

Vorwort

Die Idee zum Buch entstand mit Beginn der Pandemie im Frühjahr 2020. Diese führte uns schmerzlich vor Augen, dass wir in einer Daten- und Informationskrise stecken, die durch eine fehlende Interdisziplinarität verstärkt wird. Kaum ein Tag verging, an dem nicht über bundesweit steigende oder fallende Inzidenzzahlen berichtet wurde. Doch wie belastbar sind die Zahlen überhaupt, wie kommen sie zustande und worin liegen sie begründet? Wie das Entstehen eines Neukunden liegt auch eine Neuinfektion immer lokal in Raum und Zeit verankert. Das war der grundlegende Gedanke zum Buch. Alles geschieht an einem bestimmten Ort zu einem bestimmten Zeitpunkt. Die Raumzeit ist immanent, spiegelt sich in Daten wider und muss für die Erklärung von Phänomenen genutzt werden.

Großflächige Daten und Modelle haben vor allem eine Schwäche: Sie bilden immer nur einen groben Durchschnitt und berücksichtigen keine regionalen Besonderheiten oder Ausreißer. Als ein Leben im Land von Mediokristan würde es Taleb in seinem Buch „Der schwarze Schwan“ beschreiben. Aber gerade in Zeiten von Big Data haben wir die Chance, beliebig und massenhaft auftretende Phänomene als ‚feinste‘ Daten gezielt zu erfassen, sie in einen gewünschten Kontext zu bringen und entsprechend auszuwerten, um wichtige Erkenntnisse daraus zu ziehen. So hätte man beispielsweise, wie vom Deutschen Dachverband für Geoinformation gefordert, die Corona-Warn-App mit Geotracking ausstatten sollen. Parallel dazu verwies die soziologische, statistische und ökonomische Forschung immer wieder auf die Datenlücken und fehlenden Evidenzen der Pandemie. Ein fachübergreifendes Corona-Gremium entstand jedoch genauso wenig wie ein proaktives, lösungsdienendes Datenschutzkonzept.

Dabei stellen wir in unserem beruflichen Alltag immer wieder fest, dass die eigentliche Weiterentwicklung unseres datengetriebenen Lernens darin besteht, sowohl einer Fragestellung als auch ihrem Lösungsweg interdisziplinär zu begegnen. Dies beginnt bei der Datenentstehung und endet bei der audiovisuellen Verarbeitung beim Informationsempfänger. Immer neuere Technologien stellen sich diesen Herausforderungen, sodass die Verschmelzung zwischen physischer und virtueller Welt unseres Lebens bevorsteht. Der Umgang mit den Daten muss aber gelernt sein, denn eine Verzerrung bei der Erfassung oder Auswertung ist oftmals vorprogrammiert. So basieren die deutschen Inzidenzen der Pandemie auf dem Testgeschehen, das vom ‚individuellem‘ Vorgehen eines Gesundheitsamts bestimmt ist oder von der Bereitschaft bestimmter Bevölkerungsgruppen, sich testen zu lassen. Wie ist demnach das Ergebnis zu bewerten, dass in der zweiten und dritten Coronawelle vor allem sozial benachteiligte Personengruppen betroffen waren? Macht Armut etwa wirklich krank? Diese komplexen Fragen von heute und in Zukunft sind nur mit interdisziplinären Ansätzen aus Wirtschaft und Wissenschaft gleichermaßen zu beantworten. Die dazu notwendige Datenkompetenz (Data Literacy) ist von allen gefordert, denn datengetriebenes Handeln birgt gesellschaftliche Chancen und Risiken.

So basiert das Buch auf dem Gedanken, dass Data Science nicht nur interdisziplinär eingesetzt wird, sondern möglichst alle Dimensionen gesellschaftlicher, technologischer und physikalischer Aspekte bedient. Das daraus resultierende breite Spektrum zeigen die Autor:innen mit ihren Beiträgen in diesem Buch auf, denen wir an dieser Stelle einen besonderen Dank für ihre Teilnahme aussprechen. Die Einnahmen aus dem Verkauf des Buchs werden dem Nachwuchs im Bereich Data Science gespendet.

Bonn, im Mai 2022

Dr. Barbara Wawrzyniak und Michael Herter

Vorwort der GERMAN DATA SCIENCE SOCIETY

Die Menschen haben schon immer versucht, die Gegebenheiten („Data“) dieser Welt und des Weltalls zu beschreiben und zu erklären. Sie sind bestrebt, Gesetzmäßigkeiten und Strukturen der Gegebenheiten zu erkennen, die Stabilität dieser Gesetzmäßigkeiten nachzuweisen oder – bei möglichen Änderungen der Strukturen – die Dynamik der Gesetzmäßigkeiten im Zeitablauf zu bestimmen und zu prognostizieren. Denn nur so können Menschen die ihnen vorgegebenen oder selbstgesetzten Ziele erreichen.

Dabei hat sich die Universitas der Wissenschaft („Science“) immer weiter in Fakultäten und Fachbereiche (Disziplinen) spezialisiert, um die verschiedenen Aspekte und Details der Gegebenheiten umfassender und präziser verstehen zu können. Die Erkenntnisse über die Gegebenheiten werden je nach Disziplin in Zeichen und Bildern, in Worten und Sprachen, in Noten und Kompositionen, in Zahlen und Formeln sehr unterschiedlich abgebildet und fixiert. Der Ruf nach Annäherung von Disziplinen, nach Interdisziplinarität und gemeinsamer Nutzung von Forschungsstrategien und Forschungsv erfahren ist deshalb nicht zu überhören.

Um die langsame Geschwindigkeit und den beschränkten Umfang der menschlichen Erkenntnisprozesse zu erhöhen, wurden in der Informatik Techniken (z. B. KI) entwickelt, die die Hypothesen und mathematischen Modelle von Gesetzmäßigkeiten der Empirie und ihre Überprüfung durch statistische Methoden schneller und umfassender realisieren können. Die Kooperation von Mathematikern, Informatikern und Statistikern hat letztlich zur Entwicklung des interdisziplinären Fachs „Data Science“ geführt.

Ende des Jahrs 2018, kurz vor Ausbruch der Coronavirus-Pandemie 2019, wurde von Mathematikern, Informatikern und Statistikern die wissenschaftliche Gesellschaft GERMAN DATA SCIENCE SOCIETY (GDS e. V.) in München gegründet. Sie versucht, die unterschiedlichen Ziele und Schwerpunkte der Aus- und Weiterbildung in Data Science an Hochschulen und Universitäten zu dokumentieren und durch Vortrags- und Tagungsveranstaltungen zu ergänzen und zu erweitern. Um neue Kontaktmöglichkeiten sowie die Vernetzung der Data Scientists und der fördernden Unternehmen vor Ort zu intensivieren, aber auch um die Anwendungsorientierung der Aus- und Weiterbildung in Data Science zu fördern, wurden regionale Schwerpunkte gebildet.

Der Vorstand der GERMAN DATA SCIENCE SOCIETY begrüßt deshalb sehr die Initiative seines fördernden Mitglieds infas 360 GmbH aus der Region Köln/Bonn, die Ziele und Aktivitäten der GDS e. V. beispielhaft durch diese umfangreiche und vielschichtige Veröffentlichung „Neue Dimensionen in Data Science“ zu erläutern und zu dokumentieren.

München, im Mai 2022

Prof. Dr. Elmar Helten

