

Betroffenenschutz durch Verfahren bei Big Data und KI

Begrenzung von Big Data-Anwendungen und KI
durch rechtliche Vorgaben

Bertram Raum

Bibliografische Information Der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.de> abrufbar.

I S B N 9 7 8 - 3 - 8 0 0 5 - 1 9 6 2 - 0

dfv Mediengruppe

© 2025 Deutscher Fachverlag GmbH, Fachmedien Recht und Wirtschaft, Mainzer Landstr. 251, 60326 Frankfurt am Main, buchverlag@ruw.de

www.ruw.de

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Druck: Beltz Grafische Betriebe GmbH, 99947 Bad Langensalza

Printed in Germany

Vorwort

Die vorliegende Monographie wurde im Sommersemester 2024 von der Juristischen Fakultät der Universität Passau als Dissertation angenommen. Literatur und Rechtsprechung wurden bis Juni 2024 berücksichtigt.

Zunächst möchte ich mich bei Prof. Dr. Kai von Lewinski bedanken, der mir die Anregung zum Thema meiner Dissertation und wichtige Hinweise bei ihrer Erstellung gab. Wichtige Anregungen habe ich auch aus seinen jährlich stattfindenden Doktorandenseminaren gewonnen. Herrn Prof. Dr. Brian Valerius danke ich für die Erstellung des Zweitgutachtens.

Ein besonderer Dank gilt meinem Sohn Philipp, ohne dessen Mithilfe bei der Beschaffung von Literatur die Dissertation am Ende nicht hätte so schnell fertiggestellt werden können. Zudem hat er insbesondere zur Organisation wichtige Hinweise gegeben. Manchmal hat er durch kleine Bemerkungen seinem verzweifelten Vater unbeabsichtigt wieder Mut gegeben. Der größte Dank gehört meiner Frau für das Korrekturlesen des Manuskripts und ihre große Geduld mit mir, besonders aber dafür, dass sie mich bedingungslos unterstützt hat.

Der Autor

MinR a.D. Dr. Bertram Raum ist Rechtsanwalt. Er berät internationale und nationale Unternehmen sowie Forschungsgremien in rechtlichen Fragen des Datenschutzrechts. Er ist zudem auf die Bereiche des Sozialdatenschutzes und des Datenschutzrechts im Gesundheitswesen spezialisiert.

Inhaltsverzeichnis

Vorwort	VII
Abkürzungsverzeichnis	XIX
A. Einleitung	1
B. Begriffsbestimmungen und das Phänomen »Big Data«	7
I. Was ist ein »Verfahren«?	7
II. »Big Data« ein schillernder Begriff – eine Annäherung	8
1. Begriffsbestimmung nach Gartner	10
a) Teilaspekt »Volume«	11
b) Teilaspekt »Variety«	12
c) Teilaspekt »Velocity«	14
d) Ergänzungen der Begriffsbestimmung durch weitere »Vs«	15
e) Erweiterung um den Teilaspekt »Analytics«	17
2. Definitionen des Begriffs »Big Data« aus Deutschland	18
3. Arbeitsdefinition des Begriffs »Big Data«	19
III. »Big Data« und »Künstliche Intelligenz« (KI)	21
1. Der Begriff »Künstliche Intelligenz« (KI)	22
2. »Big Data« bei Trainingsdaten	27
3. Gefahr von Diskriminierungen bei Nutzung von Trainingsdaten	30
4. Gefahren bei nicht-personenbezogenen Daten	33
IV. Abgrenzung zu »Data Warehouse« und »Data Mining«	33
V. Quellen für Big Data-Anwendungen	35
1. Verwaltungsinterne und weitere öffentlich-rechtliche Datenquellen	35
2. Daten aus privatrechtlichen Datenbanken und Registern	37
3. Daten aus »Open Data«	38
4. Daten aus »Social Media«	40
5. »Ubiquitous Computing«	41
VI. Ausgewählte Big Data-Anwendungen	43
1. »Online-Tracking«	46
2. »Profiling«	47
a) »Scoring«	48
b) »Social Scoring«	49
3. »Behavioral Targeting«	50
4. »Web Scraping«	52
a) Zugang über offene und nicht-öffentliche Schnitt- stellen	53
b) »Screen Scraping«	54
	IX

Inhaltsverzeichnis

5. »Data Harvesting«	58
6. Zwischenergebnis	58
C. Big Data-Anwendungen einschließlich KI und gesetzliche Regelungen	60
I. Unionsrechtliche Regelungen zu Big Data und KI	60
1. »Big Data« und die Datenschutz-Grundverordnung (DSGVO)	60
a) Personenbezogene Daten im Sinne der DSGVO	62
b) »Pseudonymisierung« personenbezogener Daten und Big Data-Anwendungen	64
c) Verschlüsselung von Daten	70
d) Nicht-personenbezogene Daten und Big Data-Anwendungen	70
aa) »Anonymisierung« personenbezogener Daten und Big Data-Anwendungen	70
bb) Anwendung der DSGVO auf reine Sachdaten?	81
e) Zwischenergebnis	82
2. Rechtsakte der EU im Zusammenhang mit der Datenstrategie 2020	83
a) Der Data Act	84
b) Der Digital Markets Act – DMA	85
c) Der Data Governance Act (DGA)	86
d) Der Digital Service Act – DSA	87
e) Die KI-VO	89
f) Auswirkungen auf »Big Data«	98
3. Weitere rechtliche Vorgaben in Verordnungen und Richtlinien der EU	101
a) Der Vorschlag für eine ePrivacy-Verordnung	102
b) Die Free Flow of Data-VO (FFD-VO) der EU	103
c) Die »Open Data-Richtlinie« der EU	104
d) Die EU-VO über Transparenz und das Targeting politischer Werbung	106
II. Beispiele für nationale und internationale Big Data-Anwendungen	108
1. Einsatz von »Big Data«-Anwendungen in der Privatwirtschaft	108
a) Vorgaben zu »Big Data« aus dem allgemeinen Zivilrecht	109
b) Nutzung von Big Data-Anwendungen im Personalwesen (»People Analytics«)	114
c) Schutz vor diskriminierender Wirkung – das AGG	126

d)	Regelungen zu »Big Data« im Finanz- und Versicherungswesen	127
aa)	Die BaFin-Prinzipien zum Einsatz von Algorithmen in Entscheidungsprozessen	128
bb)	»Big Data« und KI-Systeme im Bankensektor	129
cc)	Big Data-Scoring im Bereich der Kreditauskunft	131
dd)	»Big Data«-Anwendungen im Hochfrequenzhandel	139
ee)	Regelungen zum »Big Data« im Versicherungssektor	141
ff)	Fazit zur Nutzung von Big Data-Anwendungen und KI-Systemen im Finanz- und Versicherungswesen	143
e)	Big Data-Anwendungen zur Entwicklung von Marketing- und Werbestrategien	143
f)	Big Data-Anwendungen in weiteren Wirtschaftsbereichen	147
2.	»Big Data« im öffentlichen Recht	150
a)	»Big Data« im allgemeinen Verwaltungsrecht	150
aa)	Einschränkungen des Erlasses vollständig automatisierter VA	152
bb)	Die §§ 35a VwVfG, 31 SGB X, 155 Abs. 4 AO und Big Data	154
cc)	Verstoß der §§ 35a VwVfG, 31a SGB X und 155 Abs. 4 AO gegen Art. 22 DSGVO?	159
dd)	Gibt es ein »Recht auf eine menschliche Entscheidung«?	161
ee)	Der vollständig automatisierte VA im Sonderverfahrensrecht	164
α)	Der vollständig automatisierte VA nach § 155a Abs. 4 AO	164
β)	Der vollständig automatisierte VA nach § 31a SGB X	165
ff)	Zwischenfazit	166
b)	»Big Data« in der amtlichen Statistik	166
c)	»Big Data« im Steuerrecht	170
d)	Big Data-Regelungen im Bereich der Sicherheitsbehörden	172
aa)	»Data Mining« aufgrund des Antiterrordateigesetzes (ATDG) und des Rechtsextremismusedateigesetzes (REDG)	174

Inhaltsverzeichnis

bb)	»Predictive Policing« und »Predictive Profiling« im deutschen Polizeirecht	176
a)	Nutzung der Software »PreCobs« durch deut- sche Polizeibehörden	180
b)	Nutzung ähnlicher Software-Lösungen in anderen Bundesländern	181
c)	Nutzung der Software »Palantir Gotham« durch deutsche Polizeibehörden	181
d)	Erfahrungen mit der Nutzung von Big Data- Anwendung im Polizeibereich.	184
cc)	Die »Palantir-Paragrafen« in den Polizeigesetze der Länder	184
a)	§ 25a HSOG – die Regelung von »hessen- Data« und § 49 HmbPolDVG	185
b)	Datenabgleich und -analyse nach § 23 Abs. 5 und § 25 PolG NW.	188
dd)	Regelungen außerhalb der Polizeigesetze	189
ee)	»Smart Home« als Erkenntnisquelle für die Poli- zeibehörden?	190
ff)	Big Data-Anwendungen im Bereich der Straf- verfolgung	192
gg)	Regelungen im europäischen Polizeirecht.	194
hh)	»Big Data« im Bereich der Nachrichtendienste . .	198
ii)	Zwischenfazit	199
e)	Nutzung von Big Data-Anwendungen und KI in der Justiz	200
f)	Weitere Beispiele von Big Data-Anwendungen in der öffentlichen Verwaltung	205
g)	Zwischenergebnis	206
3.	Big Data-Anwendungen im privaten und im öffentlichen Recht	207
a)	Big Data-Anwendungen in Forschung und Wissen- schaft	208
aa)	Big Data-Anwendungen in der Forschung	209
bb)	Big Data-Anwendungen im Gesundheitsdaten- nutzungsgesetz.	213
b)	Nutzung von »Big Data« in der medizinischen Ver- sorgung	216
c)	»Big Data« im Kartellrecht	220
aa)	Kartellrechtliche Fragen aus Unternehmens- sicht	221

bb)	Kartellrechtliche Fragen aus Sicht der Kartellbehörden	224
4.	Big Data-Anwendungen aufgrund des internationalen Reiseverkehrs	225
a)	Abkommen über PNR-Daten zwischen der EU und Drittstaaten	225
aa)	Die PNR-Abkommen zwischen der EU und den USA	225
bb)	Weitere PNR-Abkommen zwischen der EU und Drittstaaten	227
b)	Rechtliche Vorgaben zu PNR-Daten auf UN-Ebene	227
c)	Die PNR-RL der EU vom 27.4.2016	229
d)	Die Entscheidung des EuGH vom 21.6.2022 zum Einsatz von KI	232
e)	Nationale Rechtsgrundlagen zur Verarbeitung von PNR-Daten	235
f)	Zwischenergebnis	241
5.	Selbstregulierungsansätze bei »Big Data«	241
a)	Regelungen zu »Big Data« aufgrund eines »virtuellen Hausrechts«	241
b)	Versuch einer Normierung durch private Regelungen	246
c)	Vorgaben durch Ethik-Richtlinien und Ethik-Kodizes	248
6.	Zwischenfazit	249
III.	Zwischenbilanz	250
D.	Beeinflussung von Big Data durch Verfahrensrecht	252
I.	Verfahrensrecht zur Sicherung von Grundrechten	252
II.	Verfassungsrechtliche Vorgaben	255
1.	Vorgaben aus dem europäischen »Verfassungsrecht«	256
2.	Vorgaben aus dem deutschen Verfassungsrecht	258
3.	Der Grundsatz der Verhältnismäßigkeit bei Big Data-Anwendungen	259
III.	Verfahrensgarantien im Datenschutzrecht	260
1.	Datenschutz-Folgenabschätzungen (DSFA) für Big Data-Anwendungen	260
a)	Datenschutz-Folgenabschätzung nach Art. 35 DSGVO	260
b)	Konformitätsbewertung und Risikomanagementsystem nach der KI-VO	266
c)	Bereits existierende Folgenabschätzungsverfahren	268
d)	Bilanz zur Datenschutz-Folgenabschätzung	269
2.	Benennung eines internen Datenschutzbeauftragten (bDSB)	270

Inhaltsverzeichnis

IV. Datenschutzprinzipien und »Big Data«	273
1. Erosion traditioneller Prinzipien des Datenschutzrechts durch Big Data?	275
a) Verzicht von Grundrechten aufgrund von Fatalismus und Ignoranz	277
b) »Post Privacy«-Bewegung versus Recht auf Datenschutz.	278
c) Folgen des »digitalen Fatalismus« und des »Post Privacy«.	280
d) Zwischenergebnis	281
2. Einwilligung bei Big Data-Anwendungen	282
a) Informierte Einwilligung als ethischer Standard.	282
b) Das klassische Modell der »informierten Einwilligung« (»informed consent«)	283
aa) Die »informierte Einwilligung« und Big Data-Anwendungen	284
bb) Kritik an und Grenzen des klassischen Modells der »informierten Einwilligung«	286
c) Neue Modelle der Einwilligung	288
aa) Die Konstruktion einer »breiten Einwilligung« (»broad consent«)	288
bb) Das Konzept der dynamischen Einwilligung (»dynamic consent«)	292
cc) Das Konzept einer übergeordneten Einwilligung (»meta consent«)	293
dd) Entwicklung einer kontextorientierten Einwilligung (»contextual consent«)?	293
ee) Entwicklung einer »Kaskadeneinwilligung« (»cascading consent«)?	295
gg) Die »Blanko-Einwilligung« (»blanket consent«)?	296
hh) Der »Einwilligungsassistent«	296
ii) Einwilligung bei besonderen Gruppen	298
α) Probleme bei Einwilligungen, die Minderjährige betreffen	298
β) Probleme bei Einwilligungen, die vulnerable Personen betreffen	299
γ) Zwischenfazit.	300
jj) »Einwilligung« trotz vorhandener gesetzlicher Verarbeitungsgrundlage.	300
kk) Fazit.	302
d) Recht auf Widerruf	304
e) Die Idee einer »Datenspende«.	305

f)	Der neue »Datenaltruismus«	307
g)	Das »Recht auf Nichtwissen« als Aspekt der »informierten Einwilligung«	308
h)	Ersetzung der Einwilligung durch eine gesetzliche Regelung	310
i)	Zwischenfazit	310
3.	Big Data-Anwendungen und die Verarbeitungsgrundsätze der DSGVO	311
a)	Der Grundsatz der Rechtmäßigkeit (»Lawfulness«)	312
b)	Der Grundsatz der Verarbeitung nach Treu und Glauben (»Fairness«)	315
c)	Der Grundsatz der Transparenz (»Transparency«)	316
d)	Der Zweckbindungsgrundsatz (»Purpose Limitation«)	322
aa)	Der datenschutzrechtliche Grundsatz der Zweckbindung	322
bb)	Der Grundsatz der kompatiblen Nutzung	325
α)	Die Weiterverarbeitung der Daten für Zwecke der Wissenschaft und Forschung	326
β)	Die Weiterverarbeitung der Daten i. S. v. Art. 6 Abs. 4 DSGVO	327
cc)	Anwendung des Zweckbindungsgrundsatzes auf Big Data-Anwendungen	329
e)	Der Grundsatz der Datenminimierung (»Data Minimization«)	332
f)	Der Grundsatz der Richtigkeit (»Accuracy«)	334
g)	Der Grundsatz der Speicherbegrenzung (»Storage Limitation«)	339
h)	Grundsatz der Integrität und Vertraulichkeit (»Integrity and Confidentiality«)	340
i)	Der Grundsatz der Rechenschaftspflicht (»Accountability«)	340
j)	Das Prinzip der datenschutzgerechten Systemgestaltung (»Privacy by Design«, »Privacy by Default«)	342
k)	Zwischenergebnis	347
4.	Betroffenenrechte und Big Data-Anwendungen	348
a)	Fehlende Nutzung von Betroffenenrechten wegen Informationsasymmetrie	348
b)	Fehlende Betroffenenrechte mangels Identifizierbarkeit	349
c)	Informationsrechte nach Art. 12, 13 und 14 DSGVO	350
d)	Auskunftsrecht nach Art. 15 DSGVO und Big Data	352
e)	Berichtigungsansprüche nach Art. 16 DSGVO	356

Inhaltsverzeichnis

f)	Das »Recht auf Löschung« und das »Recht auf Vergessenwerden«	357
aa)	Das Recht auf Löschung (Art. 17 Abs. 1 DSGVO)	357
bb)	Das Recht auf Vergessenwerden (Art. 17 Abs. 2 DSGVO)	361
cc)	Ausnahmen vom Recht auf Löschung und »Vergessenwerden« (Art. 17 Abs. 3 DSGVO)	363
g)	Das »Recht auf Einschränkung der Verarbeitung« nach Art. 18 DSGVO	363
h)	Das »Recht auf Datenübertragbarkeit« nach Art. 20 Abs. 1 DSGVO	364
i)	Das »Recht auf Widerspruch gegen die Verarbeitung« nach Art. 21 DSGVO	365
aa)	Das allgemeine relative Widerspruchsrecht (Art. 21 Abs. 1 DSGVO)	365
bb)	Das besondere relative Widerspruchsrecht bei Forschungs- und Statistikzwecken	366
cc)	Zwischenergebnis	367
j)	Durchsetzung von Betroffenenrechten auf dem Rechtsweg	367
5.	Big Data und »Profiling«	368
6.	Technische und organisatorische Maßnahmen« nach Art. 32 DSGVO	370
V.	Verfahrensschutz durch Kompensation	377
1.	Kompensation durch verfahrensrechtliche Vorschriften	378
2.	Berücksichtigung ethischer Grundsätze bei Big Data-Anwendungen und KI-Systemen	381
a)	»Ethik« im Recht	385
aa)	Begriff »Ethik«	386
bb)	Big Data, KI und angewandte Ethik	387
cc)	Ethische Grundsätze bei Big Data-Anwendungen und KI-Systemen	388
α)	Achtung der menschlichen Autonomie (»Respect for human autonomy«)	390
β)	Nichtschädigung bzw. Schadensverhütung	391
γ)	Wohltätigkeit (»Beneficence«)	391
δ)	Gerechtigkeit und Fairness	392
ε)	Erklärbarkeit (»Explicability«)	392
ζ)	Lösungen bei Kollisionen der ethischen Grundsätze	392
b)	Folgen ethischer Implikationen von Big Data-Anwendungen und KI-Systemen	393

c)	Ethische Vorgaben durch Ethik-Richtlinien und Ethik-Kodizes	393
d)	Sicherstellung ethischer Vorgaben durch Ethikgremien	395
aa)	Begriffsbestimmungen und vorhandene rechtliche Regelungen	395
α)	Ethikkommissionen	395
β)	Ethikräte	396
bb)	Entstehung von »Ethikkommissionen«	397
cc)	Aufgaben und Befugnisse der Ethikkommission	401
dd)	Kritik am derzeitigen Modell der Ethikkommission	401
ee)	Das neue Modell einer »Ethikkommission«	403
ff)	Rechtsnatur der Entscheidung der Ethikkommission	404
gg)	Zwischenfazit	405
3.	Fazit	406
E.	Möglichkeiten rechtlicher Regulierung von Big Data	407
I.	Erforderlichkeit der normativen Regelung	408
1.	Allgemeine Big Data- und KI-Regelungen im internationalen Bereich	411
a)	Big Data- und KI-Regelungen in Europa	412
b)	Big Data und KI-Regelungen in den USA	414
c)	Geplante KI-Regelungen in Kanada	416
d)	KI-Regulierung in China	417
e)	Vorgesehene KI-Regulierungen in weiteren Staaten und private Normierungen	418
f)	KI-Resolution der Vereinten Nationen	418
2.	Vorhandene Regelungen auf einfachgesetzlicher Basis	419
3.	Gibt es ein Recht auf Nutzung von Big Data-Anwendungen?	419
4.	Notwendigkeit einer rechtlichen Regelung	420
II.	Big Data-Gesetzgebung und Gesetzesfolgenabschätzung	421
1.	Systematische Stellung einer weiteren »Big Data«- und KI-Regulierung	421
2.	Gesetzesfolgenabschätzung	423
III.	Regulative Herausforderungen bei Big Data	426
1.	Festgestellte Bereiche in denen eine rechtliche Regelung fehlt und Lösungsansätze	430
a)	Fehlendes Vertrauen in die Nutzung von Big Data-Anwendungen	430
aa)	Fehlende Transparenz bei Big Data-Anwendungen und KI-Systemen	430

Inhaltsverzeichnis

bb)	Spezialfall »Forschung« – Schaffung eines Forschungsdatengesetzes	435
cc)	Problem der diskriminieren Wirkung von Big Data-Anwendungen.	437
b)	Fehlende Regelungen zur Datenerhebung.	438
c)	Fehlende Regelungen zur Nutzung bestimmter Big Data-Anwendungen.	439
d)	Fehlende Regelungen zur den Datenschutzgrundsätzen und Betroffenenrechten	441
e)	Fehlende Regelungen zur Datensicherheit	442
2.	Das verfassungsrechtliche Bestimmtheitsgebot für Regelungen	443
3.	Haftungsrechtliche Regelungen	445
a)	Der Vorschlag einer KI-Haftungsrichtlinie	448
b)	Die neue Produkthaftungsrichtlinie.	449
c)	Deliktische Haftung.	450
d)	Vertragliche Haftung.	451
e)	Staatshaftungsrecht	452
f)	Haftung nach Art. 82 DSGVO.	453
g)	Stellungnahme zu den Haftungsregelungen	454
4.	Grundvoraussetzungen für eine Regulierung zu »Big Data«.	455
IV.	Zusammenfassung der Ergebnisse	457
F.	Schlussfolgerungen und Ausblick	459
	Literaturverzeichnis	465
	Dokumente der Europäischen Union	523
	Dokumente der OECD.	529

A. Einleitung

In einem Videostatement zur Eröffnung der Cebit 2016 nannte Bundeskanzlerin Angela Merkel die Fülle der digitalen Daten *die »Rohstoffe des 21. Jahrhunderts«*.¹ Tatsächlich handelt es sich bei »Big Data« um einen der Schlüsselbegriffe der aktuellen Debatte über die technologisch induzierte gesellschaftliche Veränderung.² »Big Data« wurde zum Inbegriff eines neuen Zeitalters der Informationsanalyse erklärt³ sowie für die Rechtswissenschaft und Rechtspraxis als von herausragender Bedeutung bezeichnet.⁴ Einige Autoren prophezeien, der Begriff »Big Data« würde bald durch den Begriff »Smart Data« ersetzt werden.⁵ Allerdings hat »Smart Data« »Big Data« nicht ersetzt, sondern ergänzt. Die Frage, ob es sich bei »Big Data« um einen Hype⁶ oder ein Marketing-Buzzword⁷ handelt, wird mittlerweile nicht mehr gestellt. Allerdings ist das Thema »Big Data« zwar nicht

1 Zitiert nach <https://ap-verlag.de/bundeskanzlerin-merkel-daten-sind-die-rohstoffe-des-21-jahrhunderts/19662/> (Abruf 18.8.2024 – der Original-Podcast ist nicht mehr abrufbar); zuletzt Richter, MMR 2023, 163; Spiecker gen. Döhmman, K&R 2017, Beil. 1, 4 Nr. 1 und Ulbricht/Huch/Gens (2015), S. 1; Zech, CR 2015, 137 (138 und 139); ebenso die EU-Kommissarin Meglena Kurnewa in einer Rede am 31.3.2009 (SPEECH 09/156 – »Personal data ist the new oil of the internet and the new currency of the digital world«), Schüller in König/Schröder/Wiegand, S. 111, Müller, DuD 2019, 159 (160); zuletzt Söbbing, ITRB 2022, 206, die vom »Öl der Zukunft«, vom »Öl des 21. Jahrhunderts« sprechen, sowie Hasselbalch/Tranberg in Otto/Gräf, S. 186 (188 – »Treibstoff der Zukunft«) oder vom »Rohstoff des 21. Jahrhunderts« (Geppert, DSRITB 2017, 733; Paal/Hennemann, NJW 2017, 1697. Kritisch Hoffmann-Riem in Hoffmann-Riem, S. 11 (16 ff.); Ringeling, CRi 2015, 7. von Lewinski spricht von einer abgedroschenen Phrase, von Lewinski in Stiftung Datenschutz (Bd.3), S.209 und Funk, CR 2023, 421 von einem Gemeinplatz. Kelber geht davon aus, dass »Nicht Daten [...] der Rohstoff des 21. Jahrhunderts (sind), sondern Vertrauen«, Kelber, DuD 2020, 226.

2 Deutscher Ethikrat, Big Data und Gesundheit, S.47.

3 Martini, DVBl. 2014, 1481.

4 Döpke in Hoeren/Kolany-Raiser, S. 18.

5 Vgl. Paal/Hennemann, NJW 2017, 1697 Fsn. 1 zu Unrecht unter Berufung auf Roßnagel, NJW 2017, 10.

6 Das Präsidiumsmitglied der BITKOM Michael Kleinmeier erklärte auf der Cebit 2013, dass es sich bei Big Data nicht um einen Hype, sondern um einen wichtigen neuen Wirtschaftsfaktor handele, zitiert nach Schmidt-Wehrmann, VW 6/2013, 64; ebenso Heckmann, JRE 23 (2015), 17 (19); a.A. Wehmeier/Baumann in Langkafel, S. 141, 142: »digitales Hype-Thema«; ebenso Weichert in ULD Schleswig-Holstein (2013), Big Data und Datenschutz, S. 5 und Martini, DVBl. 2014, 1481.

7 Vgl. Behrndt/Wagner, JRE 23 (2015), S. 129 (130); Heise in Maireder/Ausserhofer/Schumann/Taddicken, S. 29 (40); Spindler/Seidel, NJW 2018, 2153. Lohr, The Age of Big Data (2012) weist darauf hin, dass »Big Data« ursprünglich als »marketing term« angesehen wurde.

A. Einleitung

verschwunden,⁸ im Thema »Künstliche Intelligenz (KI)« jedoch weitgehend aufgegangen.

Bis vor wenigen Jahren war das Sammeln, Speichern und Auswerten von Daten schwierig und teuer und daher war »Big Data« noch nicht möglich.⁹ In den Archiven des Ministeriums für Staatssicherheit der ehemaligen DDR (Stasi) finden sich jedoch mehr als 100 km Archivgut.¹⁰ Allerdings werden Jahrzehnte nach dem Ende der DDR über jeden Einzelnen mehr Daten gesammelt und gespeichert als je zuvor.¹¹ Bei der im Zuge der Fahndung nach den RAF-Terroristen vom BKA entwickelten »Rasterfahndung« wurden bereits größere Datenmengen zu einem bestimmten Zweck ausgewertet.¹² Erhoben und gespeichert wurden Daten schon früher emsig und in großen Mengen. Die Auswertung dieser riesigen und zum größten Teil unstrukturiert¹³ vorliegenden Daten war jedoch nahezu unmöglich. Zudem wurden Daten meist nur zu einem bestimmten, bei ihrer Erhebung und Speicherung bekannten Zweck gesammelt.¹⁴ Dies hat sich in den letzten Jahren grundlegend geändert. Das gilt sowohl für die Datenmengen als auch für die Möglichkeit, diese auszuwerten. Bereits 2013 ergab eine Studie, dass 90 % der vorhandenen weltweiten Daten erst in den zwei Jahren davor entstanden war. Andere Studien gehen davon aus, dass der Datenbestand jährlich um ca. 40 % wächst.¹⁵ Nicht nur die Datenmenge wächst an, diese liegt zudem digital und damit erheblich besser auswertbar vor.¹⁶ Durch »cloud computing«¹⁷

8 Erst im Jahr 2024 ist unter dem Titel »Big Data« ein 700 Seiten starkes »Big Data«-Buch erschienen, vgl. *Borges/Keil* (Hrsg.), *Big Data*, 1. Aufl. 2024; siehe auch den Titel des Beitrages von *Hornung/Schaller/Selzer/Stummer*, *Privatheit in Big-Data-Verarbeitungssystemen*, K&R 2024, 187 ff.

9 *Mayer-Schönberger*, *Die politische Meinung 2017*, S. 31 (32).

10 Der tatsächliche Umfang der Sammel Leidenschaft der Stasi ist unbekannt. Der BStU beziffert den Umfang der von ihm verwalteten Akten mit 111 Kilometer Akten und mehr als 1,4 Millionen Fotos, *Walther*, *DuD* 2016, 598. In seinem 1. TB ging der damalige BStU *Joachim Gauck* von etwa 178 km Akten aus, nachdem das Bundesarchiv vorher ca. 202 km Akten geschätzt hatte. Hinzu kommen noch ca. 17.200 Säcke mit zerrissenem Akteninhalt, die weitere 25 km Akten ergeben, *BT-Drs. 12/5100* S. 5.

11 *Mayer-Schönberger/Cukier* S. 189.

12 BVerfG, *Beschl. v. 4.4.2006 1 BvR 518/07*, *BVerfGE* 115, S. 320 ff. = *NJW* 2006, 1939; *Paall/Hennemann*, *NJW* 2017, 1697.

13 *Schüller* in *König/Schröder/Wiegand*, S. 111, die darauf hinweist, dass 90 % der Daten unstrukturiert vorliegen; ähnlich *Otte/Wippermann/Schade/Otte*, S. 15.

14 *Mayer-Schönberger*, *Die politische Meinung 2017*, S. 31 (32).

15 Zitiert nach *Thouvenin* in *Boehme-Neßler/Rehbinder*, S. 27 (28) Fsn. 1.

16 Die Stasi hatte diese Möglichkeiten glücklicherweise bis zu ihrer Auflösung im Jahr 1990 (noch) nicht. Es existieren jedoch weiterhin Länder mit vergleichbaren Unrechtssystemen mit ähnlichen Sicherheitsbehörden, die diese technischen Möglichkeiten zur Überwachung ihrer Bevölkerung nutzen.

17 Eine Definition von »cloud computing« findet sich in Art. 6 Nr. 30 NIS-2-RiLi und in dessen *ErwGr* 33.

und der dadurch gegebenen Speicherkapazität ist es möglich, enorm große Mengen an Daten zu erheben, zu speichern und auszuwerten.¹⁸

Obwohl der Begriff »Big Data« (noch) nicht umfassend benutzt wurde, wurden die technisch-konzeptionellen Grundlagen von »Big Data« weitgehend im ersten Jahrzehnt des 21. Jahrhunderts geschaffen.¹⁹ »Big Data« entwickelte sich aus bestehenden Instrumenten im Wesentlichen der amtlichen Statistik sowie soziodemografischen und ökonomischen Auswertungen, die in »Data Warehouses« strukturiert und auswertbar gemacht wurden.²⁰ Den Durchbruch schaffte der Begriff »Big Data« im Jahr 2012, sodass *Steve Lohr* feststellte: »In diesem Jahr hat Big Data den Durchbruch geschafft – als Idee, als Wort und [...] als Marketinginstrument. Big Data hat die Nische der Technologieexperten verlassen und den Mainstream erreicht.«²¹ Das Auswerten riesiger, zum Teil unstrukturiert vorhandener Daten wurde durch Big Data-Anwendungen mit wirtschaftlich akzeptablen Methoden möglich. Bei der Erhebung und Speicherung der Daten steht der Zweck ihrer Auswertung oft noch nicht fest, was den datenschutzrechtlichen Grundsätzen der Datensparsamkeit und der Zweckbindung widerspricht. Verkaufsträchtig wird »Big Data« als »Revolution, die unser Leben verändern wird« bezeichnet.²² »Big Data« stehe am Anfang einer grundlegenden Umwälzung.²³ Dabei ist »Big Data« keine Erfindung des 21. Jahrhunderts. Als erstes großes (modernes) Datenverarbeitungsprojekt wird auf die Entwicklung einer Lochkartenlesemaschine durch die Fa. IBM hingewiesen, mit deren Hilfe man die Auszahlung staatlicher Zuwendungen aufgrund des Social Securi-

18 *Leistner/Antoine/Sagstetter*, Big Data, S. 200.

19 *Geiselberger/Moorstedt* in *Geiselberger/Moorstedt*, S. 13, sprechen von der »Latenzphase des Phänomens«.

20 *Weichert*, ZD 2013, 251 (252).

21 *Lohr*, How Big Data became so big (2012): »THIS has been the crossover year for Big Data — as a concept, as a term and, yes, as a marketing tool. Big Data has sprung from the confines of technology circles into the mainstream«; vgl. *Geiselberger/Moorstedt* in *Geiselberger/Moorstedt*, S. 14.

22 So der Titel von *Mayer-Schönberger/Cukier*, Big Data – Die Revolution, die unser Leben verändern wird, Erstauflage 2013; ebenso *Mayer-Schönberger*, Bundesgesundheitsblatt 2015, S. 788; Siehe *Roßnagel/Geminn/Jandt/Richter* (2016), S. 21; die Aussage von MIT Sloan's Professor *Erik Brynjolfsson*: »This 'Big Data' revolution is occurring mainly because technology enables firms to gather extremely detailed information from, and propagate knowledge to, their consumers, suppliers, alliance partners, and competitors«; in IEDP-Editorial, The Big Data Revolution. Vgl. *Richards/King* (2014), S. 393 und *Kayyalil/Knott/Van Kuiken* (2013) S. 1; *King*, zitiert nach *Behrndt/Wagner*, JRE 23 (2015), S. 129 (130). Kritisch hierzu bereits *Lazer/Kennedy/King/Vespignani*, Science 2014 S. 1203; zuletzt *Selke et al.*, ABIDA-Gutachten Ethische Standards, S. 18 m. w. Nachw.

23 *Mayer-Schönberger/Cukier* S. 14.

A. Einleitung

ty Acts von 1937 verfolgte.²⁴ Weitere große Datenverarbeitungsprogramme folgten nicht nur in den USA, sondern in fast allen entwickelten Staaten. *Roger Mougals* prägte im Jahr 2005 den Begriff »*Big Data*«, um eine große Anzahl von Datensätzen zu beschreiben, die mit herkömmlichen Datenverarbeitungsprogrammen nahezu unmöglich zu verarbeiten war.²⁵

Wenn also nicht ganz neu, ist »*Big Data*« ein Phänomen²⁶, das seit einigen Jahren in der Öffentlichkeit von Politik, Wirtschaft, Wissenschaft etc. diskutiert wird und mit dem sich große wirtschaftliche Hoffnungen verbinden.²⁷ Es verändert nicht nur die IT-Industrie, sondern betrifft weite Teile der Weltwirtschaft und das übrige gesellschaftliche Leben. Selbst Kritiker sprachen vom Beginn des »*anbrechenden Big-Data-Zeitalters*«. ²⁸

Den größten Nutzen verspricht man sich von »*Big Data*« in der Wirtschaft. 88% der Unternehmensdaten würden brachliegen und nicht ausgewertet.²⁹ McKinsey schätzte 2018 weltweit ein Wertschöpfungspotenzial durch die Verarbeitung mithilfe von Big Data-Anwendungen und der Nutzung von »*Künstlicher Intelligenz*« (KI) in Höhe von 3,6 bis 5,8 Billionen US\$ jährlich.³⁰ Wer über die Daten seiner Kunden verfügt, hat nicht nur einen Informationsvorsprung vor seinen Mitkonkurrenten, sondern einen erheblichen Wettbewerbsvorteil. Besondere Vorteile durch Big Data-Anwendungen erhofft man sich im nicht-kommerziellen Bereich, insbesondere in der Medizin. Der Begriff »*Systemmedizin*« beschreibt eine Vision einer Medizin, die alle relevanten Daten eines Patienten und das für seine Situation relevante Wissen aus Forschung und Routineversorgung für Therapie- und Präven-

24 *van Rijmenam*, A Short History of Big Data.

25 *Press* (2013), A Very Short History Of Big Data; *van Rijmenam*, Short History of Big Data; *Hornung/Herfurth* in *König/Schröder/Wiegand*, S. 149 (151); *Kring* in *Plodereder/Grunke/Schneider/Ull* S. 551 (552).

26 Vgl. *Butarelli* in *Otto/Gräf*, S. 216 (217); *Dorschel, W.* in *Dorschel, J.*, S. 2 ff. passim, spricht konsequent vom »*Phänomen Big Data*«; ebenso *Hoeren/Sieber/Holzner* *MultimediaR-Hdb/Kolany-Raiser* Teil 15.1 Rn. 1; *Hornung/Herfurth* in *König/Schröder/Wiegand*, S. 149 (150); *Rouvroy* (2016) S. 4; *Wolf*, Big Data und innere Sicherheit, S. 21. Ebenso bereits *Diebold* (2012), S. 1 und *boyd/Crawford* (2012), S. 662 ff. Letztere sprechen von einem kulturellen, technologischen und wissenschaftlichen Phänomen. *Favaretto/De Clercq/Elger*, *J Big Data* 2019, S. 1 beschreiben es als »*sehr komplexes und umfangreiches Phänomen, das seit seinem Auftauchen in den frühen 2010er Jahren wechselnde Bedeutungen hat*«; *Schefzig*, *K&R* 2014, 772 und S. 774 bezeichnet »*Big Data*« als ein Phänomen, das sich fortwährend entwickelt und daher nicht abschließend beschrieben werden kann.

27 Siehe *Wachter* in *König/Schröder/Wiegand*, S. 17 (18).

28 *boyd/Crawford* (2012), S. 662 (664) (»*How we handle the emergence of an era of Big Data is critical.*«).

29 *Schüller* in *König/Schröder/Wiegand*, S. 111.

30 *McKinsey* (2018), S. 17 (PDF-Version). Das entspricht in etwa dem Bruttoinlandsprodukt Deutschlands, *Knorre* in *Knorre/Müller-Peters/Wagner* S. 5.

tionsentscheidungen zur Verfügung stellt.³¹ Sie soll zielgerichtete Präventions- und Therapiemaßnahmen durch die bioinformatische Verarbeitung und Modellierung großer Datenmengen aus verschiedenen Quellen (z. B. klinische, epidemiologische oder Omics-Daten³²) ermöglichen. Als eines der ersten Beispiele diente die Vorhersage von Grippeepidemien durch das Programm »Google Flu Trends (GFT)«. ³³ Verbunden war dies mit der Hoffnung, durch die Nutzung von Big Data-Anwendungen Epidemien und Pandemien vorauszusagen und zu verhindern.³⁴ Nicht nur die Corona-Pandemie im Jahr 2020 zeigte, dass diese Hoffnung viel zu optimistisch war. Mittlerweile gibt es durch KI unterstützte Verfahren, die auf einer neuronalen Netzwerkanalyse basieren, die gerade im Bereich der Vorhersage von Pandemieverläufen deutlich besser funktionieren.³⁵ Nach einer PM von BITKOM vom 13.3.2018 wuchs in Deutschland der Markt für Big Data-Anwendungen verglichen mit den Werten für 2016 in 2017 und 2018 jeweils zweistellig.³⁶

Mit der Nutzung von Big Data-Anwendungen kann in das nach Art. 7 und 8 GRCh und Art. 2 Abs. 1 i. V. m. Art. 1 Abs. 1 GG gewährleistete Recht auf informationelle Selbstbestimmung eingegriffen werden. Der Einsatz von KI intensiviert die für das informationelle Selbstbestimmungsrecht relevante Gefahr, dass aus personenbezogenen Informationen »weitere Informationen erzeugt und so Schlüsse gezogen werden, die sowohl die grundrecht-

31 Winkler, Frankfurter Forum 2017, Heft 16, S. 23.

32 Unter »Omics-Daten« versteht man Daten, die aus Teilgebieten der Medizin oder Biologie stammen, die sich mit der Analyse von Gesamtheiten ähnlicher Einzelemente beschäftigen, z. B. Genomik (engl. genomics), Glykomik (engl. glykomics), Metabolomik (engl. metabolomics oder metabonomics), Proteomik (engl. proteomics) etc., vgl. Summa in Sturma/Lanzerath, S. 74.

33 Die von 50 Mio. von US-Bürgern eingegeben Suchanfragen nach Medikamenten gegen Grippeviren, z. B. TamiFlu, verglichen Mitarbeiter der Fa. Google mit den Zahlen zu den Grippeepidemien des Centers for Disease Control and Prevention (CDC) aus 2003 bis 2008, um eine Korrelation zwischen diesen Datenbeständen herzustellen und den Verlauf einer Grippeepidemie zu verfolgen, siehe Buchner/Schnebbe in Sturma/Lanzerath, S. 49 (51 f.); Mayer-Schönberger/Cukier S. 7 f.; Reichert, Politikum 2016 (Heft 1), 20 (S. 25). Zu diesem Zeitpunkt waren die Schwächen des Programms bereits bekannt, näher hierzu Raum in Stiftung Datenschutz (Bd. 2), S. 137. 2010 war nach dem Erdbeben auf Haiti durch Big Data-Anwendungen die aktuelle Ausbreitung der Cholera fast in Echtzeit verfügbar war, Geppert, DSRITB 2017, 733. Kritisch zu Google Flu zuletzt SVR Gesundheitswesen, Gutachten 2023, BT-Drs. 20/5500 S. 549 Rn. 1065.

34 Vgl. Mayer-Schönberger/Cukier S. 9: »Klar ist, dass bei der nächsten Pandemie die Menschheit über ein Instrument verfügt, um im Falle einer Pandemie die Ausbreitung vorauszusagen und damit zu verhindern.«

35 Vgl. SVR Gesundheitswesen, Gutachten 2023, BT-Drs. 20/5500 S. 553 Tz. 1065.

36 BITKOM, PM v. 13.3.2018, <https://www.bitkom.org/Presse/Presseinformation/Marktfuer-Big-Data-waechst-in-Deutschland-zweistellig.html> (Abruf 27.8.2024).