

Smart Regulation: Quick Data or Big Data?

Herausgegeben von
RAINER NIEMANN,
MATTHIAS WENDLAND
und JOHANNES ZOLLNER

Mohr Siebeck

Smart Regulation: Quick Data or Big Data?



Smart Regulation: Quick Data or Big Data?

herausgegeben von

Rainer Niemann, Matthias Wendland und Johannes Zollner

Mohr Siebeck

Rainer Niemann, ist Professor für Unternehmensrechnung und Steuerlehre und stellvertretender Sprecher des Profilsbereichs Smart Regulation der Universität Graz
orcid.org/0000-0001-9168-1144

Matthias Wendland, ist Professor für Wirtschaftsrecht, Recht der Künstlichen Intelligenz und Datenschutzrecht am Institut für Unternehmensrecht und Internationales Wirtschaftsrecht der Universität Graz.
orcid.org/0000-0002-1834-9361

Johannes Zollner, ist Professor für Unternehmensrecht und Internationales Wirtschaftsrecht und Sprecher des Profilsbereichs Smart Regulation der Universität Graz.

ISBN 978-3-16-162202-1 / eISBN 978-3-16-162203-8
DOI 10.1628/978-3-16-162203-8

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

Publiziert von Mohr Siebeck Tübingen 2024. www.mohrsiebeck.com

© Rainer Niemann, Matthias Wendland und Johannes Zollner (Hg.); Beiträge: jeweiliger Autor/ jeweilige Autorin.

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/>.

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 dort gebunden.

Printed in Germany.

Vorwort

Im vorliegenden Sammelband finden sich die Beiträge zum Symposium „Quick Data or Big Data?“ des Profilbereichs Smart Regulation, das am 1. Juli 2022 an der Karl-Franzens-Universität Graz stattgefunden hat. Smart Regulation vereint WissenschaftlerInnen aus der Rechtswissenschaftlichen, Sozial- und Wirtschaftswissenschaftlichen, Katholisch-Theologischen und Naturwissenschaftlichen Fakultät. Das Symposium bildete nach zahlreichen vorangegangenen gemeinsamen Aktivitäten die dritte große interdisziplinäre Konferenz des Profilbereichs. Smart Regulation hat sich die Erforschung zukunftsweisender Regulierungsansätze und -methoden zum Ziel gesetzt. Dabei spielen die Chancen und Risiken technologischer Disruptionen wie der Digitalisierung und der künstlichen Intelligenz eine wesentliche Rolle.

Im Sinne dieses Ziels konzipierten die Herausgeber dieses Sammelbandes Rainer Niemann, Matthias Wendland und Johannes Zollner gemeinsam das Symposium. Es stand unter dem Generalthema „Quick Data or Big Data?“ und führte die vergangenen Symposien zu den Themen „Theorie- und evidenzbasierte Politik?“ (2021) und „Vertrag, Unternehmung und Markt“ (2020) logisch fort. Im Mittelpunkt stand die Frage nach den rechtlichen, ökonomischen und ethischen Implikationen der immer schnelleren Verfügbarkeit großer Datenbestände auf Wissenschaft, Wirtschaft und Gesellschaft.

Die in diesem Tagungsband abgedruckten Beiträge sind in drei thematische Schwerpunkte unterteilt: (1) Wissenschaft und Quick Data, (2) Ethik und Quick Data, (3) Marketing und Quick Data. Die zu diesen Fragenkreisen gehaltenen Haupt- und Koreferate aus Wissenschaft und Praxis zeigen das breite inhaltliche und methodische Spektrum der am Profilbereich beteiligten Disziplinen auf. Zusammen liefern sie Impulse für die weitere interdisziplinäre Diskussion von Smart Regulation unter dem Vorzeichen des gemeinsamen Mission Statements „Wir generieren theorie- und evidenzbasiert Wissen für Wissenschaft, Praxis und Regulierungsinstitutionen aus interdisziplinärer Perspektive, um das Regulierungsziel unter Beachtung ethischer Grundsätze möglichst effizient zu erreichen“

Die Herausgeber danken Dr. Maria Fanta und Verena Kramhöller, beide Karl-Franzens-Universität Graz, für ihren unermüdlichen Einsatz bei der Organisation des Symposiums und der Vorbereitung des Tagungsbandes. Für die Unterstützung bei der Erstellung des Sammelbands danken wir dem Redaktionsteam Stefanie Pendl, M.Sc. und Mag. iur. Lukas Soritz, beide Karl-Fran-

zens-Universität Graz. Die für den interdisziplinären Austausch wichtige Veröffentlichung des Sammelbands sowohl als Print- als auch als Open-Access-Medium hat ein Druckkostenzuschuss der Karl-Franzens-Universität Graz ermöglicht.

Graz, im September 2023

Rainer Niemann
Matthias Wendland
Johannes Zollner

Abkürzungsverzeichnis

Abb.	Abbildung
AI	Artificial Intelligence
bspw.	beispielsweise
B-VG	Bundes-Verfassungsgesetz
bzw.	beziehungsweise
CAS	Complex Adaptive System
CASE	Complex Adaptive Service Ecosystem
CDC	Center for Disease Control and Prevention
et al.	et alii
etc.	et cetera
Fig.	Figure
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
i.e.	Id est
IoT	Internet of Things
KFZ	Kraftfahrzeug
KI	Künstliche Intelligenz
lat.	lateinisch
LKW	Lastkraftwagen
m.E.	meines Erachtens
NFT	Not-Fungible Token
PKW	Personenkraftwagen
S-D	Service-Dominant
SDG	Sustainable Development Goals
SE	Service Ecosystems
u.a.	unter anderem/und andere
u.v.m.	und viele(s) mehr
US	Vereinigte Staaten
v. Chr.	vor Christus

VIII

Abkürzungsverzeichnis

VfGH

Verfassungsgerichtshof

z.B.

zum Beispiel

Inhalt

Vorwort.....	V
Abkürzungsverzeichnis.....	VII
Inhalt.....	IX

I. Wissenschaft und Quick Data

Andreas Glöckner

Big Data und Quick Data: Konsequenzen für eine effiziente Wissenschaft.....	3
--	---

Thomas Pözlner

Big Data und Quick Data in der Philosophie.....	15
---	----

II. Ethik und Quick Data

Jürgen Fleiß

Quick Data als Entscheidungsgrundlage.....	23
--	----

Thomas Gremsl

Big Data zwischen Tempo und Tragweite Gesellschaftliche Herausforderungen und ethische Perspektiven in der Ära der Echtzeitdatenverarbeitung.....	31
---	----

Elisabeth Hödl

Echtzeitdaten im urbanen Raum: Regulatorische Implikationen für die Raumordnung.....	47
---	----

III. Marketing und Quick Data

Michael Ehret, Stefan Thalmann

Real-time Service-Ecosystems: Institutional Voids and Governance Implications.....	61
---	----

Tiziana Russo-Spena, Cristina Mele Complex Adaptive Service Ecosystems (CASE): A Focus on Blockchain Case Technology.....87

Autor*innenverzeichnis 109

Sachverzeichnis..... 113

Personenverzeichnis..... 115

I. Wissenschaft und Quick Data

Big Data und Quick Data: Konsequenzen für eine effiziente Wissenschaft

Andreas Glöckner

I. Wissenschaftstheoretische Grundlagen und Begriffsklärung	3
II. Potenziale von Big Data und Quick Data	6
1. Bessere empirische Testung von Theorien	6
2. Entwicklung einer neuen Klasse von Theorien	6
3. Neue methodische Ansätze zur Weiterentwicklung von Theorien	7
III. Herausforderungen von Big Data und Quick Data.....	8
1. Schwierigkeiten bei der Identifikation kausaler Zusammenhänge	8
2. Unreliable Messung und Bedrohung der internen Validität.....	9
3. Als-Ob Modelle	9
4. Verstärkung typischer Probleme und Verzerrungen im Forschungsprozess	10
IV. Diskussion und Implikationen	12

I. Wissenschaftstheoretische Grundlagen und Begriffsklärung

Erkenntnisfortschritt in den empirischen Wissenschaften wird – nach der klassischen Perspektive des *Kritischen Rationalismus*¹ – generiert durch eine möglichst effiziente Testung starker Theorien. Theorien bezeichnen dabei eine Menge an Aussagen (*Implikationen*) der Form: Für alle X gilt, WENN X gewisse Kriterien erfüllt (*Antezedens*), DANN wird X auch gewisse andere Kriterien erfüllen (*Konsequenz*). Ein Beispiel für eine solche Theorie wäre: Wenn Person X in einer Zielerreichung frustriert wird, wird sich Person X aggressiv verhalten. Dabei sollten alle enthaltenen Konzepte (*Konstrukte*) und Begriffe sowie deren Relation klar definiert sein (bspw. was versteht man unter Frustration und Aggression, ggf. auch Person? Wie hängen diese zusammen?). Theorien sollten dabei möglichst vollständig spezifiziert werden und durch eine

¹ Die Erläuterung der Grundlagen des Kritischen Rationalismus findet sich in *Popper* (1934/2005).

kritische Prüfung aller sich aus den explizierten Aussagen logisch ableitbaren Implikationen sollte sichergestellt werden, dass logische Widersprüche (*Inkonsistenzen*) und Zirkelschlüsse (*Tautologien*) vermieden werden.

Die Qualität von Theorien kann (u.a.) anhand ihres *empirischen Gehalts* beurteilt werden. Dieser bezeichnet die Anzahl der Möglichkeiten zur Falsifikation einer Theorie.² Je mehr Möglichkeiten (bspw. Verhaltensweisen) eine Theorie verbietet bzw. ausschließt, desto mehr Vorhersagen ermöglicht sie und desto höher ist ihr empirischer Gehalt.³ Der empirische Gehalt einer Theorie steigt dabei mit steigender Breite des Anwendungsbereichs der Theorie (einer inklusiven Antezedens) und mit steigender Spezifikation der Vorhersage (einer exklusiven Konsequenz).

Die zielgerichtete Durchführung von Experimenten zur kritischen Testung wichtiger Theorien bzw. die vergleichende Testung verschiedener Theorien führt zu Erkenntnisfortschritt durch Falsifikation bzw. Bewährung, Revision und Neuentwicklung von Theorien. Laut *Platt*⁴ schreiten Wissenschaften, die sich eine Methode der Forschung auf Basis starker Schlussfolgerungen (*Strong Inference Research*) zu eigen gemacht haben, dabei deutlich schneller bei der Generierung von Wissen voran als Wissenschaften, die andere Methoden anwenden. *Platt* argumentiert dabei im Kern, dass sich Forscher*innen vor der Durchführung jeder Studie die Frage stellen sollten, welche relevante(n) Theorie(n) mit dem geplanten Experiment potenziell falsifiziert werden könnte(n). Nur Experimente, die das Potenzial haben, relevante Theorien zu falsifizieren, sollten auch tatsächlich durchgeführt werden. Ein besonderer Fokus sollte auf Forschungsvorhaben gelegt werden, die es erlauben, eine große Anzahl sich widersprechender Theorien gleichzeitig mittels kritischer Experimente zu testen.

Die klassische Perspektive des Kritischen Rationalismus wurde in verschiedener Hinsicht kritisiert und weiterentwickelt.⁵ Bei der Testung von Theorien ergibt sich bspw. die Problematik, dass Theorien in der Regel unterbestimmt sind (Duhem-Quine-These), so dass in einem Experiment nicht nur die Theorie selbst, sondern gleichzeitig auch eine Reihe von Zusatzannahmen getestet werden. Diese Unterspezifikation betrifft bspw. Zusatzannahmen darüber, wie die in der Theorie enthaltenen Konstrukte adäquat gemessen werden können (*Operationalisierungen*). Aktuelle Vorschläge dieses und weitere Probleme zu lösen oder zumindest abzuschwächen bestehen u.a. darin, eine Vielzahl von Studien durchzuführen mit einer repräsentativen Auswahl von Operationalisierungen und Stimuli⁶ oder der Entwicklung und Explikation allgemein akzeptierter,

² Für aktuelle Erläuterungen siehe dazu *Glöckner/Betsch* (2011).

³ Die Grundlagen des Konzepts sind nachzulesen in *Popper* (1934/2005).

⁴ Siehe *Platt* (1964).

⁵ Eine aktuelle Zusammenfassung findet sich in *Oreskes* (2019).

⁶ Diese Idee wird entwickelt in *Oberauer/Lewandowsky* (2019).

valider Operationalisierungen in einem in Publikationen dokumentierten Konsensus-Prozess.⁷

Wie verändern Big Data und Quick Data wissenschaftliches Arbeiten und die Effizienz von Wissenschaft vor dem Hintergrund dieser klassischen und neueren bzw. alternativen wissenschaftstheoretischen Perspektiven?⁸ Wird der klassische, inhaltlich-experimentelle Forschungsansatz gar überflüssig? Diese und weitere Fragen sollen im Folgenden diskutiert werden.

Zunächst eine Begriffsklärung: Unter *Big Data* versteht man Daten, die zu groß, zu komplex oder zu schnelllebig sind, um diese mit herkömmlichen Verfahren analysieren und auswerten zu können. Entsprechend müssen alternative Analyseverfahren wie Künstliche Intelligenz (KI) oder Maschinelles Lernen zu deren Auswertung herangezogen werden. Ein prominentes Beispiel sind Ansätze des Deep Learnings unter Nutzung komplexer neuronaler Netze.⁹ *Quick* (bzw. *Fast*) *Data* erweitert diese Perspektive dahingehend, dass eine zeitliche Dimension berücksichtigt wird. Quick Data bezeichnet die Anwendung von Big Data Analysemethoden auf aktuell anfallende Daten (oft in Echtzeit) bzw. die fortlaufende Analyse sich schnell verändernder Daten.

Dies umfasst einerseits die Analyse großer Datenmengen pro Person (bspw. produziert durch Aufzeichnungen der Blickbewegungen via Eye-Tracking) oder andererseits über viele Personen hinweg (bspw. produziert durch die Aufzeichnung von Bewegungsmustern anhand von Smartphone-Daten). *Mobile Sensing* und *Digital Phenotyping* bezeichnen die Nutzung mobiler Geräte zur Erfassung von Daten mit multiplen Sensoren zur Analyse und Vorhersage von Verhalten, psychischer Gesundheit, Persönlichkeit etc. In der Psychologie erfreut sich darüber hinaus die *Experience-Sampling* Methode großer Beliebtheit, in welcher Personen in regelmäßigen Abständen via Mobiltelefon Abfragen zu ihrem aktuellen Erleben und ihren aktuellen Erfahrungen gestellt werden (bspw. haben Sie in den letzten 10 min eine Frustration erlebt?).

Die Möglichkeit zur Nutzung von Big Data und Quick Data bringt verschiedene Herausforderungen mit sich, es ergeben sich aber auch Potenziale, welche im Folgenden auszugsweise dargestellt werden.

⁷ Leising et al. (2022) spezifizieren anhand der Persönlichkeitsforschung zehn Schritte zur Verbesserung des wissenschaftlichen Erkenntnisprozesses. Das Fördern von Prozessen zur Bildung von Konsensus bezüglich zentraler Forschungsfragen, Methoden, Theorien und der Bewertung des Status der Empirie bildet einen Kernpunkt des Vorschlags.

⁸ Für einen alternativen wissenschaftstheoretischen Ansatz siehe auch Eronen/Bringmann (2021).

⁹ Siehe Nielson (2015) für eine Einführung und einen Überblick über neuronale Netzwerkmodelle.

II. Potenziale von Big Data und Quick Data

1. *Bessere empirische Testung von Theorien*

Verfahren zum Sammeln von Big Data und Quick Data liefern per Definition eine große Menge an Daten. Oft sind diese Daten preiswert oder fallen gar kostenlos an. An die Stelle der kostenintensiven Sammlung von Primärdaten in Labor-Experimenten tritt demnach die Aufbereitung und (Re-)Analyse bestehender Datensätze. Diese Daten fallen oft im „realen Leben“ an und es werden Daten von unterschiedlichsten Personen erhoben. Entsprechend sind Befunde mit hoher Wahrscheinlichkeit generalisierbar auf andere Situationen und Personen und es liegt somit potenziell eine hohe externe Validität vor. Darüber hinaus können durch eine KI-basierte Vorverarbeitung Konstrukte direkt gemessen oder Ersatzvariablen (*Proxys*) für Theorie-relevante Konstrukte berechnet werden. So könnte man bspw. mittels neuronaler Netze Emotionen wie Aggression anhand von Gesichtsausdrücken bestimmen oder über Ersatzvariablen wie bspw. besonders dichtes Auffahren im Straßenverkehr erschließen.

Die große verfügbare Datenmenge in Anwendungen von Big und Quick Data bietet eine Vielzahl von Möglichkeiten zur Testung der Vorhersagen von Theorien. So ermöglicht die oft detaillierte, kontinuierliche und gleichzeitige Messung unterschiedlicher Aspekte des Verhaltens, Erlebens und der Umgebungsbedingungen die Entwicklung und Testung von Theorien mit besonders hohem empirischen Gehalt, welche Vorhersagen zu unterschiedlichen abhängigen Variablen gleichzeitig treffen. In typischen Experimenten im Labor wird hingegen oft nur eine (unabhängige) Variable manipuliert und eine andere (abhängige) Variable einmalig gemessen.

So wäre es beispielsweise möglich, neben aggressiven Verhaltensweisen (gemessen durch Bewegungsdaten) auch emotionale oder physiologische Reaktionen (gemessen durch den Puls) in Theorien zu berücksichtigen und für diese konkrete Vorhersagen zu machen. Die hohe Menge an Daten ermöglicht prinzipiell außerdem die Testung von Theorien mit hoher Teststärke (bzw. niedrigem beta-Fehler), so dass potenziell auch kleine vorhandene Effekte mit hoher Wahrscheinlichkeit identifiziert werden können.

2. *Entwicklung einer neuen Klasse von Theorien*

Die Nutzung von Big Data und Quick Data und den entsprechenden KI-Methoden ermöglicht die Entwicklung neuer, „perfekter“ Theorien und eine kontinuierliche Revision dieser Theorien. Speziell wird es möglich, Theorien in neuronalen Netzen (*Deep Learning Netzen*) abzubilden, indem man diese Netze unter Nutzung aller vorhandenen Daten zum Zusammenhang der in der Theorie beinhalteten Konstrukte trainiert. Neuronale Netze ausreichender Größe erlauben eine perfekte Abbildung jedes funktionalen Zusammenhangs

zwischen Input- und Output-Variablen. Bei einer ausreichend großen Datenmenge ist entsprechend eine perfekte Abbildung des tatsächlich zugrundeliegenden Zusammenhangs möglich. Es kann dabei gleichzeitig von einer typischerweise eher groben zu einer präziseren Darstellung des Zusammenhangs zwischen Konstrukten übergegangen werden (bspw. von der Annahme eines positiven Zusammenhangs zwischen Frustration und Aggression zu einem spezifischen funktionalen Zusammenhang wie linear oder quadratisch etc.).¹⁰

Eine weitere aus einer wissenschaftstheoretischen Perspektive interessante Eigenschaft netzwerkbasierter Theorien besteht darin, dass diese es ermöglichen, (i) große Datenmengen aus unterschiedlichsten Quellen und von unterschiedlichen Forscher*innen zusammenzuführen und (ii) die Theorie durch neue Daten fortlaufend graduell zu revidieren. Somit wird ein gemeinschaftlicher, kumulativer Erkenntnisprozess im Sinne *Poppers* und *Platts* befördert. Die Gesamtheit der Befunde aller Forscher*innen bildet dabei den Datenkorpus einer Big Data Analyse, die darüber hinaus auch eine zeitliche Komponente beinhalten könnte (vgl. Quick Data).

Die resultierenden netzwerkbasierten Theorien könnten auch leicht offen verfügbar gemacht werden in online zugänglichen Theoriedatenbanken¹¹ und über grafische Nutzeroberflächen.¹²

3. Neue methodische Ansätze zur Weiterentwicklung von Theorien

Neben der besseren Testung von Theorien und der Entwicklung netzwerkbasierter Theorien, eröffnen Big Data und KI auch neue methodische Ansätze zur Weiterentwicklung traditioneller Theorien. Eine dieser Methoden besteht darin, auf Basis großer Datenmengen unter Nutzung von KI Verhalten vorherzusagen und damit eine Abschätzung der oberen Grenze für die Vorhersagbarkeit dieses Verhaltens zu bestimmen (*Upper-bound-Estimation*). Diese Möglichkeit ergibt sich aus der bereits erwähnten Fähigkeit neuronaler Netze, beliebige funktionale Zusammenhänge nachzubilden, ohne dass es dazu bestimmter Vorgaben bedarf. Es entfällt somit die Notwendigkeit zur vergleichenden Prüfung einer Vielzahl potenzieller funktionaler Zusammenhänge (z.B. lineare, quadratische oder kubische Zusammenhänge; Interaktionseffekte beliebiger Ordnung zwischen Variablen). Diese wird durch das Trainieren des Neuronalen

¹⁰ Vergleiche dazu auch den aktuellen Ansatz zur mathematischen Spezifikation von Theorien von *Borsboom et al.* (2021).

¹¹ Siehe dazu den Beitrag von *Glöckner/Fiedler/Renkewitz* (2018) zur Bereitstellung aller Theorien in online verfügbaren Theorie-Datenbanken.

¹² Am Beispiel der Generierung persönlichkeitspezifischer Empfehlungen im Umgang mit COVID-19 wird bspw. in *Michels/Glöckner/Giersch* (2021) dargestellt, wie auf Basis neuronaler Netze generiertes Wissen über grafische Nutzeroberflächen breit online geteilt werden kann.

Netzwerkes mit den vorhandenen (Input- und Output-) Daten (sowie Verfahren zur Vermeidung einer zu starken Anpassung, des *overfitting*) ersetzt.

Der Vergleich der aktuellen Vorhersageleistung einer Theorie mit dieser Obergrenze ermöglicht es zu bestimmen, ob und wie viel Potenzial (bzw. Bedarf) für die Weiterentwicklung einer inhaltlichen Theorie überhaupt besteht.¹³

Darüber hinaus kann analysiert werden, wie die Wegnahme bestimmter Informationen bzw. die Bereitstellung in gewisser Weise vorverarbeiteter Daten (*Feature Construction/Feature Engineering*) die Vorhersagequalität der KI Modelle beeinflusst. Diese Informationen erlauben es noch zielgenauer, Potenziale für spezifische Weiterentwicklungen inhaltlicher Theorien zu bestimmen.¹⁴

These 1: Big Data und Quick Data haben ein großes Potenzial, Erkenntnisgewinne in der Wissenschaft zu befördern durch die Möglichkeit der besseren empirischen Testung von Theorien, der Spezifikation einer neuen Art von Theorien sowie neuer methodischer Ansätze der Theorieentwicklung.

III. Herausforderungen von Big Data und Quick Data

1. Schwierigkeiten bei der Identifikation kausaler Zusammenhänge

Die Analyse und der Nachweis kausaler Zusammenhänge auf Basis von Big und Quick Data ist oft schwierig bzw. nur bedingt möglich. Zum eindeutigen Nachweis kausaler Zusammenhänge zwischen Variablen eignen sich in besonderer Weise Experimente. Diese sind dadurch gekennzeichnet, dass eine (unabhängige) Variable systematisch manipuliert wird, Personen oder Stimuli zufällig auf die aus dieser Manipulation resultierenden Bedingungen zugeteilt werden und alle anderen Faktoren konstant gehalten werden. Diese Voraussetzungen eines Experiments lassen sich bei Big Data Analysen in der Regel nicht realisieren, da anfallende Daten lediglich aufgezeichnet werden. Manipulationen und zufällige Zuweisungen auf Bedingungen sind dann unmöglich.

Bezüglich des allgemeinen methodischen Ansatzes fallen Big Data Analysen deshalb vorwiegend in die Klasse korrelativer oder bestenfalls quasi-experimenteller Studien (d.h. Studien ohne Zufallszuweisung zu Bedingungen). Damit (i) kann nicht zweifelsfrei ausgeschlossen werden, dass beobachtete Zusammenhänge durch Drittvariablen getrieben werden (*Konfundierung*), (ii) die

¹³ Siehe *Peysakhovich/Naecker* (2017) für die Nutzung dieses Ansatzes zur Testung prominenter Theorien des Entscheidens.

¹⁴ Beispiele für eine effiziente Anwendung dieser Methode im Bereich der Entscheidungsforschung liefern die Beiträge von *Erev et al.* (2017) und von *Peterson et al.* (2021).

kausale Wirkrichtung bleibt oft unklar und (iii) wichtige vermittelnde Variablen (*Mediatoren*) können übersehen werden. So lässt sich mit Hilfe von Big Data vermutlich zeigen, dass Schuhgröße und mathematische Fähigkeiten zusammenhängen. Der Schluss, dass die Größe der Schuhe ursächlich dafür ist, dass Personen besser rechnen können, ist natürlich abwegig. Auch eine umgekehrte Wirkung der mathematischen Fähigkeit auf die Schuhgröße scheint unwahrscheinlich. Der Zusammenhang wird hingegen vermutlich durch eine Konfundierung mit dem Alter getrieben, welches höhere Ausprägungen auf beiden Variablen bedingt.

2. *Unreliable Messung und Bedrohung der internen Validität*

Eine weitere Problematik ergibt sich daraus, dass die Messungen von Konstrukten im Rahmen von Big Data Analysen oft nicht auf den konkreten theoretischen Sachverhalt zugeschnitten sind. Entsprechend sind diese Messungen oft stark fehlerbehaftet und ungenau. Auch ist nicht auszuschließen, dass neben dem Konstrukt, an dem man interessiert ist, auch noch andere Konstrukte in die Messwerte eingehen. Selektionseffekte können dabei ebenfalls eine große Rolle spielen. Dies kann die interne Validität von Analysen einschränken, das heißt es bleibt unklar, ob potenziell beobachtete Zusammenhänge wirklich durch Variationen der interessierenden Variable getrieben werden. Auch sind die Zusammenhänge zwischen (mittels KI generierten) Proxys und relevantem Konstrukt oft niedrig (bspw. zur Vorhersage von Persönlichkeit aus Online-Verhalten¹⁵ oder des Wohlbefindens¹⁶). Eine solche unreliable Messung (mit Reliabilitäten oft nur um $r = .30$) kann ausreichend und nützlich sein für praktische Zwecke, erschwert aber eine zielgenaue Theorie-Testung. Unter anderem verschärft sich bei Big Data Analysen durch die unreliable und potenziell fehlerbehaftete Messung auch die bereits erwähnte Problematik, dass gleichzeitig mit einer Theorie verschiedene Zusatzannahmen (in diesem Fall bezüglich der Messung von Konstrukten) getestet und Rückschlüsse von der Empirie auf die Theorie schwieriger werden (Duhem-Quine-These).

3. *Als-Ob-Modelle*

Durch die Fokussierung auf Big Data und Quick Data und bei der Entwicklung netzwerkbasierter Theorien besteht die Gefahr der Entwicklung von *Black-Box-Modellen* zur Vorhersage von Verhalten, die lediglich Input- und Output-Variablen berücksichtigen. Die aus einer psychologischen Perspektive besonders interessanten vermittelnden Prozesse können dabei leicht aus dem Blick geraten. Außerdem ist das Verständnis psychologischer Prozesse wichtig für

¹⁵ Siehe dazu Kosinski/Stillwell/Graepel (2013).

¹⁶ Eine breitere Diskussion der Vor- und Nachteile der Nutzung von Big Data zur Erforschung von Wohlbefinden findet sich in Luhmann (2017).

die Entwicklung möglichst sparsamer und effizienter Interventionen zur Veränderung von Verhalten und zur Behandlung psychischer Störungen.

4. Verstärkung typischer Probleme und Verzerrungen im Forschungsprozess

Durch den unreflektierten Einsatz von Big Data und Quick Data besteht die Gefahr, dass typische Verzerrungen im und Probleme des Forschungsprozesses noch verstärkt werden. Dies betrifft u.a. die Phänomene (i) der *konfirmatorischen Informationssuche*, (ii) den Rückschaufehler (*Hindsight Bias*), (iii) die *Kohärenz-Verzerrungen* und (iv) einen zu geringen Fokus auf Theorien und die Theorieentwicklung in der empirischen Forschung.

Konfirmatorische Informationssuche bezeichnet das Phänomen, dass Menschen dazu tendieren (u.a.) zur Prüfung von Hypothesen nach Informationen zu deren Bestätigung zu suchen und relevante Gegenevidenz (bzw. besonders relevante kritische Tests der Hypothese) nicht ausreichend zu betrachten.¹⁷ Dies führt zu Fehlschlüssen bezüglich der Gültigkeit von Hypothesen/Theorien und behindert den Erkenntnisprozess und die effiziente Generierung von Wissen. Diese Problematik wird bei der Nutzung von Big Data und Quick Data potenziell noch verstärkt, da aufgrund der Vielzahl an verfügbaren Informationen und Analysestrategien noch mehr konfirmatorische Information potenziell verfügbar ist. Eine ausgewogene Auswahl von Informationen und die diskonfirmatorische Testung von Theorien wird in diesem Kontext noch wichtiger. Da bei den Analysen in der Regel anfallende Daten verwendet werden, ist es oft weniger eindeutig, welche Informationen zur Beantwortung einer Fragestellung herangezogen werden sollten als in speziell zur Testung einer Theorie konstruierten Labor-Experimenten. Die Gefahr der konfirmatorischen Informationssuche im Rahmen der Analyse sowie der Auswertung und des konfirmatorischen Berichtens von Ergebnissen erhöht sich dadurch. Eine ausreichende Reflexion der Möglichkeiten zur kritischen (diskonfirmatorischen) Testung von Theorien vor der Durchführung einer Big Data Analyse ist deshalb besonders wichtig. Auch kommt der Prä-Registrierung von Hypothesen und Analysen sowie dem vollständigen, transparenten Berichten aller durchgeführten Analysen bei Big Data und Quick Data Analysen eine noch größere Bedeutung zu als bei klassischen Experimenten.

Der Rückschaufehler (*Hindsight Bias*)¹⁸ besteht darin, dass Menschen dazu tendieren, in der Rückschau das Eintreten von Ereignissen (die sich bereits ereignet haben) als wahrscheinlicher zu beurteilen als in einer Vorabbetrachtung

¹⁷ In der klassischen Arbeit von *Wason* (1960) wird das Konzept der Konfirmatorischen Informationssuche erläutert und empirisch nachgewiesen.

¹⁸ Klassische Arbeiten zum *Hindsight Bias* finden sich in *Fischhoff* (1975).

Sachverzeichnis

- Analysen
 - explorativ 9, 16
 - konfirmatorisch 8, 16
- Bias 8, 22, 36, 107
- Black-Box 7
- Blockchain technology 86, 92, 96, 97, 98, 99
- Co-adaptation 94, 96
- Co-evolution 88, 94, 99
- common-pool
 - governance 60, 78
 - resource 69, 71
- Complex adaptive service ecosystems (CASE) VI, 85, 93, 96, 98
- Complex adaptive systems (CAS) 88, 89, 91
- Complexity theory 87, 91
- Constitutive rules 61, 73, 74, 75
- Contracting theory 64
- Cultural cognitive perspective 66, 67
- Data
 - governance 63, 64, 65, 68, 70, 72, 75
 - protection 65
 - sharing platform 77
- Daten
 - schutz 30, 38
 - verfügbarkeit 24
 - vernetzung 45, 47, 48, 49, 54
- Deep Learning 3, 4
- Deliberation 36, 44
- Digital
 - Autonomous Organization 68
 - contract 65, 68
 - currencies 74
 - money 66, 74
- Digital Phenotyping 3
- Digitaler Humanismus 29, 40, 42
- Duhem-Quine-These 2, 7
- Echtzeit
 - daten V, 35, 38, 41, 45, 54
 - dienstleistungslandschaften 51
 - prozesse 48
- Empirie 1, 2, 4, 6, 8, 9, 10, 11, 13, 16, 17, 109
- Experience-Sampling 3
- Eye-Tracking 3, 16
- Feature Construction 6
- GDPR 65
- IDEAS project 75
- Information
 - technology 62, 63, 64, 66
- Institutional Emergence 94, 98, 99
- intergenerationale Gerechtigkeit 39
- Introspektion 13
- Kohärenz-Verzerrungen 8, 9
- konfirmatorische Informationssuche 8
- Konfundierung 6
- Korpusanalyse 14, 16
- Korrelative Studien 6
- Kritischer Rationalismus 1, 2
- Legalitätsprinzip 53
- Legitimation durch Verfahren 51, 53, 54
- Leib-Seele-Dualismus 17
- Manipulation 6, 10, 39
- Menschlicher Faktor 35, 37, 38
- Mobile Sensing 3
- Nachhaltigkeit 36, 38, 39, 40, 43, 44
- Nestedness 93, 95
- netzwerkbasierte Theorien 5, 7, 10
- Neuronale Netze 3, 4, 5, 108
- Oneness 23, 25
- Ostrom-framework 70
- Park-Apps 48
- Philosophie
 - empirisch informiert 13, 17
 - experimentelle 13, 15, 16
- Polycentric governance 61, 69, 72, 73, 74
- Privatsphäre 38
- Property rights 60, 64, 66, 73, 78, 80
- Quasi-experimentelle Studien 6
- Quick Data V, 21
- Rationale Reflexion 13

- Raumordnungsziele 46, 54
- Raumplanung 45, 51, 53, 54
- Real-time 60
 - data 60, 61, 65, 86, 92
 - information 62
 - service ecosystems V, 59, 61, 63, 65, 78
- Rebound-Effekt 39
- Reliabilität 7, 10
- Rückschaufehler 8, 11
- service
 - co-creation 70
 - contract 62, 63, 64
 - ecosystem 59, 61, 86, 89, 91, 93
- Smart City 50
- smart contract 64
- Smarte Straßen 39
- Sozialethik 41
- Soziotechnische Systeme 33, 34
- Systemisches Denken 51
- Teststärke 4
- Ubiquitous computing 46, 50, 54, 62, 64, 70, 79
- Umwelt 38, 40, 50, 51, 52
- Upper-bound-Estimation 5
- Validität 1, 4, 7, 10, 15, 16
- Value 22, 32
 - co-creation 65, 89, 96
- Variety 21, 32
- Velocity 21, 22, 25, 29, 32, 37
- Veracity 22, 25, 32
- Virtual Reality (VR) 16
- Volume 21, 29, 32, 37

Personenverzeichnis

Alfano, Mark 14, 15, 16

Baecker, Dirk 49

Bauman, Zygmunt 47

Coase, Ronald 64

Eisenberger, Iris 54

Eyal, Amir 48

Fasel, Daniel 32

Fuchs, Christian 32

Grunwald, Armin 35

Hess, Charlotte 70

Higgins, Andres 14, 15, 16

Kalmer, Marion 48, 49

Kent, William 23, 25

Kirchschläger, Peter G. 41

Krcmar, Helmut 29

Levernier, Jacob 14, 15

Luhmann, Niklas 53

Machlup, Fritz 69

Meier, Andreas 32

Nassehi, Armin 30

Nida-Rümelin, Julian 42

Ostrom, Elinor 70

Patil, DJ 31

Platt, John R. 2, 5, 10

Popper, Karl R. 5

Searle, John Rogers 61, 67

Simon, Fritz B. 51

Stalder, Felix 41

Weidenfeld, Nathalie 42

Weiser, Mark 46

Zielinski, Siegfried 47