
Vorwort

»Die betriebliche Informationsverarbeitung unterliegt nach wie vor einer zügigen Entwicklung« schrieben wir an dieser Stelle vor exakt zehn Jahren. »Zuweilen macht sie sogar große Sprünge«, möchte man heute spontan ergänzen. Einer dieser »Sprünge« hat es unter dem Schlagwort »Big Data« wie bislang kaum ein zweites IT-Thema weit über die Fachkreise hinaus bis auf die Titelseiten der großen Tages- und Wochenzeitungen geschafft: Neue Datenquellen wie Social Media und mobile Endgeräte bieten in bisher nicht gekannter Detailtiefe Einblicke in das Verhalten und die Wünsche der Kunden. Gleichzeitig können diese zusätzlichen Datenmengen durch große Fortschritte in der Datenverarbeitungstechnik systematisiert gespeichert und analysiert werden.

»Was immer sie produziert, eine Organisation ist in erster Linie eine Fabrik von Einschätzungen und Entscheidungen« – so *Daniel Kahnemann*, Nobelpreisträger für Wirtschaftswissenschaften. Dass die systematische Analyse kleiner und großer Datenmengen Wissen von hervorragender Qualität erzeugen kann, das in vielen Fällen sprunghaft bessere Einschätzungen und Entscheidungen ermöglicht, ist mittlerweile in vielen Chefetagen angekommen. Dasselbe gilt für die Tatsache, dass man bei der vorherrschenden Datenfülle für deren Auswertung an hochautomatisierten Data-Mining-Verfahren häufig nicht vorbeikommt. Dieser Umstand hat dem Data Mining zu neuer Popularität und zu einem spürbaren Aufschwung in der betrieblichen Praxis verholfen, wie wir in zahlreichen Kundenprojekten der letzten Jahre erfahren durften.

Trotz der Erschließung immer neuer Anwendungsfelder bleibt das Customer Relationship Management eine Kerndomäne der Datenanalyse. Dem Titel dieses Buches entsprechend konzentrieren wir uns daher weiterhin auf die Analyse von Kundendaten. Denn die Ansprüche der Kunden steigen in vergleichbarem Maße wie die Datenmengen, die im Laufe einer Kundenbeziehung anfallen. Viele Kunden erwarten mittlerweile eine hochindividualisierte Betreuung und eine Servicequalität, die nur durch Kenntnis der Kundenhistorie und regelmäßige Datenauswertungen wirtschaftlich erreichbar sind. Vielen Unternehmen gelingt es noch nicht, diesem Anspruch auch in der Weise gerecht zu werden, wie man ihn als Kunde verstanden wissen möchte.

Was hat sich geändert in der zweiten Auflage? Die 2005 vorgestellten Grundkonzepte und Fallstudien sind in ihrem Kern stabil und haben sich bewährt. Veränderungen gab es an vielen Stellen lediglich im Detail – einige neue Themen sind hinzugekommen. Dabei haben wir die zahlreichen Leserrückmeldungen und -wünsche, für die wir uns herzlich bedanken, soweit möglich berücksichtigt.

So haben wir den Themen Social Media und Big Data neue, eigene Abschnitte in Teil A gewidmet. Das Grundgerüst des Buches ändert sich dadurch nicht: Social-Media-Plattformen sind letztlich eine weitere Klasse von Kommunikationskanälen. Einige der in Teil C vorgestellten Anwendungsfälle zeigen deutlichen »Big-Data-Charakter«: So sind etwa die Menge und die Erzeugungsgeschwindigkeit von Kassendaten und die Extraktion strukturierter Informationen aus den entsprechenden Textdateien typische Merkmale »großer Daten«. In allen Kapiteln haben wir an geeigneten Stellen Hinweise eingeflochten, um Eigenheiten und neue Möglichkeiten im Kontext dieser Entwicklungen aufzuzeigen.

Bewusst verzichtet haben wir hingegen – wie schon in der ersten Auflage – auf detaillierte mathematische Erläuterungen von Analysealgorithmen oder von Speichertechniken, wenn sie nicht für einen konkreten Analysefall im Teil C erforderlich sind. Zum einen können wir im Rahmen dieses Buches nicht auf alle Aspekte in gebührender Tiefe eingehen, zum anderen wollen wir Konzepte präsentieren, die nicht nur für kurze Zeit gültig sind. Stattdessen verweisen wir auf die zahlreichen Veröffentlichungen zu diesen Themen.

In Teil B wurde das Datenschutzkapitel der aktuellen Rechtslage angepasst. In Teil C haben wir eine neue Fallstudie zur Betrugserkennung aufgenommen; die beiden Analysefälle zum Verbundkaufverhalten haben wir hingegen gestrafft und zu einer Fallstudie zusammengeführt.

Die Literaturhinweise wurden vollständig aktualisiert und leicht gekürzt. Da wir der Überzeugung sind, dass Aussagen durch wiederholtes Zitieren und »Abschreiben« nicht richtiger, sondern mitunter eher verfälscht werden, verweisen wir wo sinnvoll auf die Originalliteratur, etwa aus der Pionierzeit des Data Mining in den 1990er-Jahren. Sofern Neuauflagen existieren und diese die relevanten Aussagen weiterhin enthalten (was nicht immer der Fall ist), zitieren wir die jeweils letzte verfügbare Ausgabe. Die am Ende jedes Kapitels platzierte Liste mit kommentierten Literaturhinweisen gibt Anregungen zur Vertiefung oder Ergänzung ausgewählter Themen des behandelten Stoffes. Zur besseren Orientierung ist die Liste in der Reihenfolge des Vorkommens der betreffenden Themen im Text geordnet.

Unser Dank gilt dem dpunkt.verlag – besonders unserer Lektorin Christa Preisendanz und unserer Gutachterin Prof. Dr. Heidi Heilmann – für das Vertrauen und die Geduld bei der Erstellung der zweiten Auflage. Wir danken herzlichst unseren Familien, Freunden und Kollegen, die während der Überarbeitung des Manuskripts nicht selten auf unsere Anwesenheit verzichten mussten und

dabei großes Verständnis gezeigt haben. Ohne dieses Entgegenkommen wäre die zweite Auflage dieses Buches nicht möglich gewesen.

Wir wünschen Ihnen nun wieder einmal viel Spaß bei der Lektüre und beim Experimentieren mit neuen Datenanalysen!

Peter Neckel & Bernd Knobloch
Schwarzenbek und Bamberg/Ködnitz, Juli 2015