



DATA

kristin WEBER
christiana KLINGENBERG

GOVERNANCE

Der Leitfaden für die Praxis

HANSER



Bleiben Sie auf dem Laufenden!

Unser **Computerbuch-Newsletter** informiert Sie monatlich über neue Bücher und Termine. Profitieren Sie auch von Gewinnspielen und exklusiven Leseproben. Gleich anmelden unter:

www.hanser-fachbuch.de/newsletter



Kristin Weber
Christiana Klingenberg

Data Governance

Der Leitfaden für die Praxis

HANSER

Die Autorinnen:

Prof. Dr. Kristin Weber, Kleinostheim

Dr. Christiana Klingenberg, Karlsruhe

Alle in diesem Buch enthaltenen Informationen, Verfahren und Darstellungen wurden nach bestem Wissen zusammengestellt und mit Sorgfalt getestet. Dennoch sind Fehler nicht ganz auszuschließen. Aus diesem Grund sind die im vorliegenden Buch enthaltenen Informationen mit keiner Verpflichtung oder Garantie irgendeiner Art verbunden. Autorinnen und Verlag übernehmen infolgedessen keine juristische Verantwortung und werden keine daraus folgende oder sonstige Haftung übernehmen, die auf irgendeine Art aus der Benutzung dieser Informationen – oder Teilen davon – entsteht.

Ebenso übernehmen Autorinnen und Verlag keine Gewähr dafür, dass beschriebene Verfahren usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Buch berechtigt deshalb auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Bibliografische Information der Deutschen Nationalbibliothek:

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

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form (Fotokopie, Mikrofilm oder ein anderes Verfahren) – auch nicht für Zwecke der Unterrichtsgestaltung – reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

© 2021 Carl Hanser Verlag München, www.hanser-fachbuch.de

Lektorat: Sylvia Hasselbach

Copy editing: Walter Saumweber, Ratingen

Umschlagdesign: Marc Müller-Bremer, www.rebranding.de, München

Umschlagrealisation: Max Kostopoulos

Titelmotiv: © Max Kostopoulos

Gesamtherstellung: Kösel, Krugzell

Ausstattung patentrechtlich geschützt. Kösel FD 351, Patent-Nr. 0748702

Printed in Germany

Print-ISBN: 978-3-446-46388-2

E-Book-ISBN: 978-3-446-46674-6

E-Pub-ISBN: 978-3-446-46675-3

Inhalt

1	Einleitung	1
2	Begriffe und Grundlagen	7
2.1	Von Zeichen zu Wissen	7
2.2	Arten von Daten	8
2.3	Stammdaten	10
2.4	Open Data und externe Daten	12
2.5	Metadaten	14
2.6	Datenmodellierung	16
2.7	Governance und Organisationsgestaltung	18
3	Data Governance	23
3.1	Begriff	23
3.2	Data Governance Frameworks	25
3.2.1	The Non-Invasive Data Governance Framework	25
3.2.2	Das DGI Data Governance Framework	27
3.2.3	SAS® Data Governance Framework	29
3.2.4	DAMA-DMBOK2	31
4	Das qualitätsorientierte Data Governance Framework	35
4.1	Data Governance – a Journey	35
4.2	Zielsetzung und Fokus des Frameworks	37
4.3	Überblick über die Data-Governance-Handlungsfelder	39
4.4	Handlungsfeld der strategischen Ebene	41
4.5	Handlungsfelder der organisatorischen Ebene	44
4.5.1	Controlling	44

4.5.2	Organisation des Datenmanagements	48
4.5.2.1	Rollenmodell	48
4.5.2.2	Koordination der Rollen	51
4.5.2.3	Data Quality Awareness	53
4.5.3	Datenmanagement-Prozesse	55
4.5.3.1	Datenproduktions-Prozesse	55
4.5.3.2	Datenqualitätsmanagement-Prozesse und -Methoden	57
4.6	Handlungsfelder auf Ebene der Informationssysteme	59
4.6.1	Datenarchitektur	59
4.6.1.1	Business Data Dictionary	59
4.6.1.2	Datenobjektmodell	61
4.6.1.3	Datenhaltungs- und Datenverteilungsarchitektur ...	62
4.6.2	Systemarchitektur	65
5	Data-Governance-Rollen	69
5.1	Exekutive Ebene	69
5.2	Strategische Ebene	70
5.3	Taktische Ebene	73
5.4	Operative Ebene	78
6	Datenqualität	81
6.1	Begriffsabgrenzung	81
6.2	Dimensionen der Datenqualität	82
6.3	Konzepte zur Datenqualität	85
6.3.1	Der „Fitness for use“-Ansatz	85
6.3.2	Die sechs primären Datenqualitäts-Dimensionen der DAMA UK	87
6.4	Herausforderung Datenqualität in der Praxis	89
6.4.1	Praktische Bedeutung von Datenqualität	89
6.4.2	Informationsprobleme	92
6.5	Messen und Bewerten von Datenqualität	95
6.5.1	Messen der Datenqualität	95
6.5.2	Bewerten der Datenqualität	97
6.6	Kosten schlechter Datenqualität	99
6.7	Qualität von Metadaten	100

7	Methoden, Konzepte und Tools für Data Governance	103
7.1	Überblick	103
7.2	Strategische Tools	105
7.2.1	Data Policy	105
7.2.2	Entwicklung einer Datenstrategie	108
7.3	Tools auf Ebene der Organisation	116
7.3.1	Ursachen-Wirkungs-Diagramm	116
7.3.2	Datenqualität messen	121
7.3.2.1	Datenqualitäts-Regelwerk	121
7.3.2.2	Prüfablauf und Datenqualitäts-Regelbaum	126
7.3.2.3	Datenqualitäts-Scorecard	131
7.3.3	Datenmanagement-Organisation umsetzen	134
7.3.3.1	Unternehmensspezifische Ausprägung der Datenmanagement-Organisation	134
7.3.3.2	Umsetzung der Datenmanagement-Organisation	140
7.3.3.3	Organisatorische Hilfsmittel	143
7.3.4	Beschreibung der wesentlichen Datenproduktions-Prozesse	146
7.3.5	RACI-Matrizen	150
7.4	Tools auf Ebene der Informationssysteme	152
7.4.1	Business Data Dictionary	153
7.4.2	Datenqualitäts-Tools	156
7.4.3	Auswahl von Data Governance Tools	158
8	Anwendungsszenarien von Data Governance	163
8.1	Etablierung von Quality Gates in der Stammdatenproduktion	163
8.2	Die DSGVO als Treiber für ein Rollen- und Berechtigungskonzept	166
8.3	Einführung einer DQ-Scorecard im Marketingbereich	171
9	Zusammenfassung und Ausblick	177
	Literaturverzeichnis	183
	Die Autorinnen	189
	Index	191

1

Einleitung

Daten standen noch nie so hoch im Kurs wie heute. Noch nie war der Bedarf an Praxiswissen im Umgang mit Daten gefragter. Neue Trends und Technologien für Daten entwickeln sich immer schneller. Und es steigt der Bedarf, alle Möglichkeiten, die es heute bereits gibt und morgen geben wird, auszuschöpfen. In fast allen Bereichen der Wirtschaft und des privaten Lebens spielen Daten eine immer größere Rolle. Es gibt kein Unternehmen, welches nicht in irgendeiner Weise Daten verarbeitet. Private Nutzer teilen ihre Daten in sozialen Netzwerken, bloggen, shoppen online, bezahlen mit dem Smartphone oder investieren in digitales Geld.

Die Trends der Datenverarbeitung in der Zukunft deuten sich schon an: Diskussionen über die Abschaffung von Bargeld laufen bereits [Deut18], ebenso werden die Möglichkeiten von selbstsouveränen Identitäten (SSI), also der Verwaltung der eigenen Identität im Netz, immer umfangreicher (z.B. [Emil20]). Und mit den wachsenden Möglichkeiten, Daten zu erfassen, zu halten und zu verarbeiten, steigt unweigerlich der Bedarf nach Richtlinien, Hinweisen, Transparenz, Verantwortlichkeiten und Systemen, die den Umgang mit diesen Daten regeln.

Und spätestens dann, wenn dieser Rahmen für das Management der Daten eingefordert wird, kippt die Stimmung oft von Euphorie in Richtung Ernüchterung. Wenn deutlich wird, dass es diese innovativen Ansätze nicht umsonst gibt. Fällt der Begriff Data Governance als Zusammenfassung all der Aktivitäten, die notwendig sind, um den vollen Wert der Daten auszuschöpfen, sind folgende Aussagen zu hören:

- Data Governance ist ein Kostentreiber,
- sie erfordert ein hohes Maß an zusätzlicher Bürokratie,
- sie benötigt neue Stellen und mehr Mitarbeiter,
- sie bremst Datenmanagement-Projekte aus,
- sie ist ohne zusätzliche Software nicht umsetzbar,
- und die Ergebnisse sind erst nach langer Zeit sichtbar.

Diese Argumente sind nur Beispiele für die skeptische Grundhaltung gegenüber Data Governance. Die Kunst ist es, einen Anreiz zu schaffen und gute Gründe zu nennen, warum es sich dennoch lohnt, in Data Governance zu investieren.

Die gute Nachricht und die Herausforderung

In aller Regel befinden sich die Unternehmen in Bezug auf Data Governance nicht auf der grünen Wiese. Bei genauer Betrachtung gibt es in verschiedenen Bereichen Aktivitäten, die sich Data Governance durchaus zuordnen lassen. Das kann eine bereits implementierte Softwarelösung für bestimmte Datenqualitäts-Prüfungen sein. Oder es gibt bereits Kollegen, die bei Fragen zu bestimmten Daten gute Ratgeber sind. Nicht selten gibt es Customer-Relationship-Management- oder Enterprise-Ressource-Planning-Systeme, für die Datenpflege-Prozesse bereits beschrieben wurden. Je mehr man sucht und sich durchfragt, desto mehr Hinweise auf Data Governance wird man im Unternehmen finden. Und die verschiedenen Aktivitäten sind oft seit längerer Zeit etabliert und auf das Management der Daten abgestimmt. Das heißt, in vielen Bereichen läuft es bereits und das Rad muss nicht neu erfunden werden. Das ist die gute Nachricht.

Tabelle 1.1 stellt die Auswirkungen von Datenmanagement ohne und mit Data Governance gegenüber.

Tabelle 1.1 Ziele von Data Governance (in Anlehnung an [Kare07, S. 2])

Ohne Data Governance	Mit Data Governance
Mangel an fachlicher Verantwortung	Unterstützung und Verantwortung durch das Top Management
Datenmanagement mit geringer Priorität	Unternehmensweites Management des „Anlagegutes“ Daten
Ziele des Datenmanagements haben geringere Priorität in IT-Projekten	IT-Projekte mit Auswirkung auf kritische Daten finden unter Beteiligung der Data Stewards statt
Geschäfts- und Fachbereiche ignorieren übergreifende Auswirkungen der Datenpflege	Data-Governance-Gremien stellen übergreifende Abstimmung aller Datenmanagement-Initiativen sicher
Inkonsistente Geschäftsprozesse, Erfassungsrichtlinien und Datenmodelle	Einführung und Durchsetzung von Best Practices, inklusive standardisierter Datenmodelle, Definitionen, Regeln und Geschäftsprozesse

Die Herausforderung ist allerdings, dass bereits vorhandene Aktivitäten selten koordiniert und bereichsübergreifend stattfinden. Kollegen, die sich bereits den Ruf eines „Datenexperten“ erarbeitet haben, finden das Expertentum kaum in ihrer Stellenbeschreibung. Neue Anforderungen an die Daten, z. B. in Form von neuen Reports, münden in teilweise zeitaufwendiger und fehleranfälliger manueller Datenaufbereitung, die im schlimmsten Fall weder nachvollziehbar noch reproduzierbar ist. Es ist nicht bekannt, wer was in welchem Bereich für die Datenoptimierung bereits macht. Oder wo welche Datenpflege-Prozesse bereits implementiert sind und gut laufen. Es kommt zu Missverständnissen und doppelten bzw. mehrfachen Arbeiten, wenn die gleichen Aktivitäten an verschiedenen Stellen durchgeführt werden. Anforderungen an die Daten sind nur aus dem eigenen Bereich bekannt. Die Anforderungen anderer Bereiche sind ungeklärt und werden somit bei der Datenerfassung und Verarbeitung (unwissentlich) ignoriert. Im schlimmsten Fall werden die Datenoptimierungen von Mitarbeitern des einen Bereichs durch Mitarbeiter des anderen Bereichs überschrieben und somit zunichte gemacht. Die Koordination fehlt, und damit die Transparenz

und das konsolidierte Wissen über die Möglichkeiten, was mit den Unternehmensdaten heute schon gemacht wird und morgen noch getan werden soll. Die Herausforderung und Aufgabe ist also, eine übergreifende Struktur in das vermeintliche Chaos zu bringen und so durch Data Governance die notwendige Effizienz in das Datenmanagement zu bringen.

Quo vadis?¹

Die Einführung von Data Governance ist nicht zu unterschätzen. Sehr schnell sind wieder die oben genannten Statements zu hören und das macht es nicht einfacher, Verbündete im eigenen Unternehmen zu finden. Denn Data Governance bedeutet auch, sich mit neuen Möglichkeiten des Datenmanagements auseinanderzusetzen. Aktivitäten, die sich heute mit den Daten nicht durchführen lassen, sollen morgen möglich sein. Das bedeutet Veränderung. Veränderung braucht ein starkes Netzwerk, gute Argumente und einen langen Atem. Im Fall von Data Governance bedeutet Veränderung, dass Unternehmensdaten in den Fokus rücken und deren Nutzung dem wirtschaftlichen Erfolg des Unternehmens unterstellt ist. Und dieser Nutzungszweck wird konsequent auf alle Bereiche des Datenmanagements über das gesamte Unternehmen hinweg angewendet und in vielen kleinen Maßnahmen umgesetzt. Anders gesagt, eine Datenstrategie und Maßnahmen zur Umsetzung werden definiert.

Ein einheitliches Verständnis der Strategiemassnahmen und deren Einordnung in ein Data-Governance-Rahmenwerk helfen bei der Bewältigung der genannten Herausforderungen ungemein. Voraussetzung ist, eine gemeinsame Sprache zu sprechen und das gleiche Grundverständnis von Data Governance mit seinen unterschiedlichen Handlungsfeldern zu haben. Die gemeinsame Sprache im Kontext von Data Governance bezieht sich auf Komponenten eines Data Governance Frameworks, auf die Aufgaben und Verantwortlichkeiten der definierten Rollen sowie auf die Bedeutung und den Kontext der Daten an sich, die z. B. in Business Data Dictionaries abgebildet werden. Fehlen diese gemeinsame Sprache und damit das gemeinsame Verständnis, kommt es zu Missverständnissen und Reibungsverlusten und man arbeitet den Skeptikern zu, die sich zu den oben genannten Aussagen verleiten lassen.

Das vorliegende Buch unterstützt dabei, bereits laufende Data-Governance-Aktivitäten in ein Framework einzuordnen. So können diese Aktivitäten besser in einen gemeinsamen Kontext gestellt und auf ein gemeinsames Ziel hin koordiniert werden. Nicht nur die Symptome und Auswirkungen von „schlechten“ Daten sollen angegangen werden, sondern nach deren Ursachen gesucht und diese in weiteren Schritten behoben werden.

Ist diese Einordnung geschehen, kann auch viel einfacher identifiziert werden, an welchen Stellen von Data Governance noch Unterstützung benötigt wird. Anforderungen an interne Projekte oder externe Beratungen können präzise formuliert werden. Eine Unterhaltung auf Augenhöhe wird möglich.

Die hohe Relevanz von Datenmanagement in Kombination von Data Governance wird in einer aktuellen Studie [BuKn20] bestätigt. Demnach gehört Data Governance neben Datenstrategie und Datenarchitektur zu den Topthemen aus der Perspektive Datenmanagement. Die gleiche Studie zeigt auch, dass sich Data Governance in den meisten Unternehmen in

¹ Aus dem Lateinischen für „Wohin gehst du?“

einem sehr frühen Implementierungsstadium befindet. Einige Unternehmen planen zumindest, Data Governance einzuführen.

Damit ist jetzt der richtige Zeitpunkt, das vorliegende Buch zu lesen und mit einfachen und pragmatischen Aktivitäten Data Governance umzusetzen und das Datenmanagement zum Erfolg zu führen.

Wie dieses Buch zu lesen ist

Das vorliegende Buch kann auf unterschiedliche Weise zum Einsatz kommen. Wer an der theoretischen Einführung zu Data Governance interessiert ist und sich aus den vorgestellten und selbst entwickelten Frameworks selbst ein Bild machen will, der kann das Buch von vorne bis hinten durchlesen.

Wer sich nur für einzelne Abschnitte interessiert, kann das Buch auch abschnittsweise lesen. Die Themen sind in den Kapiteln so aufbereitet und beschrieben, dass sie (fast) unabhängig voneinander gelesen werden können. An manchen Stellen wird auf andere Abschnitte verwiesen und bei Bedarf kann der Wissensdurst dann an den entsprechenden Stellen gestillt werden.

Für die Praktiker, die nach einer schnellen Möglichkeit für die Umsetzung der einzelnen Handlungsfelder suchen, ist *Kapitel 7* gedacht. Hier werden ganz viele Methoden, Konzepte und Tools vorgestellt, mit denen sich Aspekte von Data Governance einfach verwirklichen lassen. Und die allermeisten Tools lassen sich mit den üblichen Office-Programmen umsetzen.

Es gibt dazu noch verschiedene Informationen, die in Kästen stehen. Hier wird zwischen Praxistipps, Hinweisen und Fallbeispielen unterschieden. Besonders interessant sind die Fallbeispiele, denn sie zeigen, wie in Projekten konkrete Data-Governance-Themen adressiert und umgesetzt wurden. Bei den Hinweisen handelt es sich um ergänzende Hinweise zu den Themen, die in Abhängigkeit des Kontexts und der Umsetzung zum Tragen kommen können oder auch nicht.

Für wen dieses Buch hilfreich ist

Dieses Buch enthält Hinweise für die Umsetzung von Data Governance in der Praxis für alle, die ...

- ... bei ihren täglichen Aufgaben über Ungereimtheiten bei Daten stolpern und überlegen, was die Ursache ist und wie diese behoben werden kann. Oft haben diese Personen die Rolle eines Data Stewards inne und sind (implizit) für die Qualität von Daten verantwortlich.
- ... im Unternehmen bekannte Herausforderungen bei DatenmanagementThemen fachlich einordnen möchten, um die richtigen Maßnahmen zur Verbesserung der Situation einzuleiten.
- ... oft mit großen (Stamm-)Datenbeständen aus unterschiedlichen Quellen arbeiten und diese zur Beantwortung bestimmter Fragestellungen untersuchen. Bei der Zusammenführung der Daten zur Auswertung kommt es immer wieder zu Herausforderungen, deren Auflösung sehr viel Zeit in Anspruch nimmt und manuelle Anpassungen an den Daten erfordern.

- ... in verantwortlicher Position den effizienten Einsatz von Unternehmensdaten sicherstellen müssen, z. B. Marketing Manager oder E-Commerce-Verantwortliche.
- ... ein neues Stammdatenmanagement (MDM) System einführen wollen oder eingeführt haben, und nun sicherstellen müssen, dass die Stammdatenpflege nach einheitlichen Prozessen durchgeführt wird.
- ... die Rolle des Chief Data Officers innehaben und für das Management der Unternehmensdaten verantwortlich sind. Sie geben ihren Kollegen Orientierung bei der Umsetzung von Data Governance und stellen sicher, dass die Daten ein echter Produktionsfaktor sind.
- ... neu in einer verantwortlichen Rolle für Daten sind und eine Orientierungshilfe und Best-Practice-Angebote suchen.
- ... beratend tätig sind und immer wieder mit Data-Governance-Herausforderungen konfrontiert werden und praktische Hinweise für einen möglichen Lösungsweg suchen.

Aber auch alle anderen Personen, die sich nicht in den oben genannten Situationen wiederfinden, sind eingeladen, dieses Buch zu lesen. Zu Bedenken ist bei der Lektüre, dass es sich um ein Handbuch für die Praxis handelt und der Schwerpunkt auf der praktischen Anwendbarkeit und Umsetzung liegt.

Wie dieses Buch aufgebaut ist

Das Buch beschreibt in den folgenden Kapiteln die verschiedenen Aspekte von Data Governance. Dabei wird sowohl auf theoretische Grundlagen eingegangen als auch auf Hinweise für die Praxis. Wo immer es möglich ist, sind Fallbeispiele angeführt. Die Struktur ergibt sich wie folgt.

Kapitel 2, „Begriffe und Grundlagen“, erläutert die wesentlichen Grundlagen zu Daten und Informationen. Es werden die verschiedenen Datenarten beschrieben. Zudem werden die Themen Open Data und externe Daten adressiert – beides Themen, die heute bereits und in der Zukunft noch viel mehr an Relevanz gewinnen werden. Spricht man über Daten, müssen Metadaten unweigerlich genannt werden. Mit Metadaten legt man in der Regel die Grundlage für Datenqualität, da sie beschreiben, wie Daten beschaffen sein müssen. Wer Datenmanagement betreibt, kommt um Datenmodellierung nicht herum. Damit ist nicht nur die Strukturierung von Attributen in Entitäten und Objekten gemeint, sondern auch die Beschreibung und Dokumentation des Datenmodells in Form von Glossaren. Im Abschnitt zu Governance und Organisationsgestaltung geht es um Fragen zur Organisation, Kompetenzarten, Verantwortung und Kongruenzprinzip. All das sind Punkte, die für die Definition einer funktionierenden Data-Governance-Organisation bekannt sein sollten.

In **Kapitel 3, „Data Governance“**, werden verschiedene, in der Wissenschaft und Praxis bekannte Definitionen von Data Governance vorgestellt. Zusätzlich werden vier Data Governance Frameworks vorgestellt. Jedes dieser Frameworks hat einen individuellen Schwerpunkt. Damit wird deutlich, dass die Entwicklung und der Einsatz eines Data Governance Frameworks abhängig vom jeweiligen Unternehmens- bzw. Datenmanagement-Kontext ist. Es gibt kein Richtig und kein Falsch.

Kapitel 4, „Das qualitätsorientierte Data Governance Framework“, stellt das gleichnamige Framework vor. Dieses Framework beschreibt Struktur und Rahmen für die praxisorientierte Umsetzung in den folgenden Abschnitten. Grundsätzliche Zielsetzung des Frame-

works ist es, Daten in hoher Qualität für alle Nutzenden bereitzustellen. Es werden sechs Handlungsfelder auf den Ebenen Strategie, Organisation/Prozesse und Informationssysteme beschrieben: Datenstrategie, Controlling, Organisation, Datenprozesse, Datenarchitektur und Systemarchitektur. Datenqualität ist als Querschnittsaufgabe definiert. Die Handlungsfelder oder Gestaltungsbereiche zeigen als eine Art Best Practice, woran Unternehmen denken sollten, wenn sie Data Governance definieren und umsetzen wollen.

Kapitel 5, „Data-Governance-Rollen“, beschreibt eines der Kernelemente des qualitätsorientierten Data Governance Frameworks: die Struktur der Datenmanagement-Organisation und die Ausgestaltung der Rollen. Letztendlich übernehmen Inhaber von Rollen konkrete Aufgaben und Verantwortlichkeiten des Datenmanagements. Umso wichtiger ist eine gute Basis, wie die Rollenorganisation aussehen könnte. Die verschiedenen hierarchischen Ebenen werden adressiert sowie Möglichkeiten der Kommunikation zwischen diesen Ebenen. Das Rollenmodell ist flexibel gestaltet, sodass es auf Unternehmen jeder Größe anpassbar ist. Wer im eigenen Unternehmen nach Möglichkeiten der Definition und Besetzung von Datenmanagement-Rollen sucht, wird hier hilfreiche Informationen finden.

Im Datenmanagement nimmt die Qualität der Daten eine herausragende Rolle ein. Der Name des eingeführten qualitätsorientierten Data Governance Frameworks gibt bereits einen Hinweis darauf. **Kapitel 6, „Datenqualität“**, gibt eine ausführliche Beschreibung von Datenqualität und stellt zwei Konzepte dazu vor. Ebenso wird auf die Bedeutung von Datenqualität in der Praxis eingegangen mit Hinweisen, wie die Datenqualität zum einen gemessen und zum anderen bewertet werden kann. Auch wird die Frage adressiert, welche Kosten schlechte Datenqualität mit sich bringt. Hinweise auf die Qualität von Metadaten werden gegeben.

Das größte Kapitel des Buchs, **Kapitel 7**, beschreibt diverse **„Methoden, Konzepte und Tools“** für die praktische Umsetzung von Data Governance. Für jedes der in *Kapitel 4* vorgestellten Handlungsfelder werden verschiedene Möglichkeiten vorgestellt, diese ganz oder teilweise in der Praxis umzusetzen. Auf die Empfehlung von zusätzlichen Softwarelösungen wird weitestgehend verzichtet. Vielmehr lassen sich die meisten Methoden mit den üblichen Office-Programmen umsetzen. Letztendlich spielen die Idee und die Methode der Umsetzung eine größere Rolle als die Anwendung eines spezifischen Toolsets. Das bedeutet, dass dieser Abschnitt für die Anwendung in der Praxis geschrieben wurde. Ideen für mögliche Templates und Fallbeispiele ergänzen die unterschiedlichen Methoden. Die vorgestellten Ansätze können sehr leicht auf die individuellen Anforderungen angepasst werden.

In **Kapitel 8, „Anwendungsszenarien“**, werden drei Anwendungsszenarien aus der Praxis vorgestellt, bei denen verschiedene Methoden und Konzepte aus *Kapitel 7* zum Einsatz kommen. Es geht um die Etablierung von Quality Gates in der Stammdatenproduktion, die DSGVO als Treiber für ein Rollen- und Berechtigungskonzept sowie die Einführung einer Data Quality Scorecard im Marketing.

Kapitel 9, „Zusammenfassung und Ausblick“, fasst das Buch kurz zusammen. Es beschreibt Themen, die bei der Weiterentwicklung des qualitätsorientierten Data Governance Frameworks berücksichtigt werden können.

2

Begriffe und Grundlagen

Dieses Kapitel legt die Grundlagen für die weiteren Ausführungen zu Data Governance. Es erklärt die Definition des Begriffs Daten und beschreibt verschiedene Arten von Daten. Ein kurzer Abriss über Datenmodellierung zeigt Begriffe, die häufig im Zusammenhang mit Data Governance gebraucht werden. Da Governance auch viele Gemeinsamkeiten mit Organisationsgestaltung hat, werden zur Organisationsgestaltung ebenfalls grundlegende Konzepte erläutert.

■ 2.1 Von Zeichen zu Wissen

Das der Semiotik entlehnte Ebenenmodell (siehe Bild 2.1) erklärt den Zusammenhang zwischen Zeichen, Daten, Informationen und Wissen. **Zeichen** und Signale (technisch codierte Zeichen) werden in der Kommunikation verwendet. Beispielsweise die Zeichen „.“, „0“, „2“, „8“, „M“, „r“, „z“ und „ä“. Zeichen stehen in keinem besonderen Zusammenhang und sind wertfrei zu betrachten. Auf der syntaktischen Ebene werden die verschiedenen Zeichen miteinander in Verbindung gebracht. Sie werden nach bestimmten Regeln in eine formale Struktur überführt und darin eingebunden. Durch diesen Schritt entstehen **Daten**, die einer einheitlichen Syntax gehorchen und dieser Syntax entsprechend mit technischen Mitteln verarbeitet werden können. Ein Beispiel ist „8. März 2020“. Der Datenbegriff ist der (informations-)technischen Ebene der Generierung, Verarbeitung, Weiterleitung und Speicherung von Zeichen zugeordnet.

Auf der semantischen Ebene erhalten die Daten zusätzlich eine inhaltliche Bedeutung für den Empfänger der Daten. Damit werden die Daten zur **Information**. Im o.g. Beispiel handelt es sich um das Datum, an dem dieser Text überarbeitet wurde. Daten sind also zweckneutrale Fakten, die zu Informationen werden, wenn ihnen im unternehmerischen Kontext eine Bedeutung zugewiesen wird. Der Informationsbegriff ist subjektiv. Was für die eine Person eine wichtige Information darstellt, schätzt eine andere Person als nebensächlich oder gar gänzlich unbedeutend ein.

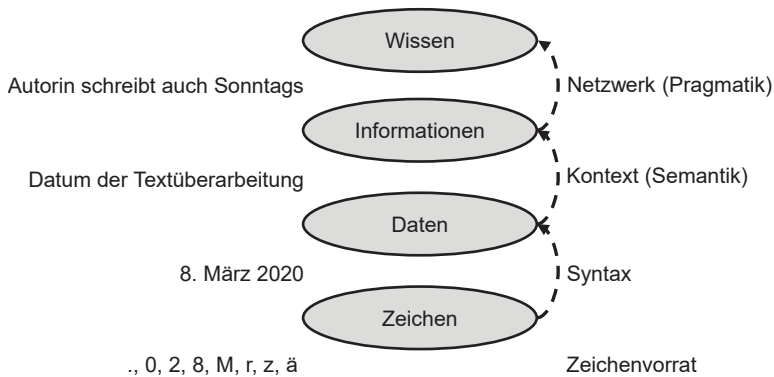


Bild 2.1 Ebenenmodell der Semiotik (in Anlehnung an [Krcm15, S.12])

Auf der obersten Ebene des Modells steht das **Wissen**. Informationen werden zu Wissen, wenn sie miteinander vernetzt werden bzw. interpretiert werden. Um auf das vorhergehende Beispiel zurückzukommen, könnte die Information des Datums der Überarbeitung dieses Textes mit einer weiteren verknüpft werden, nämlich, dass der 8. März 2020 ein Sonntag war. Somit könnte man daraus das Wissen ableiten, dass die Autorin auch am Sonntag schreibt – also vermutlich außerhalb ihrer regulären Arbeitszeit.

Wissen ermöglicht es Personen, Aufgaben in größere Sinnzusammenhänge, Entwicklungen, Arbeitsprozesse etc. einzuordnen [Klot11]. Aus ihrer Erfahrung heraus kennen sie erfolgversprechende Herangehensweisen für die Aufgabenbearbeitung und wählen dafür passende Methoden und Techniken aus. Durch Wissen können Personen die ihnen zur Verfügung stehenden Daten sichten, auswählen und als Informationen für ihre Aufgabenbearbeitung nutzen.

■ 2.2 Arten von Daten

Aufgrund der prominenten Stellung des Begriffes Data (also Daten) in Data Governance lohnt sich ein näherer Blick auf die Daten. Nach dem Verwendungszweck (zustands- oder abwicklungsorientiert) und der Veränderbarkeit (geringe oder hohe Änderungshäufigkeit) können vier Arten von Daten unterschieden werden ([LeOt07, Sche08, S.19 f], Bild 2.2): Stammdaten, Bestandsdaten, Bewegungsdaten und Änderungsdaten.

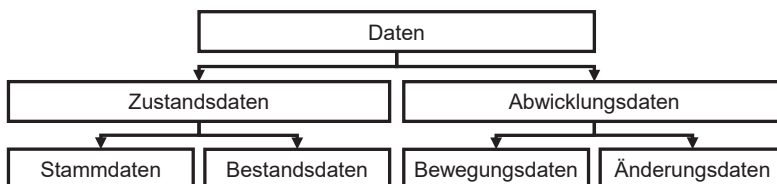


Bild 2.2 Datenarten [vgl. Sche08, S.20]

Stammdaten repräsentieren die Kernentitäten bzw. Kernobjekte eines Unternehmens, z. B. Kunden, Materialien, Produkte, Personal und Lieferanten. Stammdaten sind zustandsorientierte Daten, die sich selten ändern. Stammdaten werden von Bewegungsdaten zur Beschreibung von Geschäftsvorfällen referenziert.



Beispiel: Verwendung von Stammdaten in Geschäftsvorfällen

Die Bestellung (= Geschäftsvorfall) einer Kundin enthält typischerweise Daten aus dem Stammdatensatz der Kundin, wie z. B. ihre Adresse und die vereinbarten Zahlungsbedingungen. Die Bestellung enthält aber auch Daten aus einem oder mehreren Produktstammdatensätzen, nämlich die Nummern, Bezeichnungen, den Preis und eventuell weitere Merkmale der bestellten Produkte.

Bestandsdaten beschreiben die betriebliche Mengen- und Wertstruktur. Typische Bestandsdaten sind Lagerbestand und Kontostand. Bestandsdaten sind wie Stammdaten zustandsorientiert, sie ändern sich aber häufig durch geschäftsbedingte Zu- und Abgänge.

Bewegungsdaten bilden betriebswirtschaftliche Vorgänge ab. Typische Beispiele sind Fertigungsaufträge, Lieferscheine, Bestellungen und Rechnungen. Bewegungsdaten sind abwicklungsorientiert und verändern Bestandsdaten durch mengen- oder wertmäßige Zu- und Abgänge. Zum Beispiel erhöht ein Wareneingang den Bestand an Rohstoffen und eine bezahlte Lieferantenrechnung verringert den Kontostand bei der Bank.

Änderungsdaten sind abwicklungsorientierte Daten. Sie lösen Änderungen von Stammdaten aus. Ein Beispiel ist die Hochzeit einer Mitarbeiterin, welche den Familienstand im Mitarbeiterstammsatz ändert. Die Entwicklung eines besseren Wirkstoffs ändert die Zusammensetzung der Produktionsstückliste eines Düngemittels. Eine Änderung der Rechtsform eines Unternehmens wirkt sich auf die Bezeichnung des Unternehmens in dessen Lieferantenstammsatz aus.

Die Unterscheidung der Datenarten ist nicht immer eindeutig und hängt auch vom Unternehmenszweck ab. Beispielsweise ist ein Umzug für die meisten Unternehmen ein Änderungsdatum. Für ein Postunternehmen hat die Meldung über die erfolgte oder bevorstehende Adressänderung der Kunden eher den Charakter eines Bewegungsdatums. Es ist ein typischer Geschäftsvorfall, der vielfältige Aktivitäten auslöst, z. B. die Einrichtung eines Nachsendeauftrags.

Im Zusammenhang mit betrieblichen Daten können auch die Begriffe Geschäftsobjekt und Datenobjekt (bzw. Datenelement) unterschieden werden [vgl. ÖsHO11]. **Geschäftsobjekte** sind reale oder gedachte Gegenstände, welche in Geschäftsprozessen verwendet oder bearbeitet werden. Dabei handelt es sich meistens um Stammdaten (z. B. Material, Kunde, Anlage) oder Bewegungsdaten (z. B. Rechnung, Auftrag, Vertrag). **Datenobjekte** sind die informationstechnische Repräsentation dieser Geschäftsobjekte in Anwendungssystemen. Das Geschäftsobjekt Material wird demnach als Materialstammdaten in einem ERP-System abgebildet. Geschäftsobjekte sind somit eine fachliche Sicht auf Datenobjekte. Sie abstrahieren von einer konkreten Repräsentation in einem Anwendungssystem.

Die Notwendigkeit dieser Unterscheidung zeigt sich speziell in der Kommunikation mit verschiedenen Adressaten von Data Governance. Fachexperten sprechen über die Verwen-

derung von Daten (Informationen) in Geschäftsprozessen, also den Geschäftsobjekten. IT-Experten interessieren sich für die Repräsentation von Daten in Anwendungssystemen und deren Speicherung in Datenbanken und meinen daher Datenobjekte.

■ 2.3 Stammdaten

Stammdaten repräsentieren häufig die wichtigsten Geschäftsobjekte eines Unternehmens, z.B. Kunden, Artikel, Lieferanten oder Mitarbeiter. Sie sind für die Abwicklung von Geschäftsprozessen von zentraler Bedeutung. Stammdaten werden in Geschäftsvorfällen (Bewegungsdaten) referenziert. Sie haben somit eine steuernde Wirkung von IT-gestützt ablaufenden Geschäftsprozessen. Beispielsweise steuern die Maße und das Gewicht der Produkte die Beladung eines LKWs. Anhand der aufsummierten Maße und Gewichte der bestellten Produkte wird über das Versandsystem automatisch ermittelt, welche und wie viele LKWs für die Lieferung eingesetzt werden müssen.



Beispiel: Materialstammsatz (siehe Bild 2.3)

„Ein Materialstammsatz speichert beispielsweise sämtliche Informationen zu den Artikeln, Teilen und Dienstleistungen, die ein Unternehmen beschafft, fertigt und lagert. Er wird von den verschiedenen funktionalen Bereichen (z. B. Einkauf, Verkauf, Logistik, Produktion oder Buchhaltung) genutzt. Diese benötigen neben den allgemeinen Grunddaten, wie z. B. einer Materialnummer zur eindeutigen Identifikation und einer Bezeichnung, in der Regel noch funktionspezifische Informationen, die in Sichten gruppiert werden: So müssen dem Einkauf insbesondere Preise und Konditionen zur Verfügung stehen, während für die Fertigungsplanung (Disposition) Sicherheitsbestände und Losgrößen relevant sind und die Buchhaltung vor allem an Kontierungsinformationen interessiert ist. Stammdaten können von der Organisationsstruktur abhängig sein bzw. in mehreren Sprachen gepflegt werden (...). Durch die Integration zum Beispiel aller materialspezifischen Informationen in einem einzigen Stammsatz entfällt dafür die redundante Datenhaltung.“ [LeOt07]

Stammdaten passieren während ihres Lebenszyklus oft verschiedene Unternehmensfunktionen. Ein Beispiel sind Produktstammdaten, die in der Forschung & Entwicklung entstehen und später in der Produktion, im Marketing und in der Auftragsabwicklung genutzt werden. Die Stammdaten Name, Adresse und Ansprechpartner einer neuen Kundin zum Beispiel haben ihren Ursprung häufig bei Mitarbeitern des Außendienstes, das Controlling ordnet der Kundin verschiedenen Kategorien für das Berichtswesen zu, der Vertrieb legt Zahlungs- und Lieferkonditionen fest, und die IT ist für die Bereitstellung und Wartung der Systeme zuständig, welche die Kundenstammdaten vorhalten (z. B. ERP- und CRM-Systeme). Otto und Österle [OtÖs16, S.29 f] unterscheiden „globale“ und „lokale“ Stammdaten. Globale Stammdaten – oder auch **Konzerndaten** – sind Stammdaten, die für das gesamte

Unternehmen gültig sind. Lokale Stammdaten werden nur in einem Teil des Unternehmens verwendet, z. B. in einer Filiale, in einem Geschäftsbereich, in einer Abteilung oder in einem Werk. Im Fokus von Data Governance stehen meist global gültige Stammdaten.

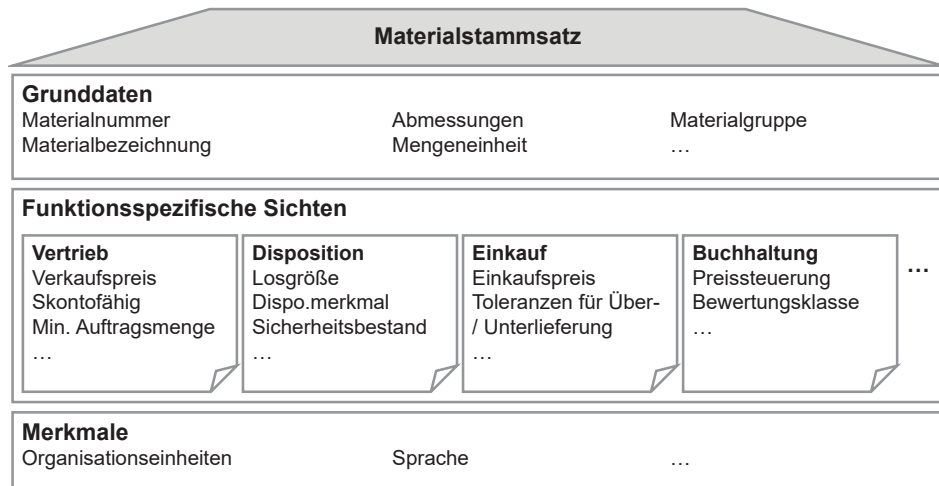


Bild 2.3 Beispiel für den Aufbau des Materialstammsatzes am Beispiel SAP R/3 (in Anlehnung an [LeOt07])

Eine spezielle Art von Stammdaten sind **Referenzdaten**. Referenzdaten haben ähnliche Eigenschaften wie Stammdaten, ihr Ursprung und auch die Verantwortung für deren Qualität liegen jedoch außerhalb des eigenen Unternehmens [OtÖs16, S. 30]. Typische Beispiele sind Währungs- und Ländercodes, Geodaten und Klassifikationsstandards wie eCl@ss. Auch viele Adressdaten sind Referenzdaten, wie z. B. Straßennamen, Hausnummern, Ortsbezeichnungen und Postleitzahlen. Für die Vergabe (und Pflege) der deutschen Postleitzahlen ist beispielsweise die Deutsche Post AG zuständig.

Eine andere Sichtweise ist, dass auch unternehmensinterne Daten als Referenzdaten bezeichnet werden können, wenn sie einen „referenzierenden“ Charakter haben. Sie ändern sich noch seltener als Stammdaten, sind deutlich weniger komplex und werden nicht nur in Bewegungsdaten, sondern auch in den Stammdaten selbst referenziert. Ein Beispiel sind Organisationsbezeichnungen, wie sie in ERP-Systemen verwendet werden, also z. B. Buchungskreise, Vertriebsbezirke, Kunden- oder Produktgruppen und Kostenschlüssel. Diese Referenzdaten haben meist eine ordnende Funktion oder dienen der Charakterisierung anderer Daten [HeED17, S. 350ff]. In einer Hochschule könnte man Studiengänge oder Semesterzahlen (1 bis 7) als Referenzdaten bezeichnen. In dem Fall werden die Studierenden (= Stammdatum) einem Studiengang und dem aktuellen Semester (= Referenzdaten) zugeordnet.