168
- Empirical Legal Studies*, 12
- Entscheidungssammlung, 76
- Euler, Leonhard*, 1, 4
- Europa
  - BVerfGE-Entscheidungen, 293, 302, 303
- Evaluation
  - als Beziehung, 43
- Explikation, 65
- Exploration, 62
- Exponentialverteilung, 180
- Extraktionsfehler, 87
- Extrapolation, 68
  
- F-Maß, 244
- False Negative*, 242
- False Positive*, 242
- Finanzmarkt
  - Netzwerkstudien, 24
  - Regulierung, 25
- First In, First Out* (FIFO), 214
- Fitnessmodell, 210
- Flatland*, 88, 327
- Fraktal, 180
  
- Gemeinschaft (Netzwerkforschung), 146
  - Gemeinschaftserkennung, 147, 289
  - Gemeinschaftsgraph, 295, 297
  - Gemeinschaftsstruktur, 147
    - hierarchische, 151
    - überlappende, 150
- Genauigkeit (Suche), 244
- Gerichtsentcheidung
  - als Objekt, 32
  - Netzwerkstudien, 13
  - Urheberrecht, 82

- Gesetzestext
  - als Objekt, 33
  - Netzwerkstudien, 15
- Gesetzgebungsorgane
  - Netzwerkstudien, 19
- Gittergraph, 199
- Gleichverteilung, 199
- Grad, 131
  - Gradassortativität, 185, 186
  - Gradsequenz, 186
  - Gradverteilung, 175
  - Gradzentralität, 131, 132
- Graph, 2
  - azyklischer, 8
  - bipartiter, 21, 54, 92
  - Digraph, 55
  - Multigraph, 55
  - regulärer, 200
  - schlichter, 55
  - Subgraph, 128
    - textbasierter, 301
  - unverbundener, 124
  - verbundener, 8, 124
    - schwach, 124
    - stark, 124
  - vollständiger, 148
- Graphdiagramm, 109
- Graphenensemble, 197
- Graphentheorie, 4, 17
- Graphisches Symbol, 104
- Graphmatrix, 118
- Grundgesamtheit, 93, 245
- Grundgesetz (GG)
  - Baumdarstellung, 8
- Grundwahrheit, 157
- Gruppierungsfrage, 73, 142, 282
  
- Häufigkeitsverteilung, 98, 174
  - absolute, 98, 174
  - kumulative, 106, 174
  - relative, 98, 174
- Hamming-Distanz, 296
- Heterophilie, 183
- Hierarchisches Clustering, 153
- Histogramm, 106, 265
- Homophilie, 183
- Hubs & Authorities*, 138
- Hyperkante, 54
  
- igraph*, 134, 154, 158, 330
  
- Implementierung, 112
- Individuum
  - als Objekt, 29
- Informationsextraktion, 81
- Informationsweiterverwendungsgesetz (IWG), 79, 83
- Inklusionsproblem, 57
- Innengrad, 131, 185, 270
  - Innengradverteilung, 175, 264
  - temporaler, 192, 273
- Interdisziplinarität, 12
- Interlocking Directorate*, 22
- Interpretationsproblem, 114
- Interquartilsabstand, 262
- Islam
  - BVerfGE-Entscheidungen, 308, 312
  
- Judentum
  - BVerfGE-Entscheidungen, 307, 311
- Jupyter Notebook, 330
- juris*, 82, 236, 241
  
- Kante (Graphentheorie), *siehe* Linie (Graphentheorie)
  - gewichtete, 52
- Kardinalskala, 94
- Kartesisches Produkt, 49
- Kleine-Welt-Effekt, 168
- Kleine-Welt-Modell, 200
- Knoten (Graphentheorie), *siehe* Punkt (Graphentheorie)
- Königsberger Brückenproblem, 1, 2, 119, 123, 125
- Kohäsion, 145
- Kollektiv
  - als Objekt, 31
- Kombinatorische Explosion, 152
- Kommunale Selbstverwaltung
  - BVerfGE-Entscheidungen, 286
- Kompartimentmodell, 217
- Komponente, 149
  - größte, 167
  - Riesenkomponekte, 203
- Konfigurationsmodell, 201
- Konsensgemeinschaft, 299
- Konsensgraph, 299
- Kontingenztafel, 181
- Kopiermodell, 209
- Korrelationskoeffizient, 184

- Kovarianz, 184
- Kriminalität
  - Netzwerkstudien, 25
- Kürzeste-Wege-Baum, 127
- Kumulationsmodell, 206
- Kumulativer Vorteil, 206
  
- Label-Propagierung, 153
- Lagemaß, 95
- Last In, First Out* (LIFO), 214
- Layout, 110
- Linie (Graphentheorie), 3
  - gerichtete, *siehe* Pfeil (Graphentheorie)
- Liniendiagramm, 99
- Literatur
  - als Objekt, 36
- Logarithmic Binning*, 179
  
- Machine Learning* (ML), 69
- Makroebene, 73, 252
- Maß, 129
  - mediales, 141
  - radiales, 141
- Mathematik, 9
- Matrix, 118
  - quadratische, 136
- Matthäus-Effekt, 207
- maximal (Eigenschaft), 148
- Median, 95
- Mehrfachkante, 52
- Meinungsfreiheit
  - BVerfGE-Entscheidungen, 280, 293
- Menge (Mathematik), 3
- Merkmal, 94
  - Ausprägung, 94
  - diskretes, 95
  - qualitatives, 95
  - quantitatives, 95
  - quasi-stetiges, 96
  - Skala, 94
  - stetiges, 96
  - Typ, 96
- Mesoebene, 73, 282
- Messproblem, 129
- Mikroebene, 72, 269
- Modell, 193
  - agentenbasiertes, 194, 211
  - algorithmisches, 193
  - mathematisches, 193
    - physisches, 193
- Modellierung, 193
- Modul, *siehe* Gemeinschaft (Netzwerkforschung)
- Modularität, 156
- Momentaufnahme, 187
- Moreno, Jacob Levy*, 12, 65
- Motiv, 173
- Multidimensionale Skalierung, 116
- Multikante, *siehe* Mehrfachkante
- Multimenge, 3
  
- n-Gram*, 190
- Nachbar (Graphentheorie), 118
- Näherungslösung, 152
- Nähezentralität, 140
- Natural Language Processing* (NLP), 37
- network*, 113, 196, 204, 209, 248, 330
- Netzwerk, 5
  - Definition, 28, 229
    - beziehungsgeleitete, 56
    - objektgeleitete, 55
  - Evolution, 187
  - Konstruktion, 235
  - mehrschichtiges, 224
  - multiplexes, 225
  - skalenfreies, 178
  - Subnetzwerk, 128
- Netzwerkforschung, 4
- Nominalskala, 94
- Normalverteilung, 176
- Nullmodell, 166
  
- Objekt (Netzwerkforschung), 5
- Öffentliches Interesse, 64
- One-Mode Projection*, 21, 54
- Optical Character Recognition* (OCR), 79
- Optimierungsproblem, 151
- Ordinalskala, 94
  
- PageRank*, 137
- Patent
  - als Objekt, 34
  - Netzwerkstudien, 26
- Patentfamilie, 26
- PatentsView*, 81
- Personalgraph, 93
- Personalverflechtungen, 22
- Pfeil (Graphentheorie), 17
- Phasenübergang, 202

- Poissonverteilung, 198
- Position (Netzwerkforschung), 146, 161
- Positionierung, 110
  - distanzbasierte, 116
  - kräftebasierte, 112
  - merkmalsbasierte, 111
- Positionszentralität, 139
- Potenzgesetz, 178, 265
- Präferenzanbindung, 207
- Prior Art*, 26
- Probability Density Function* (PDF), 178, 265
- Produktion
  - als Beziehung, 40
- Programmierschnittstelle, 81, 317
- Punkt (Graphentheorie), 3
  - adjazenter, 118
  - erreichbarer, 124
- Python, 241, 330
  
- Quantifizierung, 122
- Quantil, 95
- Quartil, 95, 262
  
- Randhäufigkeit, 181
- Reaktion
  - als Beziehung, 41
- Recherchegraph, 213
- Rechtsbegriff
  - als Objekt, 36
- Rechtsdogmatik, 63, 71–73
- Rechtsgeschichte, 63, 67, 70
- Rechtswissenschaften, 12, 82
- Rechtsphilosophie, 71
- Rechtspolitik, 67
- Rechtssprechungsziernetzwerk, 78, 187, 209
- Rechtssoziologie, 67, 70, 72, 73, 180
- Rechtsstrukturvergleiche, 326
- Rechtssubjekt
  - als Objekt, 29
- Rechtstext
  - als Objekt, 32
- Rechtstheorie, 9, 63, 67, 70, 72, 73, 180, 326
- Rechtsvergleiche, 63, 70, 72, 73, 180
  - intrasystemische, 181
- Rechtswissenschaftstheorie, 67
- Redundanz, 172
- Regulärer Ausdruck (Regex), 240, 301
- Regulierungstheorie, 63
- Rekursionsproblem, 59
- Relation (Mathematik), 49
- Religion
  - BVerfGE-Entscheidungen, 306, 309
- Reziprozität, 170
- Richter
  - Netzwerkstudien, 19
- Rolle (Netzwerkforschung), 146, 161
- RSS-Feed, 81
- Rückkopplungszentralität, 135
  
- Schleife (Graphentheorie), 53
- Schwellenwert, 117
- Selbstähnlichkeit, 180
- Selektionsproblem, 56
- SI-Modell, 217
- Simulation, 217
- Skalierungsproblem, 114
- Soziale Netzwerkanalyse, 12, 65
- Soziogramm, 65
- Soziomatrix, 118
- Spaltenvektor, 136
- Stabilitätsmotiv, 173
- Stammbaum, 6
- Standardabweichung, 95
- Statistik
  - bayesianische, 69
  - deskriptive, 62, 94
  - explorative, 62
  - frequentistische, 69
  - induktive, 69
- Statistisches Ensemble, 157, 197
- Status
  - als Beziehung, 43
- Stichprobe, 244
  - Stichprobenvarianz, 95
  - Zufallsstichprobe, 244
    - einfache, 245
- Stirlingzahl, 152
- Streudiagramm, 108
- Streuungsmaß, 95
- Stumpf (Graphentheorie), 201
- Suchbegriff, 86
- Suchprozess, 212
- Suchraum, 152
- Supreme Court Database*, 33
- System, komplexes adaptives, 9, 63, 181
- Systemisches Risiko, 25, 63, 221
- Szientometrie, 36, 42

- Teilnehmerperspektive, 129
- Text Mining*, 81
- Texterkennung, 79
- Tiefensuche (DFS), 214, 215
- Transaktion
  - als Beziehung, 40
- Transitivität, 171
- Trefferquote (Suche), 244
- Triadenzensus, 173
- Tupel, 3
  
- U.S. House of Representatives*, 19, 21
- U.S. Supreme Court*, 12, 13, 16
- überabzählbar, 96
- Überanpassung, 180
- Überdeckung, 152
- überzufällig, 172
- Unternehmensgraph, 93
- Unternehmensverflechtungen
  - Netzwerkstudien, 22
- Untersuchungseinheit, 93
  
- Variable, 29
  - Ausprägung, 29
  - binäre, 97
  - visuelle, 104, 106
- Varianz, 95
- Verdrahtungsmodell, 195, 196
- Verflechtungsgraph, 92
- Verhältnisskala, 94
- Vermischung
  - assortative, 183
  - disassortative, 183
- Verteilung
  - eindimensionale, 174, 263
  - gemeinsame, 181
  - schiefe, 177
  - Verteilungsfunktion, 178
  - zweidimensionale, 181, 267
- Vertrag
  - als Objekt, 35
- Verwaltungsentscheidung
  - als Objekt, 34
- Verweis, 15
  - als Beziehung, 42
- Visualisierung, 99
  - analytische, 100
  - illustrative, 100
  - Visualisierungsprozess, 103
- Visuelle Kodierung, 104
- Visuelle Struktur, 103
  
- Wachstumsmodell, 195, 204
- Wahlrecht
  - BVerfGE-Entscheidungen, 281, 284, 285
- Wahrscheinlichkeitsfunktion, 178
- Wahrscheinlichkeitsverteilung, 174
- Web Scraping*, 248
- Weg, 124
  - kürzester, 126
  - Länge, 124
- Wichtigkeit
  - abstrakte, 130, 269
  - konkrete, 130, 278
- Wichtigkeitsfrage, 72, 129, 269
  
- XML-Schema, 81
  
- Z-Score*, 186
- Zeilenvektor, 136
- Zeitreihe, 187
- Zentralität, 130
  - Zentralitätsmaß, 130
  - Zentralitätswert, 130
- Zentrum-Peripherie-Struktur, 186
- Zerlegung, 143
  - Abdeckung, 155
  - feinste, 153
  - gröbste, 153
- Zitat
  - als Beziehung, 42
  - binäre Zählung, 78, 192, 270, 274
  - gewichtete Zählung, 78, 192, 271, 275
  - Zählmodus, 78, 132, 272
  - Zitatextraktion, 14, 233
  - Zitationsanalyse, 14
  - Zitierblock, 77, 190
    - Zitierblockzeichen, 235
  - Zitierbreite, 132
  - Zitiergraph, 91
  - Zitierkurve, 192, 273, 276
  - Zitierprofil, 132, 133
  - Zitierrauschen, 163
  - Zitiertiefe, 132
  - Zitierverhältnis, 271
- Zufallsgraph, 197
- Zufallslauf, 125
- Zufallssurfer, 137
- Zufallsvariable, 174
- Zusammengehörigkeit, 145
- Zwischenzentralität, 140
- Zyklus (Graphentheorie), 8