

Peter Gluchowski
Peter Chamoni *Hrsg.*

Analytische Informationssysteme

Business Intelligence-Technologien
und -Anwendungen

5. Auflage



Springer Gabler

Analytische Informationssysteme

Peter Gluchowski • Peter Chamoni
(Hrsg.)

Analytische Informationssysteme

Business Intelligence-Technologien
und -Anwendungen

5., vollständig überarbeitete Auflage

Herausgeber

Peter Gluchowski
Technische Universität Chemnitz
Chemnitz
Deutschland

Peter Chamoni
Universität Duisburg-Essen
Duisburg
Deutschland

ISBN 978-3-662-47762-5
DOI 10.1007/978-3-662-47763-2

ISBN 978-3-662-47763-2 (eBook)

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

Springer Gabler

© Springer-Verlag Berlin Heidelberg 1998, 2004, 2006, 2010, 2016

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt 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. Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Berlin Heidelberg ist Teil der Fachverlagsgruppe Springer Science+Business Media
(www.springer.com)

Vorwort zur 5. Auflage

Als Herausgeber des Sammelbandes „Analytische Informationssysteme“ freuen wir uns, Ihnen als Leser fünf Jahre nach Erscheinen der vierten Auflage eine aktualisierte und mit vielen neuen Themen und Autoren versehene neue Auflage vorlegen zu können.

Das behandelte Themengebiet erweist sich nach wie vor als wachstumsstark und innovativ. Stetig drängen neue Anbieter und Technologien auf den Markt und versprechen, mit einer andersartigen Herangehensweise bessere Lösungen für Anwenderunternehmen aufbauen zu können. Nicht zuletzt durch die große öffentliche Aufmerksamkeit, die das Begriffsgebilde Big Data (Analytics) auf sich zieht, werden auch das Data Warehousing und die Business Intelligence neu belebt. Allerdings ergeben sich ebenfalls erweiterte technologische und organisatorische Herausforderungen durch ein vergrößertes Aufgabengebiet und gestiegene Anforderungen.

Somit adressiert der überarbeitete Sammelband nicht nur die Leserschaft, die sich seit längerer Zeit mit Data Warehousing und Business Intelligence beschäftigt und die eigene Perspektive aktualisieren oder erweitern will, sondern auch neu hinzukommende Experten für Informationsbereitstellung und betriebswirtschaftliche Analyse.

Wir freuen uns darüber, dass es erneut gelungen ist, eine Vielzahl renommierter Autoren für einen Beitrag im Sammelband gewinnen zu können, und danken ihnen dafür, dass sie ihre spezifischen Expertisen eingebracht haben. Nur so ist es möglich, das Spektrum der unterschiedlichen Perspektiven auf den Themenbereich gebündelt zu präsentieren.

Dem Leser wünschen wir, dass er durch die Lektüre des Sammelbandes interessante und hilfreiche Informationen für seine Praxisprojekte oder zur eigenen wissenschaftlichen Arbeit erhält.

Dem Springer-Verlag danken wir für die langjährige gute Zusammenarbeit. Natürlich bietet ein derartiger Sammelband Anlass zur kritischen Reflexion, so dass wir uns wie immer über konstruktive Hinweise freuen.

Chemnitz, Duisburg, im April 2015

Peter Gluchowski und Peter Chamoni

Vorwort zur 4. Auflage

Als Herausgeber des Sammelbandes „Analytische Informationssysteme“ freuen wir uns, Ihnen als Leser, nun 12 Jahre nach der Erstauflage eine in wesentlichen Teilen überarbeitete und mit vielen neuen Themen und Autoren versehene neue Auflage vorlegen zu können.

Die eingesetzten Technologien zu Data Warehousing und Business Intelligence haben sich flächendeckend in der Praxis etabliert, bieten aber weiterhin noch vielfältige Forschungsgebiete, so dass es auch für die Wissenschaft noch spannend bleibt. Insbesondere der Beitrag zur Chronologie dieses Forschungsfeldes zeigt, dass wir uns auch nach 50 Jahren Geschichte der entscheidungsunterstützenden Systeme nach wie vor den Herausforderungen der dynamischen Unternehmensprozesse stellen müssen.

In den letzten Jahren ist der Markt der analytischen Informationssysteme zu einem der wichtigsten Wachstumsträger in der IT-Branche geworden. Aktuelle Studien weisen für die nächsten Jahre auch weiterhin ungebrochene Wachstumsraten auf, so dass wir hoffen, mit dem überarbeiteten Sammelband nicht nur die Leserschaft zu erreichen, die uns in Theorie und Praxis seit Jahren begleiten, sondern auch neu hinzukommende Experten, die sich dem komplexen Gebiet der Informationsbereitstellung und betriebswirtschaftlichen Analyse stellen wollen.

Wir bedanken uns bei allen Autoren, die ihre spezifische Expertise in Form von Fachbeiträgen zeitgerecht eingebracht haben. Insbesondere freuen wir uns über die Bereicherung durch neue Autoren, denn damit wird das Spektrum der unterschiedlichen Sichtweisen und Schulen erkennbar. Dem Leser gibt es die Möglichkeit, aktuelle und umfassende Kenntnisse im Bereich der analytischen Informationssysteme aufzunehmen und hoffentlich gewinnbringend in Praxisprojekten oder zur eigenen wissenschaftlichen Arbeit einzusetzen.

Dem Springer-Verlag danken wir für die langjährige gute Zusammenarbeit und speziell Herrn Dr. Müller für seine Geduld. Natürlich fordert ein Sammelband auch zu Kritik auf, so dass wir uns wie immer über Rückmeldungen und Kommentare freuen.

Duisburg, Chemnitz, im Oktober 2009

Peter Chamoni und Peter Gluchowski

Vorwort zur 3. Auflage

Nach dem Erscheinen der 1. Auflage (1997) und der 2. Auflage (1999) des Sammelbandes „Analytische Informationssysteme“ hat sich in den letzten Jahren ein enormes Wachstum der unter diesem Begriff zusammengefassten Informationsverarbeitung ergeben. Ebenso erfolgte ein Begriffswandel, so dass der dem Zeitgeist entsprechende Titel des vorliegenden Bandes nun „Business Intelligence“ heißen müsste. Um einerseits die Kontinuität zu wahren und andererseits den Aspekt der betriebswirtschaftlichen Analyse ökonomischer Zusammenhänge in den Vordergrund zu stellen, halten wir dennoch im Haupttitel am Begriffsgebilde „Analytische Informationssysteme“ fest. Nach 6 Jahren war naturgemäß eine vollständige Überarbeitung der Beiträge notwendig, was sich in der Neugestaltung der nun vorliegenden 3. Auflage widerspiegelt.

Das Thema hat an Breite und Tiefe zugelegt, was die Autoren in ihrer weiteren Arbeit, sei es in der Forschung, in der Beratung oder in der Anwendung bestärkt. Einen vollständigen Überblick geben zu wollen, ist deutlich schwieriger geworden als vor knapp 10 Jahren, da die Entwicklung analytischer Applikationen sich noch in einer Frühphase befand. Technologie, Anwendungsfelder und Professionalisierung des Betriebs haben nun einen Reifegrad erreicht, der diese Klasse von IV-Systemen aus dem Nischendasein in das Bewusstsein allgegenwärtiger IT-Strategien der Unternehmen rückt. Die Herausgeber freuen sich über den regen Zuspruch der Leser, aber auch über die Bereitschaft alter und neuer Autoren, dem Sammelband nun eine längst überfällige Aktualität zu verleihen. Wie immer war es für alle Beteiligte ein Kraftakt, die Manuskripte zeitgerecht zu er stellen. Hierfür bedanken wir uns bei allen Autoren und speziell bei Herrn Dr. Müller, der in Geduld und hoher Verlässlichkeit die professionelle Abwicklung beim Springer-Verlag sichergestellt hat.

Allen Lesern wünschen wir, dass bei der Lektüre der 3. Auflage neue Aspekte und Erkenntnisse die Arbeit beim Aufbau von analytischen Informationssystemen bereichern. Wie immer freuen wir uns über konstruktive Kritik und Rückmeldungen.

Duisburg, Düsseldorf, im Oktober 2005

Peter Chamoni und Peter Gluchowski

Vorwort zur 2. Auflage

Die Themen, die sich mit dem Aufbau analytischer Informationssysteme befassen, sind in den letzten Jahren stark angewachsen, so dass eine kurzfristige Erweiterung und Neufassung des vorliegenden Bandes notwendig wurde. Kaum ein Gebiet der betriebswirtschaftlich orientierten Informatik unterliegt derzeit einem schnelleren Wandel als der Einsatz von Data Warehouses in der Praxis. Aber auch der Fokus der wissenschaftlichen Arbeit hat sich gemessen an der Anzahl von Veröffentlichungen und Konferenzen intensiv dem Bereich der Aufbereitung und Analyse entscheidungsrelevanter betriebswirtschaftlicher Daten zugewandt. Multidimensionale Datenbanken und Data Mining sind Forschungsgegenstände der Informatik geworden, so dass heutige Anwendungen auf eine wesentlich breitere Basis gestellt werden können. Auch liegen jetzt mehr Erfahrungen über Data Warehouse-Projekte vor, die zum einen die Vielfältigkeit der Einsatzgebiete zeigen, zum anderen aber auch auf die „pitfalls“ hinweisen, die einer erfolgreichen Implementation entgegenwirken.

Die vorliegende 2. Auflage der „Analytischen Informationssysteme“ versucht, die neuen Entwicklungstendenzen aufzunehmen, aktuelle Produktübersichten zu geben und aus der Anwendersicht kritisch Position zu beziehen. Entgegen der 1. Auflage sind nicht nur die marktbezogenen Informationen aktualisiert worden, sondern die meisten Beiträge sind überarbeitet worden; auch einige neue Aufsätze tragen der Tatsache Rechnung, dass das Gebiet des Data Warehousing in seiner Dynamik ständige Erweiterungen erfährt.

Die Herausgeber danken wiederum allen Autoren, die sich der Mühen einer Überarbeitung unterzogen haben oder neue Beiträge zeitgerecht einbrachten. Die Korrektur des Manuskripts lag wie bei der 1. Auflage in den bewährten Händen von Herrn Dipl.-Kfm. Steffen Stock, dem an dieser Stelle gedankt sei. Die reibungslose Abwicklung und angenehme Zusammenarbeit mit Herrn Dr. Müller und seinen Mitarbeitern ermöglichte die schnelle Realisierung der 2. Auflage.

Wir hoffen mit dem nunmehr vorliegenden Sammelband allen Praktikern und Wissenschaftlern erneut Impulse für ihre fachliche Arbeit geben zu können und freuen uns über Ihre Rückmeldungen.

Duisburg, Düsseldorf, im April 1999

Peter Chamoni und Peter Gluchowski

Vorwort zur 1. Auflage

„Nichts ist beständiger als der Wandel.“ Diese von zahlreichen Zukunftsforschern und Wirtschaftsaguren proklamierte Feststellung wird – so die einhellige Meinung – die Geschicke vieler Wirtschaftsunternehmen in den 90er Jahren und vielleicht auch darüber hinaus bestimmen. Nur die Marktteilnehmer, denen es gelingt, sich rasch auf wechselnde Umweltbedingungen einzustellen, haben mittel- und langfristig Chancen, sich im harten Verdrängungswettbewerb zu behaupten. Dem effizienten Einsatz der Informationstechnologie wird hierbei in vielen Bereichen eine besonders hohe strategische Bedeutung beigemessen.

Dabei ist heute festzustellen, dass die operativen betrieblichen Abläufe fast flächendeckend durch Transaktionssysteme unterstützbar sind und eine zumindest partielle Integration in weiten Teilbereichen vollzogen ist. Moderne betriebswirtschaftliche Anwendungssoftwarepakete bieten funktionsbereichsübergreifende Module mit hoher Leistungsbreite. Sehr heterogen bezüglich ihrer Mächtigkeit und der gewählten Realisierungsform präsentiert sich dagegen die Systemlandschaft im Bereich der Informationsversorgung und Entscheidungsunterstützung betrieblicher Fach- und Führungskräfte. Allerdings werden derzeit von Anbietern und Anwendern massive Anstrengungen unternommen, um vorhandene Defizite abzubauen. So verwundert es nicht, dass Schlagworte wie „Data Warehouse“, „On-Line Analytical Processing“ und „Data Mining“, die in diesem Umfeld einzuordnen sind und deren logische Zusammenfassung hier unter dem Oberbegriff „Analytische Informationssysteme“ erfolgt, intensiv und teilweise kontrovers diskutiert werden.

Vor diesem Hintergrund sollen im vorliegenden Sammelband die aktuellen Konzepte und Techniken aufgegriffen und aus unterschiedlichen Perspektiven diskutiert werden. Das Spektrum reicht hier von der internen, eher informatikorientierten Funktionsweise über Gestaltungsaspekte bis zur Endbenutzerschicht der einzelnen Systeme und Ansätze. Neben diesen technischen Gesichtspunkten sind es auch die betriebswirtschaftlich-organisatorischen Aspekte, die keinesfalls vernachlässigt werden dürfen. Um die vielschichtige Thematik einer gewissen Struktur zuzuführen, wurde eine Untergliederung des Sammelbandes in fünf Hauptabschnitte vorgenommen.

Der erste Teil stellt eine Einführung in das Themengebiet dar und nimmt eine Abgrenzung und Einordnung der wesentlichen Begriffe vor. Zudem finden sich hier zwei empi-

risch fundierte Übersichten über den Markt für „Analytische Informationssysteme“, wie er sich derzeit international und im deutschsprachigen Raum präsentiert.

Der zweite Teil des Sammelbandes hat das Thema „Data Warehouse“ zum Gegenstand. Der Vielschichtigkeit dieses Konzeptes wird durch eine Reihe unterschiedlicher Fachbeiträge genüge geleistet, die sich mit den einzelnen Komponenten einer ganzheitlichen Data Warehouse-Architektur auseinandersetzen.

Speziell die multidimensionale Sichtweise des Anwenders auf verfügbare Datenbestände steht beim „On-Line Analytical Processing“ im Vordergrund. Die Beiträge im Teil 3 beschäftigen sich mit der angemessenen Umsetzung dieser Philosophie auf der internen, konzeptionellen und externen Systemebene.

Wegen des hohen Anspruchs, der mit einer aktiven Informationsversorgung betrieblicher Fach- und Führungskräfte verbunden ist, und der viel versprechenden Auswirkungen in wirtschaftlicher Hinsicht finden die Ansätze, die mit einem Knowledge Discovery in Databases verknüpft sind, in Wissenschaft und Praxis große Beachtung. Teil 4 setzt sich mit diesem Thema unter der Bezeichnung „Data Mining“ auseinander und erläutert vor allem die zur Verfügung stehenden Algorithmen und Visualisierungstechniken.

Eine Betrachtung der „Analytischen Informationssysteme“ aus rein technischer Sicht, wie sie in den Teilen 1 bis 4 im Vordergrund steht, gibt ein unzureichendes Bild der gesamten Thematik wieder. Vielmehr sind es in Praxisprojekten insbesondere auch die betriebswirtschaftlich-organisatorischen Gesichtspunkte, die es zu beachten gilt und die letztlich über Erfolg oder Misserfolg entscheiden. Aus diesem Grund wurde im vorliegenden Sammelband ein separater, fünfter Teil den hier auftretenden Fragestellungen gewidmet.

Als Herausgeber des Sammelbandes hoffen wir, der komplexen Thematik mit den zahlreichen Beiträgen aus Praxis und Wissenschaft gerecht geworden zu sein. Da sich die meisten Autoren mit ihren Beiträgen in ihren aktuellen Arbeits- bzw. Forschungsschwerpunkten bewegen, ist neben der gebotenen Vielfalt auch eine gewisse inhaltliche Qualität garantiert.

Als primäre Zielgruppe für das Buch kommen aus unserer Sicht betriebliche Anwender und Entscheider aus den DV-Abteilungen aber auch aus den Fachbereichen in Betracht. Insbesondere gehen wir davon aus, dass Systemgestaltern bei der Konzeptionierung und Realisierung von Data Warehouse-, OLAP- und Data Mining-Systemen wertvolle Hinweise und hilfreiche Anregungen gegeben werden können. Schließlich sind darüber hinaus vor allem auch Leser aus dem wissenschaftlichen Umfeld anvisiert, die einen aktuellen Überblick über den Stand der Technik sowie mögliche Entwicklungstendenzen gewinnen wollen.

Der Ansatz dieses Sammelbandes besteht darin, eine Vorstrukturierung mit abgegrenzten Teilgebieten vorzunehmen und durch ausgewiesene Fachexperten für den jeweiligen Aspekt mit Inhalt füllen zu lassen. Leider weist das Buch immer noch Teilaspekte auf, die nicht oder zu kurz behandelt werden, weil sich hier keine Fachexperten motivieren lassen konnten oder diese es nicht geschafft haben, ihre Beiträge fristgerecht fertig zu stellen. Auf's herzlichste wollen wir uns dagegen bei den einzelnen Autoren bedanken, die trotz voller Terminkalender und vielfältiger anderer Verpflichtungen sowie des engen

Zeitrahmens ihre Aufsätze einbringen konnten. Bedanken wollen wir uns auch für die redaktionelle Überarbeitung bei Frau Elke Ochs, Sekretärin im Fachgebiet Statistik und Ökonometrie in Düsseldorf, sowie Herrn Dipl.-Kfm. Steffen Stock und Herrn René Angenheister. Als besonders angenehm empfanden wir die Zusammenarbeit mit dem Springer-Verlag und besonders mit Herrn Dr. Müller und seinen Mitarbeitern.

Abschließend wünschen wir den Leserinnen und Lesern des Sammelbandes eine angenehme und ergiebige Lektüre. Für kritische oder bestätigende Anmerkungen stehen wir unter den E-Mail-Adressen *Chamoni@uni-duisburg.de* sowie *Gluchows@uni-duesseldorf.de* zur Verfügung.

Duisburg, Düsseldorf, im Juni 1997

Peter Chamoni und Peter Gluchowski

Inhaltsverzeichnis

Teil I Grundlagen und Organisatorische Aspekte

1 Analytische Informationssysteme – Einordnung und Überblick	3
Peter Chamoni und Peter Gluchowski	
1.1 Motivation	4
1.2 Einordnung der Analytischen Informationssysteme in die betriebliche Informationsverarbeitung	6
1.3 Grundlagen und organisatorische Aspekte	8
1.4 Architektur und Technologien	9
1.5 Betriebswirtschaftliche Anwendung und spezielle Aspekte der Analytischen Informationssysteme	10
Literatur	12
2 Historische Fragmente einer Integrationsdisziplin – Beitrag zur Konstruktgeschichte der Business Intelligence	13
Christian Schieder	
2.1 Einleitung	14
2.2 Historische Fragmente zur Chronologie der Business Intelligence	16
2.3 Etymologische Annäherung an Business Intelligence – Taxonomie des State-of-the-Art	21
2.4 Situation Awareness als Zielkonstrukt einer kognitionsorientierten Gestaltung der Business Intelligence	25
2.5 Zusammenfassung und Ausblick	27
Literatur	30
3 Von der Business-Intelligence-Strategie zum Business Intelligence Competency Center	33
Andreas Totok	
3.1 Einführung	34
3.2 Entwicklung einer BI-Strategie	36

3.2.1	Business-Intelligence-Strategie	36
3.2.2	Vorgehensmodell zur Entwicklung einer BI-Strategie	36
3.3	Aufbau eines BI Competency Center	45
3.4	Ausgewählte Aspekte einer BI-Strategie	46
3.4.1	Kosten und Nutzen bewerten	46
3.4.2	Auf Trends reagieren	48
3.4.3	Architektur gestalten	48
3.4.4	Softwareportfolio festlegen	49
3.4.5	Spannungsfeld erkennen	50
3.5	Fazit	52
	Literatur	53
4	Die nächste Evolutionsstufe von AIS: Big Data	55
	Carsten Dittmar	
4.1	Big Data und Business Intelligence	56
4.2	Erweiterung der analytischen Anwendungsfelder durch Big Data	58
4.3	Überblick über Technologien für Big Data	61
4.4	Erweiterung klassischer BI-Architekturen durch Big Data	62
4.5	Fazit	64
	Literatur	65
5	Analytische Informationssysteme aus Managementsicht:	
	lokale Entscheidungsunterstützung vs. unternehmensweite	
	Informations-Infrastruktur	67
	Robert Winter	
5.1	Analytische Informationssysteme aus Managementsicht	68
5.1.1	Dimensionen „Umsetzungsunabhängigkeit“ und „Reichweite“ ...	68
5.1.2	Infrastruktur vs. Analytik	71
5.1.3	Unternehmensweite Informations-Infrastruktur	72
5.1.4	Analytik	74
5.1.5	Gestaltungsaufgaben für Informations- Infrastruktur und Analytik	76
5.2	Unternehmensweite Informations-Infrastruktur aus Managementsicht ...	78
5.2.1	Projekt- vs. Betriebssicht	78
5.2.2	Informationsinfrastruktur-Strategie	79
5.2.3	Informationsinfrastruktur-Organisation	82
5.2.4	Finanzielle Aspekte der Informationsinfrastruktur	84
5.2.5	IT/Business Alignment für die Informations-Infrastruktur	86
5.3	Analytik aus Managementsicht	87
5.3.1	Projekt- vs. Betriebssicht	88
5.3.2	Anwendungspotenziale und Wertbeitrag von Analytik	88

5.3.3	Realisierungsformen von Analytik	90
5.3.4	IT/Business Alignment in der Analytik	91
5.4	Ausblick	92
	Literatur	93
6	Werkzeuge für analytische Informationssysteme	97
	Carsten Bange	
6.1	Einsatzbereiche und Werkzeugkategorien	98
6.2	Marktentwicklung	100
6.2.1	Trends im Markt für Datenmanagement	104
6.2.2	Trends im Markt für BI-Anwenderwerkzeuge	105
6.3	Datenintegration und -aufbereitung	105
6.4	Datenspeicherung und -bereitstellung	108
6.5	Anwendungen	111
6.5.1	Cockpits und Scorecards	112
6.5.2	Standard Reporting	114
6.5.3	Ad-hoc Reporting	117
6.5.4	Analyse	118
6.5.5	Planung und Simulation	123
6.5.6	Legale Konsolidierung	124
6.5.7	Data Mining	124
6.6	Fazit	125
	Literatur	126
Teil II Architektur und Technologien		
7	Transformation operativer Daten	129
	Hans-Georg Kemper und Ralf Finger	
7.1	Operative und dispositive Daten	130
7.2	Transformation – Ein Schichtenmodell	131
7.2.1	Filterung – Die Extraktion und Bereinigung operativer Daten	133
7.2.2	Harmonisierung – Die betriebswirtschaftliche Abstimmung gefilterter Daten	137
7.2.3	Aggregation – Die Verdichtung gefilterter und harmonisierter Daten	139
7.2.4	Anreicherung – Die Bildung und Speicherung betriebswirtschaftlicher Kenngrößen	141
7.3	Aktuelle Konzepte – Datentransformation wird nicht obsolet	142
7.4	Fazit	143
	Literatur	144

8 Architekturkonzepte und Modellierungsverfahren für BI-Systeme	147
Michael Hahne	
8.1 Einleitung	148
8.2 Traditionelle Business-Intelligence-Architekturen	149
8.2.1 Stove-Pipe-Ansatz	149
8.2.2 Data Marts mit abgestimmten Datenmodellen	151
8.2.3 Core Data Warehouse	151
8.2.4 Hub-and-Spoke-Architektur	153
8.2.5 Data-Mart-Busarchitektur nach Kimball	155
8.2.6 Corporate Information Factory nach Inmon	156
8.2.7 Architekturvergleich Kimball und Inmon	158
8.3 Core-Data-Warehouse-Modellierung in Schichtenmodellen	158
8.3.1 Aufgaben und Komponenten in Multi-Layer-Architekturen	159
8.3.2 Eignungskriterien für Methoden der Core-Data-Warehouse-Modellierung	162
8.4 Star-Schema-Modellierung im Core Data Warehouse	164
8.4.1 Granulare Star-Schemata im Core Data Warehouse	164
8.4.2 Bewertung dimensionaler Core-Data-Warehouse-Modelle	167
8.5 Normalisierte Core-Data-Warehouse-Modelle	168
8.5.1 Core-Data-Warehouse-Modellierung in 3NF	168
8.5.2 Historisierungsaspekte von 3NF-Modellen	168
8.5.3 Bewertung 3NF-Modellierung im Core Data Warehouse	170
8.6 Core Data Warehouse mit Data-Vault	171
8.6.1 Hub-Tabellen	172
8.6.2 Satellite-Tabellen	173
8.6.3 Link-Tabellen	176
8.6.4 Zeitstempel im Data Vault	179
8.6.5 Harmonisierung von fachlichen Schlüsseln	180
8.6.6 Agilität in Data-Vault-Modellen	181
8.6.7 Bewertung der Data-Vault-Methode	182
8.7 Zusammenfassung	183
Literatur	184
9 Grundlagen und Einsatzpotentiale von In-Memory-Datenbanken	187
Tobias Knabke und Sebastian Olbrich	
9.1 Einleitung und Motivation	188
9.2 Grundlagen In-Memory-Datenbanken	189
9.2.1 Aktuelle Entwicklungen im Hauptspeicherbereich	189
9.2.2 Datenorganisation	191
9.3 Konsequenzen für Business Intelligence und Business Analytics	193
9.3.1 Klassische Business Intelligence-Architekturen	193

9.3.2	Zukünftige Business Intelligence-Architektur als semi-virtuelles Data-Warehouse	195
9.3.3	Konsequenzen für die Informationslandschaft	197
9.4	Aktuelle Beispiele aus der betrieblichen Praxis	198
9.4.1	Steigende Datenmenge und zunehmende Integration unstrukturierter Daten erhöhen die Komplexität von Business Intelligence	198
9.4.2	Potentiale von In-Memory-Technologie am Beispiel von Handelsunternehmen	199
9.4.3	Unterstützung von Geschäftsprozessen durch In-Memory-Technologie im Bereich Predictive Maintenance	200
9.5	Fazit und Ausblick	201
	Literatur	201
10	NoSQL, NewSQL, Map-Reduce und Hadoop	205
	Jens Lechtenböcker und Gottfried Vossen	
10.1	Einleitung und motivierende Beispiele	206
10.2	NoSQL- und NewSQL-Datenbanken	207
10.2.1	Grundlagen: Partitionierung, Replikation, CAP-Theorem, Eventual Consistency	208
10.2.2	NoSQL	211
10.2.3	NewSQL	212
10.3	Big Data und Map-Reduce/Hadoop	213
10.3.1	Technologie zur Handhabung von Big Data: Map-Reduce	214
10.3.2	Nutzung von Big Data	219
10.4	Zusammenfassung und Ausblick	221
	Literatur	222
11	Entwicklungstendenzen bei Analytischen Informationssystemen	225
	Peter Gluchowski	
11.1	Motivation	226
11.2	Technologische und organisatorische Entwicklungen bei Analytischen Informationssystemen	226
11.2.1	Advanced/Predictive Analytics	227
11.2.2	Datenmanagement von BI-Systemen	228
11.2.3	Agile BI	232
11.2.4	Self Service BI	233
11.2.5	Mobile BI	236
11.3	Zusammenfassung	237
	Literatur	237

Teil III Betriebswirtschaftliche Anwendung und spezielle Aspekte

12 Planung und Informationstechnologie – Vielfältige Optionen	
erschweren die Auswahl	241
Karsten Oehler	
12.1 Einleitung	242
12.2 Werkzeug-Kategorien	242
12.2.1 ERP (Enterprise Resource Planning)	243
12.2.2 Tabellenkalkulation	243
12.2.3 Generische OLAP Werkzeuge (Generisches BI)	244
12.2.4 Dedizierte Planungsanwendungen (BI Anwendungen)	244
12.3 Die Beurteilung im Einzelnen	245
12.3.1 Zentrale Administration	245
12.3.2 Workflow-Unterstützung	246
12.3.3 Planung auf verdichteten Strukturen	247
12.3.4 Validierung	247
12.3.5 Automatisierte Top-Down-Verteilungen	248
12.3.6 Simulationsrechnungen	248
12.3.7 Vorbelegung	249
12.3.8 Aggregation	250
12.3.9 Kommentierungen	250
12.3.10 Eingabe über das Internet	251
12.3.11 Detailplanungen	251
12.3.12 Abstimmung des Leistungsnetzes	252
12.3.13 Übernahme von Stamm- und Bewegungsdaten aus ERP-Systemen	252
12.3.14 Verknüpfung von Teilplanungen	253
12.3.15 Abbildung abhängiger Planungsgrößen	254
12.3.16 Nachvollziehbarkeit	255
12.4 Zusammenfassende Bewertung	255
12.5 Ausblick	258
Literatur	258
13 Operational Business Intelligence im Zukunftsszenario der Industrie 4.0	259
Tom Hänel und Carsten Felden	
13.1 Einleitung	260
13.2 Entwicklung der IT in der Produktion bis zur Industrie 4.0	262
13.2.1 Unterstützung technischer Aufgaben und Funktionen	262
13.2.2 Integration von Aufgaben und Funktionen	263
13.2.3 Integration von Prozessen und Wertschöpfungsketten	265

13.2.4	Digitalisierung und globale Vernetzung	268
13.2.5	Status Quo der Operational Business Intelligence	269
13.3	Industrie 4.0	270
13.3.1	Technologische Grundlagen	271
13.3.2	Das Zukunftsszenario „Smart Factory“	275
13.3.3	Operational Business Intelligence in der Industrie 4.0	278
13.4	Fazit	279
	Literatur	280
14	Innovative Business-Intelligence-Anwendungen in Logistik und Produktion	283
	Henning Baars und Heiner Lasi	
14.1	Entscheidungsunterstützung in Logistik und Produktion – auf dem Weg zur Industrie 4.0	284
14.2	Relevante Entwicklungen in der Business Intelligence	285
14.3	Logistik und Produktion: Rahmenbedingungen, Herausforderungen und Entwicklungen	287
14.3.1	Logistik	287
14.3.2	Produktion	290
14.4	Innovative BI-Anwendungen in der Logistik	292
14.4.1	Produktionslogistik und Ladungsträgermanagement	293
14.4.2	Cloud- und RFID-basierte Retail Supply Chain	294
14.5	Innovative BI-Anwendungen in der Produktion	295
14.5.1	Produktentwicklung und Produktdatenmanagement	296
14.5.2	MES als Datenquelle für Produktions- und Qualitätsoptimierung	297
14.6	Diskussion und Ausblick	298
	Literatur	299
15	Der Markt für Visual Business Analytics	303
	Jörn Kohlhammer, Dirk U. Proff und Andreas Wiener	
15.1	Einführung	304
15.2	Das Visual-Business-Analytics-Modell	305
15.3	Information Design und Reporting	307
15.3.1	Dekoration	307
15.3.2	3-D-Diagramme	309
15.3.3	Skalierung	309
15.3.4	Einsatz von Farben, schlanke Visualisierung und hohe Informationsdichte	309
15.4	Visual Business Intelligence und Dashboarding	311
15.5	Visual Analytics und Big Data	315

15.5.1	Big Data	315
15.5.2	Visual Analytics	317
15.5.3	Aktueller Einsatz in Unternehmen	318
15.5.4	Anwendungsbeispiel	321
15.6	Zusammenfassung	322
	Literatur	323
16	Effektive Planung und Steuerung erfolgskritischer	
	Komponenten eines Geschäftsmodells	325
	Markus Linden und Frank Navrade	
16.1	Einführung	326
16.2	Evolutionsstufen in der Unternehmenssteuerung	327
16.2.1	Finanzorientierte Steuerungskonzepte	327
16.2.2	Ausgewogene Steuerungskonzepte	329
16.2.3	Gegenüberstellung aktueller Steuerungskonzepte	331
16.3	Erfolgskritische Komponenten eines Geschäftsmodells	333
16.3.1	Abdeckungsgrad der Geschäftsmodellkomponenten	333
16.3.2	Steuerungsdomänen eines Geschäftsmodells	336
16.3.3	Zuordnung steuerungsrelevanter Kennzahlen	338
16.4	Planung und Steuerung mit analytischen Informationen	340
16.4.1	Herausforderung in der Informationsbedarfsanalyse	340
16.4.2	Identifikation von Kennzahlen-Dimensionen-Kombinationen ...	341
16.4.3	Informationssystem zur Steuerung des Geschäftsmodells	345
16.5	Fazit und Ausblick	347
	Literatur	348
	Stichwortverzeichnis	351

Autorenverzeichnis

Dr. Henning Baars Universität Stuttgart, Lehrstuhl für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik I, Stuttgart

Dr. Carsten Bange Geschäftsführer, Business Application Research Center, Würzburg

Prof. Dr. Peter Chamoni Universität Duisburg-Essen, Lehrstuhl für Wirtschaftsinformatik und Operations Research, Duisburg

Dr. Carsten Dittmar Head of Business Development Big Data & Analytics, NTT DATA Deutschland GmbH, Köln

Prof. Dr. Carsten Felden Technische Universität Bergakademie Freiberg, Professur für ABWL, insbes. Informationswirtschaft/Wirtschaftsinformatik, Freiberg

Dr. Ralf Finger Geschäftsführer, Information Works Unternehmensberatung und Informationssysteme GmbH, Köln

Prof. Dr. Peter Gluchowski Technische Universität Chemnitz, Lehrstuhl für Wirtschaftsinformatik II, Chemnitz

Dipl. Wi.-Ing. Tom Hänel Technische Universität Bergakademie Freiberg, Professur für ABWL, insbes. Informationswirtschaft/Wirtschaftsinformatik, Freiberg

Dr. Michael Hahne Geschäftsführer, Hahne Consulting GmbH, Bretzenheim

Prof. Dr. Hans-Georg Kemper Universität Stuttgart, Lehrstuhl für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik I, Stuttgart

Dr. Heiner Lasi Universität Stuttgart, Lehrstuhl für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik I, Stuttgart

Dr. Jens Lechtenbörger Westfälische Wilhelms-Universität, Institut für Wirtschaftsinformatik, Münster

Dipl.-Kfm. Markus Linden Senior Consultant, SAP Deutschland SE & Co. KG, Ratingen

Prof. Dr.-Ing. Jörn Kohlhammer Abteilungsleiter, Fraunhofer-Institut für Graphische Datenverarbeitung, Darmstadt

M. Sc. Tobias Knabke Universität Duisburg-Essen, Lehrstuhl für Wirtschaftsinformatik, insbesondere Business Intelligence, Duisburg

Dr. Frank Navrade Head of Business Unit BI Strategy, PwC cundus AG, Duisburg

Prof. Dr. Karsten Oehler VP CPM Solutions, pmOne, Alzenau

Dr. Sebastian Olbrich Universität Duisburg-Essen, Lehrstuhl für Wirtschaftsinformatik, insbesondere Business Intelligence, Duisburg

Dipl.-Wirt.-Inf. Dirk U. Proff Geschäftsführer, blueforte GmbH, Hamburg

Dr. Christian Schieder Technische Universität Chemnitz, Lehrstuhl für Wirtschaftsinformatik II, Chemnitz

Dr. Andreas Totok Enterprise Information Management, Finanz Informatik Solutions Plus GmbH, Frankfurt am Main

Prof. Dr. Gottfried Vossen Westfälische Wilhelms-Universität, Institut für Wirtschaftsinformatik, Münster

M. A. Andreas Wiener Geschäftsführer, reportingimpulse GmbH, Hamburg

Prof. Dr. Robert Winter Universität St. Gallen, Institut für Wirtschaftsinformatik, St. Gallen

Teil I

Grundlagen und Organisatorische Aspekte

Peter Chamoni und Peter Gluchowski

Inhaltsverzeichnis

1.1 Motivation. 4

1.2 Einordnung der Analytischen Informationssysteme in die betriebliche
Informationsverarbeitung 6

1.3 Grundlagen und organisatorische Aspekte 8

1.4 Architektur und Technologien. 9

1.5 Betriebswirtschaftliche Anwendung und spezielle Aspekte der Analytischen
Informationssysteme. 10

Literatur. 12

Zusammenfassung

Analytische Informationssysteme – dieses Wortgebilde wird im vorliegenden Beitrag nicht etwa als bislang unbekanntes Konzept oder neuartiger Ansatz propagiert, sondern als logische Klammer, welche die gängigen Schlagworte wie „Data Warehouse“, „On-Line Analytical Processing“ und „Data Mining“ sowie „Big Data Analytics“ aber auch betriebswirtschaftlich-dispositive Anwendungslösungen umschließt.

P. Chamoni (✉)
Lehrstuhl für Wirtschaftsinformatik und Operations Research, Universität Duisburg-Essen,
Duisburg, Deutschland
E-Mail: peter.chamoni@uni-due.de

P. Gluchowski
Lehrstuhl für Wirtschaftsinformatik II, Technische Universität Chemnitz, Chemnitz, Deutschland
E-Mail: peter.gluchowski@wirtschaft.tu-chemnitz.de

Dabei erscheint es wichtig, dass sich eine Betrachtung des Problembereichs nicht nur auf informationstechnologische Aspekte beschränkt, sondern die in diesem Kontext immer wichtigeren betriebswirtschaftlich-organisatorischen und strategischen Fragestellungen gleichsam Beachtung finden.

1.1 Motivation

Das heutige Wirtschaftsleben ist für alle Marktteilnehmer durch hohe Komplexität und rasche Veränderungen interner und externer Rahmenbedingungen gekennzeichnet. Die Anforderungen, die daraus an den Produktionsfaktor Information erwachsen, sind in den letzten Jahren rapide gestiegen. Nur durch eine ausgereifte Informationslogistik kann sichergestellt werden, dass die benötigten Informationen zeitgerecht in der erforderlichen Qualität und am richtigen Ort vorliegen.

Zunächst waren es dabei schwerpunktmäßig die operativen Geschäftsabläufe, die es zu unterstützen galt und für die informationstechnologische Lösungen zu konzipieren und realisieren waren. Nachdem jedoch eine flächendeckende Versorgung mit betrieblicher Basisdatenverarbeitung (On-Line Transaction Processing (OLTP)) weitgehend gewährleistet und eine zumindest partielle Integration in vielen Teilbereichen vollzogen ist, verschiebt sich der Fokus zunehmend auf die entscheidungsgerechte Versorgung betrieblicher Fach- und Führungskräfte mit Informationen.

Unterstützungsbedarf wurde dabei zunächst von den oberen Führungsebenen sowie in den Controlling-Abteilungen artikuliert. Mittlerweile jedoch sind es die Mitarbeiter aus allen Unternehmensebenen und -bereichen, die vehement den Zugang zum Produktionsfaktor Information einklagen, um diesen als Wettbewerbsfaktor für Entscheidungen nutzbar zu machen.

Aufgabe der Informationstechnologie muss es sein, hier angemessene Zugänge zu eröffnen und dabei aus den umfangreichen Datenbeständen, mit denen sich Unternehmen heute konfrontiert sehen, personen-, problem- und situationsgerechte Angebote zusammenzustellen. Die Kunst besteht darin, dauerhafte Lösungen zu konzipieren, die sowohl den aktuellen Anforderungen der Anwender gerecht werden als auch informationstechnologisch mittel- und langfristig tragfähig sind.

Neue Trends und Entwicklungen, die in immer rascherer Folge aufgenommen und verarbeitet werden müssen, stellen die Wandlungs- und Anpassungsfähigkeit von Unternehmen und damit die Reaktionsfähigkeit von Fach- und Führungskräften heute auf eine harte Probe. Dazu tragen zunehmend Globalisierungstendenzen bei, die einhergehen mit globalen und lokalen Instabilitäten und verstärktem Wettbewerbsdruck, ebenso wie die rasante technologische Dynamik, die in allen Lebensbereichen einschneidende Verbesserungen aber auch immense Herausforderungen mit sich bringt.

Darüber hinaus finden in zahlreichen Unternehmen immer noch tiefgreifende Umstrukturierungen statt, die sich sowohl in Veränderungen der Aufbaustrukturen als auch

der Prozessabläufe dokumentieren. Eine Hinwendung zu einer kunden- und vorgangsorientierten Unternehmensorganisation, wie im Rahmen des Customer Relationship Management und des Business Process Reengineering gefordert, erweist sich nur mit geeigneter informationstechnologischer Unterstützung als durchführbar.

Der äußere Druck für die Konzeption und Realisierung ausreichender Unterstützungssysteme für das Management scheint somit gegeben. Nachdem jedoch in der Vergangenheit vielfältige Versuche in dieser Richtung gescheitert sind oder jedenfalls oftmals zu unbefriedigenden Lösungen geführt haben, muss nach neuen Wegen gesucht werden.

Moderne DV-Landschaften, wie sie in den letzten Jahren in vielen Unternehmen entstanden, sollen dabei hilfreich zur Seite stehen. Weder monolithische Großrechnerlösungen mit ausgeprägter Zentralisierung von Rechenkapazitäten noch die PC-orientierte Maximierung von Benutzerautonomie am Arbeitsplatz stellten letztlich geeignete Architekturparadigmen dar, mit denen eine flächendeckende Informationsversorgung in zufriedenstellender Art zu gewährleisten war. Erst durch die Verbindung unterschiedlicher Rechnersysteme im Rahmen verteilter und kooperativer Architekturen, gegebenenfalls in Verbindung mit den allgegenwärtigen Internet-Technologien, sind die technischen Voraussetzungen gegeben, um auf große Datenvolumina endbenutzergerecht und mit der erforderlichen Performance zugreifen zu können.

In den Unternehmungen sind vor allem die Konzepte und Technologien des Data Warehousing, des On-Line Analytical Processing, des Data Mining und des betrieblichen Berichtswesens sowie in den letzten Jahren auch Big Data Analytics aufgegriffen und umgesetzt worden, um tragfähige Lösungen für eine unternehmensweite Informationsversorgung mit Analysefunktionalitäten zu etablieren. Im Rahmen des vorliegenden Sammelbandes erfolgt die Behandlung entsprechender Systeme und Konzepte sowie der darauf aufsetzenden betriebswirtschaftlichen Anwendungslösungen unter dem Oberbegriff „**Analytische Informationssysteme**“ (AIS).

Dabei sind die unterschiedlichen technischen Realisationsalternativen aufzuzeigen und auf ihre Tauglichkeit im betrieblichen Einsatz hin zu untersuchen. Nicht zuletzt wird angestrebt, die wesentlichen Schlagworte und Akronyme, die in diesem Zusammenhang immer wieder gebraucht werden, zu entmystifizieren und in den Gesamtkomplex logisch einzuordnen.

Die Beschränkung auf informationstechnologische Aspekte des Themas Analytische Informationssysteme jedoch kann keine hinreichende Erörterung bieten, da es deren Zweck ist, den Endbenutzer besser bei seinen Aufgaben zu unterstützen und ihm zu helfen, sein Geschäft zu verstehen und zu beherrschen. Deshalb sollen auch die zugehörigen betriebswirtschaftlich-organisatorischen sowie strategischen Gesichtspunkte eingehend behandelt werden.

Vor diesem Hintergrund ist es Ziel des vorliegenden Beitrags, einerseits das zugrundeliegende Begriffsverständnis zu erläutern, andererseits jedoch auch die Einordnung der behandelten Facetten in den thematischen Kontext vorzunehmen. Zu diesem Zweck soll im folgenden Kapitel zunächst eine Einordnung und Abgrenzung Analytischer Informationssysteme erfolgen, bevor die einzelnen Beiträge des Sammelbandes in einen logischen

Gesamtzusammenhang gebracht werden. Eine aktuelle Bestandsaufnahme mit Hinweisen auf zukünftige Entwicklungspotenziale Analytischer Informationssysteme beschließt den Beitrag.

1.2 Einordnung der Analytischen Informationssysteme in die betriebliche Informationsverarbeitung

Bereits oben wurde auf den grundlegenden Wandel eingegangen, dem betriebliche Aufbau- und Ablauforganisationen seit einigen Jahren unterworfen sind. Für den einzelnen Mitarbeiter erwächst aus dem Leitbild einer schlanken Unternehmensorganisation in der Regel eine Ausweitung des individuellen Entscheidungsspielraumes. Längst ist es nicht mehr das klassische Management, dem Entscheidungen vorbehalten sind. Vielmehr werden in zunehmendem Umfang auch kompetente Fachkräfte in die Pflicht genommen, wenn es darum geht, anstehende Probleme schnell, unbürokratisch und ohne Inanspruchnahme des betrieblichen Instanzenweges zu lösen. Dies jedoch impliziert, dass dispositive und planerische Tätigkeiten durchaus zum Arbeitsinhalt des modernen Sachbearbeiters gehören.

Aus diesem Grund erscheint eine personengruppenspezifische Klassifikation betrieblicher Informationssysteme (wie in der Vergangenheit beispielsweise durch die Führungsinformationssysteme bzw. Executive Information Systems gegeben (Gluchowski et al. 2008, S. 74–82) nicht mehr zeitgemäß. Als sinnvoller erweist sich dagegen eine tätigkeitsorientierte Unterteilung der Systeme nach der Art der unterstützten Arbeitsinhalte. Grob lassen sich hier die beiden Klassen der operativen und der dispositiven bzw. analytischen Aufgaben voneinander abgrenzen.

Systeme, die mit einer ausgeprägten Transaktionsorientierung auf eine Unterstützung operativer Anwendungsfelder ausgerichtet sind, leisten heute in jeder Unternehmung unverzichtbare Dienste. So bilden Administrationssysteme den Einsatz der Elementarfaktoren (Potenzial- und Verbrauchsfaktoren) im Leistungsprozess einer Unternehmung ab und stellen damit Dokumentations- und Bewertungsfunktionalität bereit. Weiterhin können die Dispositionssysteme, die Steuerungs- und Lenkungsaufgaben im Falle klar strukturierter Entscheidbarkeit und Delegationsfähigkeit übernehmen, den **operativen Systemen** zugerechnet werden. Klassische Einsatzgebiete für die Administrations- und Dispositionssysteme sind u. a. die Verwaltung von Kunden-, Lieferanten- und Produktstammdaten oder die Erfassung, Bearbeitung und Kontrolle von Kundenaufträgen, Lagerbeständen, Produktionsvorgaben und Bestellungen. Operative Systeme sind heute in schlüsselfertiger Form auf der Basis betriebswirtschaftlicher Standardsoftware für nahezu jeden Anwendungsbereich erhältlich und präsentieren sich aufgrund des z. T. langjährigen Einsatzes als ausgereift und stabil.

Weit diffuser dagegen stellt sich die Situation bei der Unterstützung **dispositiver bzw. analytischer Tätigkeiten** dar. Die auftretenden Schwierigkeiten reichen von fehlender Interaktivität und übermäßiger Starrheit über Informationsüberflutung einerseits und man-

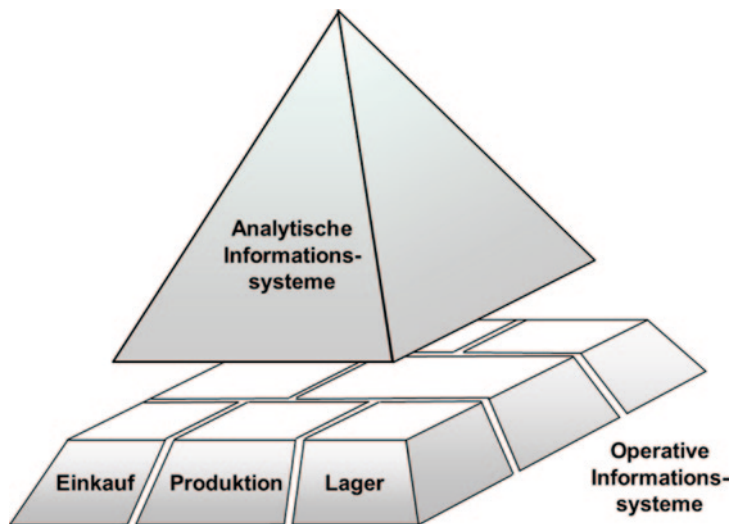


Abb. 1.1 Betriebliche Informationssystempyramide

gelnde Datenanbindung andererseits bis zur unzureichenden Integrierbarkeit in bestehende DV-technische und organisatorische Infrastrukturen.

In diesem Kontext sind Begrifflichkeiten und Konzepte wie „Data Warehouse“, „On-Line Analytical Processing“, „Data Mining“ und „Reporting“ sowie in den letzten Jahren auch „Big Data Analytics“ aufgegriffen und umgesetzt worden, die versprechen, Lösungen für die angeführten Probleme bieten zu können. Zum Teil erweisen sich diese Ansätze als schwer zugänglich und nur mit erheblichem Aufwand umsetzbar, zumal hierdurch nicht etwa „fertige Informationssysteme von der Stange“ repräsentiert werden, sondern Konzepte, die im spezifischen Einsatzbereich einzelfallbezogen umgesetzt werden müssen (Kemper et al. 2010).

Auf der Grundlage dieser technologischen Basiskonzepte lassen sich dann betriebswirtschaftliche Anwendungslösungen etwa für die Bereiche Planung und Budgetierung, Konsolidierung sowie analytisches Customer Relationship Management entwerfen und implementieren.

Diese aufgezählten technologischen Ansätze, die derzeit die Diskussion um die Ausgestaltung dispositiver Systeme dominieren, sowie die darauf aufsetzenden fachlichen Lösungen für dispositive Aufgabenstellungen werden im vorliegenden Sammelband unter der begrifflichen Klammer „**Analytische Informationssysteme**“ diskutiert. Diese Bezeichnung wurde gewählt, da sie einerseits noch nicht durch ein bestimmtes Begriffsverständnis belegt ist und andererseits zum Ausdruck kommt, dass die **Informationsversorgung und funktionale Unterstützung betrieblicher Fach- und Führungskräfte zu Analysezwecken** im Vordergrund stehen. Inhaltlich bilden damit die „Analytischen Informationssysteme“ das logische Komplement zu den operativen Informationssystemen (vgl. Abb. 1.1).

Synonym hierzu lässt sich das derzeit sehr populäre Wortgebilde Business Intelligence-Systeme verwenden (Kemper et al. 2010; Mertens 2002), das allerdings aufgrund der nahezu inflationären Verwendung zu verwässern droht (Gluchowski 2001; Grothe und Gentsch 2000). Bisweilen dient auch das Wortkonstrukt Performance Management-System zur Bezeichnung der Systemkategorie, wenngleich die ursprünglichen Gedanken hinter dem Business bzw. Corporate Performance Management eine stärkere betriebswirtschaftliche Fokussierung aufwiesen und sich explizit auch der engeren Verschmelzung von operativer und dispositiver Informationsverarbeitung verschrieb. Die traditionellen Ansätze (wie z. B. auch die Erkenntnisse des Operations Research) sollen an dieser Stelle nicht etwa vergessen werden, sondern gehen als integrativer Bestandteil darin auf.

Im vorliegenden Sammelband erfolgt eine Konzentration auf die aktuellen Schlagworte und Entwicklungstendenzen. Dazu gehören nach heutigem Verständnis insbesondere

- die **strategische** und **betriebswirtschaftlich-organisatorische** Einbettung entscheidungsunterstützender Systemlösungen in die Aufbau- und Ablaufstrukturen von Unternehmen, die zu einer tragfähigen **Informationslogistik** führen sollen (Teil 1 des Sammelbandes),
- der Aufbau eines unternehmensweiten, entscheidungsorientierten Datenpools, mit dem sich die unterschiedlichen analytischen Aufgaben wirksam unterstützen lassen und der als **Data Warehouse** bezeichnet werden kann, einschließlich der neuen und innovativen Speicher- und Ablageformen für Daten, mit denen sich auch unstrukturiertes Datenmaterial effektiv verarbeiten lässt (Teil 2 des Sammelbandes), sowie
- die **Anwendungen** analytischer Informationssysteme, die sich beispielsweise in den Themenbereichen **Planung** und **Budgetierung** manifestieren, dabei die Spezifika einzelner betrieblicher **Funktionsbereiche** (wie der Logistik und der Produktion) beachten müssen und sich fortgeschrittener **Visualisierungstechniken** bedienen (Teil 3).

Die folgenden Ausführungen sowie der gesamte vorliegende Sammelband folgt der logischen Einteilung in die Themenkomplexe Organisation, Technik und Fachlichkeit von Analytischen Informationssystemen.

1.3 Grundlagen und organisatorische Aspekte

Das Thema „Analytische Informationssysteme“ präsentiert sich bei näherer Betrachtung als komplex und vielschichtig. Um die einzelnen aktuellen Strömungen und Entwicklungslinien richtig einordnen und bewerten zu können, erweist sich ein Blick in die Historie der Analytischen Informationssysteme als guter Ausgangspunkt. Aus diesem Grunde reicht der Beitrag von **Schieder** mit dem Titel „*Historische Fragmente einer Integrationsdisziplin – Beitrag zur Konstruktgeschichte der Business Intelligence*“ bis zu den Wurzeln der Systemkategorie zurück und zeigt die wesentlichen Meilensteine mehr über ein halbes Jahrhundert chronologisch auf.

Beim der Gestaltung Analytischer Informationssysteme gilt es, für eine erfolgreiche Umsetzung tragfähigen Lösungen nicht nur technologische Aspekte zu beachten, sondern auch den betriebswirtschaftlich-organisatorischen Problemen sowie strategischen Fragestellungen angemessene Beachtung zu schenken.

Als Teil der IT-Strategie erweist sich auch die Strategie einer Unternehmung für den Themenkomplex der Analytischen Informationssysteme eng mit der grundlegenden Geschäfts- bzw. Unternehmungsstrategie verknüpft. Aufgabe ist es hier, einen langfristigen und unternehmungsweiten Orientierungsrahmen für die Ausgestaltung der zugehörigen Lösungen zu schaffen. Vor diesem Hintergrund löst sich der Beitrag von **Totok** mit dem Titel *„Von der Business-Intelligence-Strategie zum Business Intelligence Competency Center“* von technologischen Implementierungsdetails und konzentriert sich auf die Gestaltung langfristig tragfähiger Rahmenkonzepte – auch im Hinblick auf zu beachtende Organisationskonzepte und Wirtschaftlichkeitsgesichtspunkte.

Gänzlich neue Anwendungsbereiche erschließen sich die Analytischen Informationssysteme vor dem Hintergrund der zunehmenden Diskussion um Big Data. Der Beitrag von **Dittmar** widmet sich unter der Überschrift *„Die nächste Evolutionsstufe von AIS: Big Data – Erweiterung klassischer BI-Architekturen mit neuen Big Data Technologien“* dieser Facette des Themenkomplexes.

Speziell auf die Führungsaufgaben im Umfeld Analytischer Informationssysteme konzentriert sich **Winter** mit dem Beitrag *„Analytische Informationssysteme aus Managementsicht: lokale Entscheidungsunterstützung vs. unternehmensweite Informations-Infrastruktur“* und legt dabei den Fokus auf die fachliche Seite der Betrachtung zugehöriger Planungs- und Steuerungsaufgaben.

Der Aufsatz von **Bange** mit dem Titel *„Werkzeuge für analytische Informationssysteme“* leistet anhand möglicher Architekturbausteine einen Marktüberblick über die verfügbaren Produkte und Anbieter zum Aufbau analytischer Lösungen. Dabei erfolgt eine Orientierung an den drei Bereichen „Datenintegration und -aufbereitung“, „Datenspeicherung und -bereitstellung“ und „Anwendungen“.

1.4 Architektur und Technologien

Beim Aufbau analytischer Informationssysteme fallen 80% und mehr des Aufwandes für die Etablierung leistungsfähiger Zugriffsstrategien auf die vorgelagerten Informationsspeicher an. Der Schnittstelle zwischen operativen und analytischen Informationssystemen muss folglich besondere Beachtung geschenkt werden. Im vorliegenden Sammelband beginnt Teil II aus diesem Grund mit einer Erörterung der zentralen Aspekte der Datentransformation zwischen den Systemkategorien.

Dabei stellen **Kemper** und **Finger** mit dem Beitrag *„Transformation operativer Daten – Konzeptionelle Überlegungen zur Filterung, Harmonisierung, Aggregation und Anreicherung im Data Warehouse (DWH)“* die transformationsorientierten Gestaltungsaspekte in

den Vordergrund, die beim Aufbau konsistenter analytischer Datenspeicher zu bedenken sind.

Data Warehouse-Architekturen setzen sich aus unterschiedlichen logischen Komponenten zusammen, die im Idealfall optimal zusammenwirken. Bei der Ausgestaltung des Zusammenwirkens der Komponenten ergeben sich vielfältige Freiheitsgrade, die in verschiedenen Architekturvarianten münden. Der Beitrag von **Hahne** mit dem Titel „Architekturkonzepte und Modellierungsverfahren für BI-Systeme“ zeigt die Gestaltungsalternativen auf und präsentiert darüber hinaus unterschiedliche Modellierungsverfahren für das Core Data Warehouse.

Das über Jahrzehnte vorherrschende Paradigma der festplattenorientierten, relationalen Speicherung wird in letzter Zeit durch die verstärkte Verbreitung neuartiger Datenbanktechnologien in Frage gestellt. **Knabke** und **Olbrich** stellen in ihrem Beitrag „Grundlagen und Einsatzpotentiale von In-Memory-Datenbanken“ die Funktionsweise hauptspeicherorientierter Datenbanksysteme vor, ordnen diese in den BI-Kontext ein und diskutieren ausgewählte Anwendungsbeispiele.

Ebenfalls auf die Datenspeicherung konzentrieren sich **Lechtenböcker** und **Vossen** unter dem Titel „NoSQL, NewSQL, Map-Reduce und Hadoop“, wobei hier die Datenverarbeitung großer, gegebenenfalls unstrukturierter Datenbestände im Vordergrund steht.

Gluchowski greift in seinem Aufsatz „Entwicklungstendenzen bei Analytischen Informationssystemen“ aktuelle technologische, aber auch organisatorische Trends im Umfeld der betrachteten Systemkategorie auf, die sich vor allem in der Praxis derzeit ausmachen lassen.

Wenngleich das aufgezeichnete Spektrum informationstechnologischer Aspekte beim konkreten betrieblichen Einsatz genügend Problempotenzial aufweist, um entsprechende Projekte scheitern zu lassen, liegen die Schwierigkeiten oftmals in der Anwendung analytischer Informationssysteme im Kontext konkreter betriebswirtschaftlicher Problemstellungen, wie der folgende Abschnitt aufzeigt.

1.5 Betriebswirtschaftliche Anwendung und spezielle Aspekte der Analytischen Informationssysteme

Informationstechnologie und betriebswirtschaftliche Organisationsformen gehen häufig ein bemerkenswertes Wechselspiel ein. Einerseits gilt die technikunterstützte Informationsverarbeitung als „Enabler“ für die Anwendung neuer Organisationskonzepte, andererseits wird durch neuartige Aufbau- und Ablaufstrukturen oftmals der Bedarf nach entsprechenden Systemlösungen erst geweckt. So dürfte der Einsatz moderner Informationstechnik in einem starren Umfeld mit antiquierten Strukturen i. d. R. wenig erfolgversprechend sein.

Für den Einsatz Analytischer Informationssysteme gewinnen diese Überlegungen besondere Relevanz, da die herkömmlichen Informationswege und -inhalte sowie das bislang benutzte analytische Instrumentarium gänzlich neu überdacht und konzipiert werden

müssen. Aus diesem Grunde erlangen betriebswirtschaftlich-fachliche Fragestellungen hier ein erhebliches Gewicht. Nicht selten scheitern diesbezügliche Praxisprojekte durch innerbetriebliche Widerstände, sei es aus Angst vor dem Verlust von Wissens- und damit Machtmonopolen oder aus Kompetenz- und Besitzstreitigkeiten, wenn es um die Zurverfügungstellung und Verwaltung von Datenmaterial geht.

Deshalb wurde im vorliegenden Sammelband ein spezieller Abschn. (Teil 3) der Betrachtung betriebswirtschaftlicher Anwendungen Analytischer Informationssysteme sowie besonders interessanter Querschnittsthemen gewidmet.

In jeder Unternehmung erweist sich die Planung zukünftiger Geschäftszahlen und die Zuweisung ökonomischer Größen zu einzelnen Bereichen als zeitaufwändiges und anstrengendes Unterfangen. Oftmals führt nicht zuletzt die Verwendung unzureichender Werkzeuge zur Planung und Budgetierung zu diesem Ergebnis. Der Beitrag von **Oehler** widmet sich vor diesem Hintergrund mit dem Beitrag „*Planung und Informationstechnologie – Vielfältige Optionen erschweren die Auswahl*“ der Fragestellung, welche Funktionalitäten moderne Planungslösungen mit sich bringen müssen.

Große Umwälzungen durch den zunehmenden Einsatz von Sensorik und die damit einhergehende Datenverfügbarkeit sind derzeit vor allem im industriellen Sektor auszumachen. Unter der Überschrift „*Operational Business Intelligence im Zukunftsszenario der Industrie 4.0*“ greift der Aufsatz von **Hänel** und **Felden** die besonderen Herausforderungen auf, die dabei für Analytische Informationssysteme mit operativer Ausrichtung erwachsen.

Der Einsatz von Informationssystemen muss den Spezifika des jeweiligen Anwendungsfeldes gezielt Rechnung tragen. Dass sich dabei in den einzelnen Funktionsbereichen einer Unternehmung ganz besondere Herausforderungen ergeben, belegt der Beitrag von **Baars** und **Lasi** zum Thema „*Innovative Business-Intelligence-Anwendungen in Logistik und Produktion*“.

Der Endanwender bewertet die Leistungsfähigkeit eines Analytischen Informationssystems letztlich danach, was er am Bildschirm sieht und wie er mit den angebotenen Inhalten interagieren kann. Dabei nehmen beispielsweise Faktoren wie Antwortzeitgeschwindigkeit, Stabilität und Navigationsfunktionalität eine wesentliche Rolle ein. Daneben erweist sich die Art und Weise der Aufbereitung und Präsentation von Informationen als zentrales Kriterium, um auch komplexe Zusammenhänge rasch und fehlerfrei verstehen und interpretieren zu können. An dieser Stelle setzt der Text von **Kohlhammer**, **Proff** und **Wiener** zur „*Der Markt für Visual Business Analytics*“ an, der die herkömmlichen Verfahren zur grafischen Darstellung von Geschäftsdaten kritisch hinterfragt, Verbesserungspotenziale aufzeigt und innovative Präsentationsalternativen anbietet.

Schließlich widmen sich **Linden** und **Navrade** einer Betrachtung der inhaltlichen Ausgestaltung eines Analytischen Informationssystems, indem Sie die „*Effektive Planung und Steuerung erfolgskritischer Komponenten eines Geschäftsmodells*“ in den Fokus der Betrachtung rücken und dabei auch auf Verfahren zur Informationsbedarfsanalyse blicken.

Insgesamt kann festgehalten werden, dass sowohl Data Warehouse-Konzepte als auch OLAP- und Data Mining-Technologien heute zumindest in größeren Unternehmen

gängige Praxis darstellen. Durch die zunehmende Nutzung als Basistechnologien für betriebswirtschaftliche Anwendungen erlangen die Schlagworte auch in den Fachabteilungen einige Verbreitung. Dabei löst sich der Schwerpunkt der Diskussion immer stärker von technischen Fragestellungen und wendet sich betriebswirtschaftlich-fachlichen Aspekten zu.

Literatur

- Gluchowski, P.: Business Intelligence. Konzepte, Technologien und Einsatzbereiche. HMD – Theorie Praxis Wirtschaftsinform. **38**(222), 5–15 (Dez. 2001)
- Gluchowski, P., Gabriel, R., Dittmar, C.: Management Support Systeme und Business Intelligence. Computergestützte Informationssysteme für Führungskräfte und Entscheidungsträger, 2. Aufl. Springer Berlin (2008)
- Grothe, M., Gentsch, P.: Business Intelligence. In: Aus Informationen Wettbewerbsvorteile gewinnen, München (2000)
- Kemper, H.-G., Baars, H., Mehanna, W.: Business Intelligence – Grundlagen und praktische Anwendungen, 3. Aufl. Vieweg Wiesbaden (2010)
- Mertens, P.: Business Intelligence – ein Überblick. Arbeitspapier Nr. 2/2002, Bereich Wirtschaftsinformatik I. Universität Erlangen-Nürnberg (2002)

Historische Fragmente einer Integrationsdisziplin – Beitrag zur Konstruktgeschichte der Business Intelligence

Christian Schieder

Inhaltsverzeichnis

2.1	Einleitung	14
2.2	Historische Fragmente zur Chronologie der Business Intelligence	16
2.3	Etymologische Annäherung an Business Intelligence – Taxonomie des State-of-the-Art	21
2.4	Situation Awareness als Zielkonstrukt einer kognitionsorientierten Gestaltung der Business Intelligence	25
2.5	Zusammenfassung und Ausblick	27
	Literatur	30

Zusammenfassung

Ein Vergewegenwärtigen der Ursprünge erlaubt ein tieferes Verständnis rezenter Erscheinungsformen jeglicher zu beobachtender Phänomene. Dieses Wissen wiederum eröffnet Möglichkeiten zur Bewertung ihrer Entwicklungstendenzen. Das Phänomen der Business Intelligence als integrierte, IT-basierte Management- und Entscheidungsunterstützung ist keine neue Erscheinung des sogenannten Informationszeitalters. Vielmehr ist sowohl die Verwendung des Begriffs, als auch die wissenschaftliche und praktische Beschäftigung mit den damit bezeichneten Aufgaben, Methoden und Techniken auf die Mitte des 20. Jahrhunderts zurück zu datieren – in ihren Grundlagen sogar noch weiter.

C. Schieder (✉)
Technische Universität Chemnitz, Lehrstuhl für Wirtschaftsinformatik II, Chemnitz, Deutschland
E-Mail: christian.schieder@wirtschaft.tu-chemnitz.de

Um den Boden für ein tieferes Verständnis der aktuellen und der sich andeutenden zukünftigen Gestalt der Business Intelligence zu bereiten, zeichnet der vorliegende Beitrag – ausgehend von diesen Grundlagen – die Entwicklungslinien analytischer Informationssysteme bis zu den heutigen Business-Intelligence-Systemen nach. Ziel des Beitrags ist es, im Sinne einer Phänomenologie der Business Intelligence, chronologische, morphologische und taxonomische Fragmente zu sammeln, aufzubereiten und zu untersuchen. Das Leitbild des Beitrags liegt statt in einer vollständigen, in einer ganzheitlichen und statt in einer bewertenden, in einer beschreibenden Analyse des Phänomens Business Intelligence. Den Schwerpunkt der Untersuchung bilden dabei die derzeit in der Praxis erfahrbaren und in der Wissenschaft thematisierten Erscheinungsformen des Untersuchungsgegenstands.

2.1 Einleitung

Das legendäre Blechtley-Park, in dem ein Team um Alan Turing 1942 den Code der Enigma entschlüsselte, war nur ein kleiner Teil eines umfassenden „Intelligence“-Systems der britischen Regierung unter W. Churchill (Kozak-Holland 2005). Zusammen mit den Einrichtungen in Storey’s Gate (einem unterirdischen Komplex in London) sollte eine ausgefeilte Informationsinfrastruktur dabei helfen, sowohl die aktuelle Lage auf den Kriegsschauplätzen, als auch die Erfüllung von Produktionszielen kriegswichtiger Industrien über die gesamte Wertschöpfungskette hinweg nahezu in Echtzeit zu überwachen und zu steuern. Im „Cabinet War Room“ traf sich Churchill mit seinen Ministern, um jene strategischen Entscheidungen zu treffen, die letztlich dazu führten, dass Großbritannien von einer Invasion verschont blieb und sich langfristig im 2. Weltkrieg behaupten konnte. Zur Vorbereitung und Unterstützung dieser Entscheidungen kam ein breites Methodenspektrum zur Anwendung, wie zum Beispiel Portfolioanalysen, Szenariotechniken und Multiprojektmanagement. Einige Methoden wurden in diesem Kontext überhaupt erst entwickelt. Seltsam vertraut erscheint dabei eine Analysten-Beschreibung der zu diesem Zweck eingerichteten „Intelligence“-Infrastruktur: „*It’s purpose was getting the right information to the right person in the right timeframe.*“ (Kozak-Holland 2005).

Zentrale Treiber der methodischen, organisatorischen und technischen Entwicklungen im Umfeld der britischen „Intelligence“ war die eklatante Ressourcenknappheit Großbritanniens zu Beginn des 2. Weltkriegs. Dieser grundlegende Umstand galt für alle kriegsführenden Parteien gleichermaßen. So ist es nicht verwunderlich, dass viele Entwicklungen im Bereich der Informationsverarbeitung ihren Anfang in militärischen Anwendungsbereichen nahmen. Die Operations Research ist dabei nur ein weiteres Beispiel, bei dem die Optimierung militärischer Ressourcenallokationen im Zentrum der Bemühungen stand. Entsprechende Fragestellungen beschäftigten sich mit der Entwicklung von Heuristiken zur Bestimmung der optimalen Größe des Begleitschutzes in Abhängigkeit der Anzahl von Schiffen in Konvois. Der bekannte OODA-Loop (Observe, Orient, Decide,

Act) von Boyd (Curts und Campbell 2001) ist ein weiteres Beispiel. Als Air-Force-Analystiker war es Boyds Absicht, das Entscheidungsverhalten von Kampfpiloten zu analysieren und Muster für erfolgreiche Luftkampfeinsätze bei technischer Unterlegenheit abzuleiten (Houghton et al. 2004). Er kam zu dem Ergebnis, dass diejenigen Piloten erfolgreicher waren als ihre Gegner, die in der Lage waren, eben jenen OODA-Zyklus schneller zu durchlaufen. In den 1990er Jahren wurde dieses Konzept von verschiedenen Autoren zur Strukturierung des Entscheidungsverhaltens im Unternehmensumfeld popularisiert und erfährt dort ungebrochen hohe Aufmerksamkeit (El Sawy und Majchrzak 2004).

Neben zahlreichen Impulsen aus dem militärischen Bereich, gehen wichtige Grundlagen der modernen Massendatenverarbeitung allerdings auch auf zivile Anwendungsfelder zurück. Ein Beispiel hierfür ist die Lochstreifenmaschine: zur Bewältigung der großen Anzahl an Datensätzen, die bei der Durchführung der Volkszählung von 1890 anfielen, setzte das US-amerikanische Bureau of the Census im selben Jahr ein solche Maschine zum ersten Mal zur automatisierten Datenverarbeitung ein (Boyd und Crawford 2012). Damit wurde eine Entwicklung begründet, die in den 1960er Jahren in die Implementierung relationaler Datenbanken mündete. Fortschritte im Bereich der relationalen Datenbanken wiederum machten den Aufbau dedizierter, themenorientierter Berichts- und Analysedatenbestände möglich, die ab Ende der 1980er Jahre unter dem Begriff „Data Warehouse“ Verbreitung fanden. Die Entwicklung der integrierten, IT-basierten Management- und Entscheidungsunterstützung ließe sich sicherlich noch weiter zurückverfolgen – im Grunde bis zur Erfindung der Schrift, deren Notwendigkeit gleichsam im Monitoring von Mengen- und Wertgrößen begründet lag und die damit die Basis für die Analyse und Optimierung von Güter- und Wertströmen legte. Die erforderliche Sammlung, Integration und Aufbereitung von Informationen zum Zwecke der Entscheidungsunterstützung – nicht nur aber vor allem auch im Management – fungierte daher seit jeher als Innovationsmotor und führte zu einigen der bemerkenswertesten Errungenschaften unserer Zivilisation.

Das vorerst letzte Kapitel dieser Entwicklungen wurde vor nunmehr fast 60 Jahren eröffnet und legte den Grundstein für die moderne, integrierte, IT-basierte Entscheidungsunterstützung im Management. Die zugehörigen Themen werden derzeit sowohl in der Praxis (Luftman und Ben-Zvi 2011) aber auch in der Wissenschaft (Steininger et al. 2009) unter dem Begriff „Business Intelligence“ (BI) thematisiert. In der Wissenschaft findet sich in weitgehend synonyme Verwendung hierzu vermehrt auch der Begriff der analytischen Informationssysteme (AIS). Chameni und Gluchowski (2010) führen die Wahl dieser Bezeichnung auf eine komplementäre Abgrenzung zur Kategorie der operativen, transaktionsorientierten Informationssysteme zurück. So soll der Zweckfokus der AIS betont werden, der auf der Informationsversorgung und der funktionalen Unterstützung betrieblicher Fach- und Führungskräfte zur Bewältigung von Analyseaufgaben im Rahmen der Entscheidungsfindung liegt. Die BI dient somit zur Realisierung einer integrierten Informationslogistik (IIL) zur Planung, Steuerung, Durchführung und Kontrolle der Gesamtheit der Datenflüsse in Unternehmen zur Unterstützung von Entscheidungen sowie die Aufbereitung und Speicherung der durch sie bewegten Daten (Winter 2010).

Historische Entwicklungen nachzuzeichnen haftet immer der Makel der subjektiven und zum Teil willkürlichen Auswahl der verwendeten Quellen und Entwicklungslinien an – dies darf für die Business Intelligence gleichermaßen gelten, wie für ihre Mutterdisziplin, der Wirtschaftsinformatik (Heinrich 2005). Gleichzeitig liegen zur Geschichte der BI und der angrenzenden Disziplin der Decision-Support-Systeme (DSS) bereits etliche Beiträge aus verschiedenen Perspektiven vor (Adam 2012; Gluchowski und Kemper 2006; Hosack et al. 2012; Power et al. 2011; Power 2003). Die nachfolgenden Ausführungen stellen deshalb vielmehr einen Versuch dar, bisher unberücksichtigte Fragmente ans Licht zu bringen. Dabei gelingt es mit Hilfe einer phänomenologischen Beschreibung dem Vorhaben ein gewisses Maß an Struktur zu verleihen.

Der Beitrag gliedert sich daher in drei Teile. Er befasst sich mit der Chronologie der BI, der Taxonomie ihres State-of-the-Arts und der zukünftigen Gestalt einer kognitionsorientierten BI. Die historisch-chronologische Sicht fokussiert auf die zeitliche Zuordnung einzelner Schlaglichter technologischer Innovationen, welche die heutige Gestalt der BI maßgeblich mitprägten. Kapitel 2 des Beitrag legt hierauf den Fokus und betrachtet insbesondere die Konstruktgeschichte der BI, also was wann zum Bedeutungsfeld des Begriffs gehörte. Die morphologischen Untersuchungen in Kap. 3 befassen sich mit dem State-of-the-Art der Gestalt beziehungsweise der Erscheinungsformen moderner BI-Systeme. Die Struktur der Darstellung orientiert sich dabei an etymologischen Betrachtungen des BI-Begriffs. Kapitel 4 führt das Konstrukt der Situation Awareness in den Kontext der BI als Zielkonstrukt für eine kognitionsorientierte Gestaltung der zukünftigen BI. Kapitel 5 fasst den Beitrag abschließend zusammen und gibt einen Ausblick, indem es die Popularitätstrends der Begriffe BI und Big Data gegenüberstellt.

2.2 Historische Fragmente zur Chronologie der Business Intelligence

Der nachfolgende Abschnitt befasst sich mit einer Chronologie der Business Intelligence. Dabei meint Chronologie die Beschreibung der Abfolge des Auftretens und der Verwendung des Begriffs und seiner Bestandteile im Sinne einer Chronik. Zu diesem Zweck müssen zunächst historische Vorkommen der Begriffsverwendung lokalisiert und gesammelt werden, um sie anschließend zu katalogisieren und insbesondere zeitlich zu ordnen. Begriffe und die ihnen unterliegenden Bedeutungsfelder sind einem steten Wandel unterworfen. Das gilt auch für künstlich geschaffene Konstrukte wie das der „Business Intelligence“. Um das Bedeutungsfeld des Begriffs zu erschließen wird er in der Folge mit verwandten Konzepten in Zusammenhang gesetzt und von ihnen abgegrenzt.

Die ursprüngliche Wortschöpfung „Business Intelligence“ als zusammengesetztes Kunstwort wird allgemein dem deutschstämmigen IBM-Forscher Hans Peter LUHN zugeschrieben, der in einem 1958 erschienenem Artikel erstmalig ein technisches BI-System beschreibt. Er versteht darunter ein

„automatic system [...] to disseminate information to the various sections of any industrial, scientific or government organization. This intelligence system will utilize

data-processing machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the ‘action points’ in an organization. Both incoming and internally generated documents are automatically abstracted, characterized by a word pattern, and sent automatically to appropriate action points.” (Luhn 1958, S. 314).

LUHNS BI entsprach damit einem regelbasierten Wissensmanagementsystem. Es verfolgte den Zweck, erforderliche Wissensfragmente in Form von Dokumenten den Stellen (,action points‘) zu zuführen, die auf diese Information zur Durchführung ihrer betriebswirtschaftlichen (Entscheidungs-)Aufgaben angewiesen waren. Er verstand in diesem Zusammenhang die Rolle der ,Intelligence‘ als Fähigkeit zur Antizipation von Abhängigkeiten zwischen Fakten, um die zur Zielerreichung notwendigen Aktivitäten anzuleiten. Abbildung 2.1 zeigt eine Reproduktion der schematischen Darstellung von LUHNS

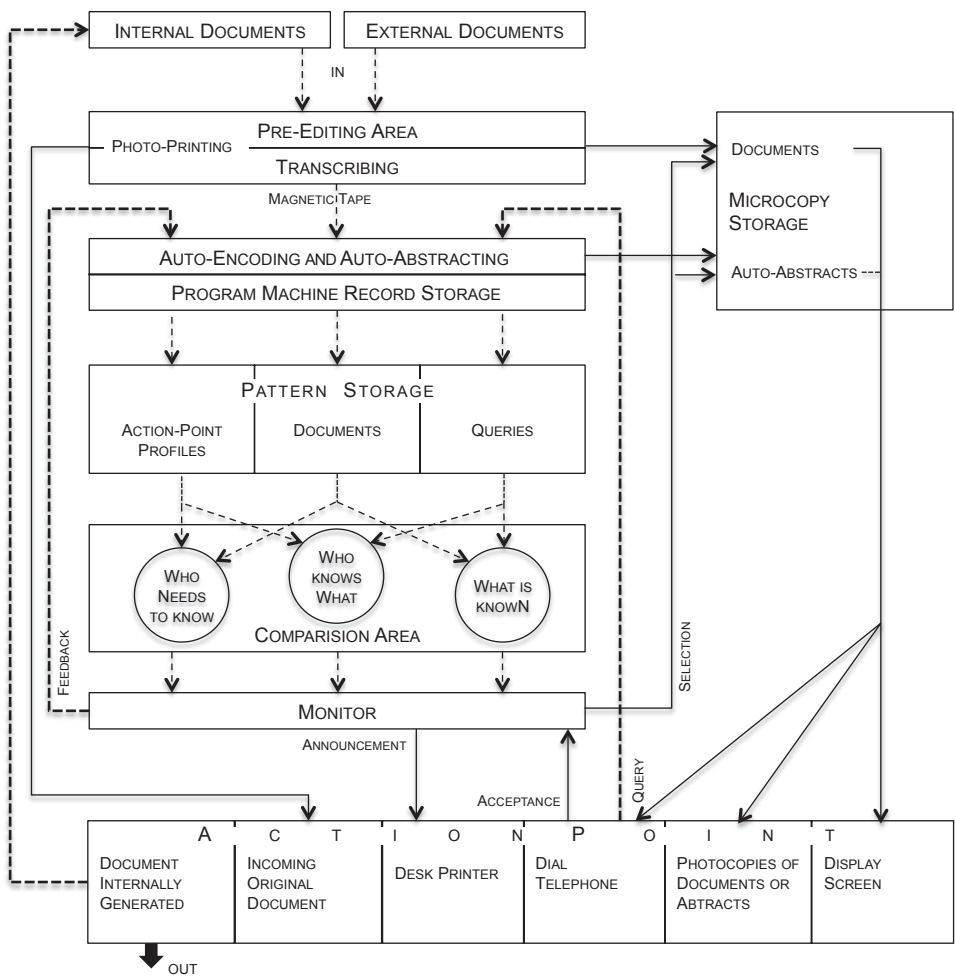


Abb. 2.1 LUHNS Business Intelligence System von 1958

System. Aus der Abbildung gehen die Informations- und Feedbackflüsse hervor (Pfeile). LUHN geht davon aus, dass eingehende interne und externe Dokumente in einer Pre-Editing-Area zunächst transkribiert werden müssen, bevor sie anschließend automatisch mit Metadaten versehen werden können (Auto-Encoding and Auto-Abstracting). Die Originaldokumente werden dabei entsprechend archiviert, während die Informationen auf Basis der Metadaten in mit Hilfe eines Pattern Storage mit Action-Point-Profilen, anderen Dokumenten und Anfragen in Zusammenhang gesetzt werden (Comparison Area). Auf diese Weise wird ermittelt, wer die verfügbaren Informationen benötigt, wer zusätzliche Informationen dazu hat und was darüber hinaus an Informationen zur Verfügung steht. Eine Monitor-Komponente leitet diese Information dann an die betroffenen Stellen weiter.

Abgesehen von der heute antiquiert anmutenden technischen Umsetzung mit Hilfe von Druckern, Telefonen und Magnetbändern, erweist sich LUHNS Konzeption als nachhaltig gültig und seiner Zeit weit voraus. Insbesondere die hervorgehobene Bedeutung, die er der Verarbeitung von Metadaten zuweist, erfährt heute vor dem Hintergrund des zunehmenden Anteils polystrukturierter, großer Mengen an Daten wieder steigende Relevanz (Schieder et al. 2015).

LUHNS modern erscheinendes, umfassendes Verständnis einer BI, konnte sich allerdings zunächst nicht durchsetzen. Entgegen der Konzeption LUHNS stand in den darauffolgenden Jahrzehnten die Kategorie der Decision-Support-Systeme (DSS) beziehungsweise Entscheidungsunterstützungssysteme (EUS) im Fokus der Betrachtungen. Maßgeblichen Einfluss darauf hatten die Arbeiten zur Automatisierung von Entscheidungsprozessen insbesondere von (Simon 1966), aber auch von (Scott 1967). DSS wurden neben Managementinformationssystemen (MIS) unter dem Oberbegriff der Management-Support-Systeme (MSS) diskutiert. Während bei historischen MIS häufig eine Abbildung des gesamten Unternehmens angestrebt wurde (Weber und Schäffer 2006), ging es bei der Gestaltung von DSS um die partielle „Abbildung des Verhaltens von Fach- und Führungskräften bei der Lösung von [speziellen] Fachproblemen“ (Gluchowski et al. 2008). Beiden Systemkategorien gemein war das Ziel, mit Hilfe ihres Einsatzes unter Verwendung des jeweils verfügbaren Entwicklungsstands der Informationstechnologie, einen Teil des Spektrums des Unterstützungsbedarfs von Managern abzudecken (Schieder 2014).

Intensive Recherchen deuten darauf hin, dass es tatsächlich bis ins Jahr 1986 dauerte, bevor erste Arbeiten den Begriff der BI wieder explizit aufgriffen. Sie erschienen im Sloan Management Review und stammen von Gilad und Gilad (1986) und Ghoshal und Kim (1986). Sowohl (Gilad und Gilad 1986) als auch (Ghoshal und Kim 1986) sehen BI darin vor allem unter dem Aspekt der Gewinnung von geschäftskritischen Erkenntnissen aus externen Informationen. Zu diesen externen Informationen zählen solche über Märkte, neue Technologien, Kunden, Wettbewerber oder auch breite soziale und politische Trends. Im angelsächsischen Sprachraum wurde der Teilaspekt der Aufklärung externer Erkenntnisse über Wettbewerber später als ‚Competitor Analysis‘ bezeichnet. Im deutschsprachigen Raum verwendete (Lange 1994) hierfür in den 1990-iger Jahren zunächst den Begriff der „technologischen Konkurrenzanalyse“, bevor sich auch hierfür der englische Terminus „Competitive Intelligence“ durchsetzte (Meier und Mertens 2004).

In ihrem Beitrag beschreiben (Gilad und Gilad 1986, S. 53) einen Prozess zur Generierung von (Business) Intelligence und betten diesen in ein entsprechendes organisationales System ein, ein Netzwerk aus Regeln und Rollen, welches einen BI-Prozess ausführt. Sie unterteilen den Prozess in fünf Phasen, die dem prinzipiellen Aufbau heutiger BI-Prozesse entsprechen:

1. Sammlung von Daten,
2. Bewertung der Gültigkeit und Zuverlässigkeit der Daten,
3. Datenanalyse,
4. Speicherung von Daten und Erkenntnissen sowie
5. Verteilung von Daten und Erkenntnissen.

Gegenüber dem heutigen Begriffsverständnis stand, wie auch beim ursprünglichen Wort-sinn der ‚intelligence‘ (siehe Kap. 3), die Verarbeitung unstrukturierter Daten und die Nutzbarmachung externer Informationen im Vordergrund. In ihren Ausführungen fokussieren die Autoren auf die Formalisierung und Institutionalisierung des vorgeschlagenen Prozesses und liefern zwei entgegengesetzte Vorschläge für eine zentralisierte und eine dezentralisierte, intelligence-zentrierte Geschäftsarchitektur (siehe Abb. 2.2). Dabei stehen jeweils dedizierte BI-Units im Zentrum der Betrachtungen. In der zentralisierten Variante (linker Teil der Darstellung) existiert eine Corporate BI Unit (CBIU), die für die zentrale Sammlung, Bewertung und Auswertung von (strukturierten) Daten verantwortlich zeichnet. Sie

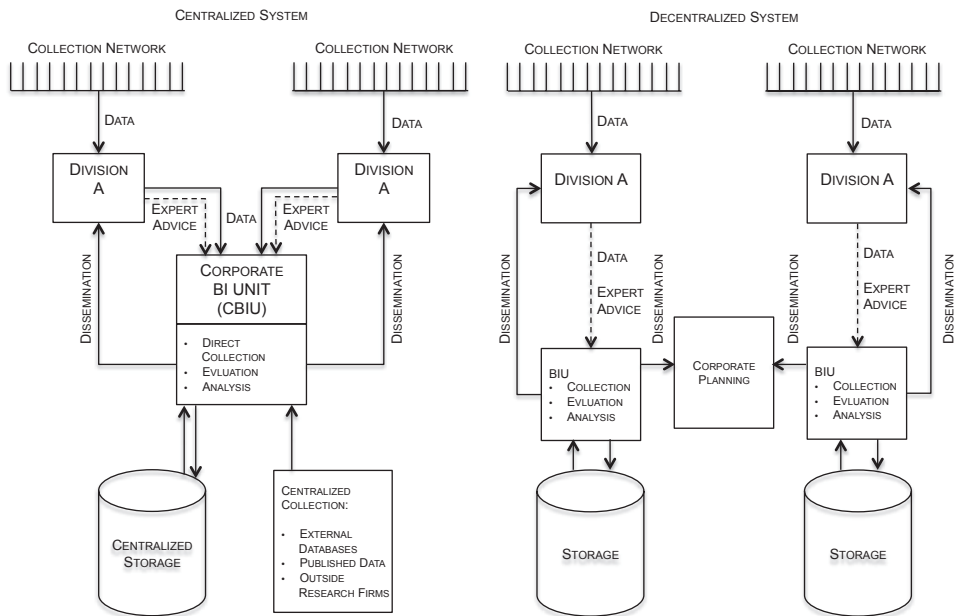


Abb. 2.2 Zentralisierte vs. dezentralisierte BI-Architektur nach. (Ghoshal und Kim 1986)

holt sich dazu fachliche Expertise aus den datenliefernden Geschäftseinheiten (Divisions) ein und stellt diesen die ausgewerteten Informationen bereit. Die dezentralisierte Variante zeichnet sich demgegenüber durch geschäftsbereichsbezogene BI-Units aus, die lediglich durch eine Corporate-Planning-Einheit miteinander verbunden sind.

Ghoshal und Kim (1986) hingegen stellen die „Gestalt“ eines BI-Systems als sozio-technisches System zur Generierung wettbewerbsrelevanter Erkenntnisse in den Vordergrund. Sie verstanden BI als „*collection and analysis of information on markets, new technology, customers, competitors, and broad social trends*“ (Ghoshal und Kim 1986, S. 49). Sie gehen von der These aus, dass das Vorhandensein einer BI-Organisation für sich genommen keine Garantie dafür ist, dass Daten und Informationen die Entscheider erreichen und von diesen auch wahrgenommen werden. Dabei richten sie das Augenmerk ihrer Analyse darauf, wie die intrinsischen sozialen und politischen Prozesse eines Unternehmens zur Errichtung einer gewinnbringenden BI-Organisation beitragen oder dies verhindern. Basierend auf einer Fallstudie leiten sie organisatorische Gestaltungsempfehlungen ab, deren Umsetzung die Gewinnung und die Verwendung von nutzenstiftenden Informationen befördern sollen.

In beiden vorgenannten Veröffentlichungen nehmen Aspekte der organisatorischen Umsetzung von BI-Systemen großen Raum ein. Sie sehen Technologie lediglich als Vehikel für Verdichtung von Daten und Gewinnung von Information an, die unabdinglich in einen organisationalen Kontext eingebunden werden muss. Diese sozio-technischen Aspekte der BI rücken auch in letzter Zeit wieder ins Zentrum der Betrachtungen und zeugen von zeitlos gültigen Grundprinzipien für die Gestaltung der BI (Foley und Guillemette 2010; Horakh et al. 2008).

Wenig später, im Jahr 1989, übernahm der spätere „Gartner Group“-Analyst Howard DRESNER den Terminus „Business Intelligence“ als Sammelbegriff, zur Beschreibung von „*concepts and methods to improve business decision making by using fact-based support systems*“ (Power 2003). Damit legte er den Grundstein für ein weithin antizipiertes und – in Anbetracht der vorstehenden Ausführungen – tendenziell verengtes, technikzentriertes BI-Verständnis. Gleichwohl muss DRESNER für seine Rolle in der Popularisierung des BI-Begriffs, vor allem in der Praxis, gewürdigt werden. Wichtige Prinzipien, um die das BI-Konzept zu dieser Zeit angereichert wurde, umfassen vor allem die Erkenntnis nach der Notwendigkeit einer integrierten Datenhaltung (im Gegensatz zum Datensilo) und einer entsprechenden Infrastruktur. Im Laufe der 1990er Jahre gelangten die unter dem Begriff BI diskutierten Anwendungen und Infrastrukturkomponenten in der Breite zur Anwendung. Möglich wurde dies durch die Verbreitung neuer Basisinformationstechnologien wie dem Internet und großen Datenbanksystemen (Data Warehousing).

Im Zuge der Popularisierung von DRESNERS BI-Begriff war BI lange Zeit auf technologische Aspekte fokussiert. Dazu trugen weitere Entwicklungen in der Anwendung analyseorientierter Informationssysteme durch die Ansätze des Online Analytical Processing und des Data Mining bei, die in den 1990er Jahren große Popularität erfuhren (Chaudhuri und Dayal 1997; Codd et al. 1993; Fayyad et al. 1996; Gluchowski 2001). Gleichzeitig rückten weitere technologieorientierte Aspekte von BI in den Vordergrund der Betrachtung.

tungen. Dazu gehörten Themen wie Integration von Massendaten mit Hilfe von Extraktions-, Transformations- und Ladeprozessen (ETL), die Vorteilhaftigkeit und die Gestaltung unterschiedlicher Architekturen von Data-Warehouse-Systemen (INMON vs. KIMBALL, die Modellierung von abfrageorientierten Datenschemata (Star Schema) sowie Bestrebungen zur Verbesserung und Sicherstellung einer angemessenen Datenqualität (Wang und Strong 1996; Wand und Wang 1996).

Mit Beginn des 21. Jahrhunderts setzte eine Konsolidierung von Technologien im Kontext der BI ein, die sich gegenwärtig durch Trends wie BI-as-a-Service und Cloud Computing weiter zu verstärken scheint (Gluchowski et al. 2011). Zunehmend realistischer erscheint damit die Verwirklichung alter Versprechen der BI von einer fachlich getriebenen Lösungsentwicklung an Stelle einer technik-zentrierten Softwareentwicklung. So ist es nicht verwunderlich, dass nach einer Phase der Technikdominanz wieder vermehrt strategische (Totok 2010), organisatorische (Dittmar und Ossendoth 2010) und integrative (Dinter et al. 2010) Dimensionen der BI im Interessensfokus von BI-Forschern und -Praktikern stehen.

Basierend auf dem historischen Verständnis der BI gilt der nächste Abschnitt der Untersuchung der Verständnisebenen des modernen BI-Begriffs. Auf diese Weise lassen sich die Dimensionen einer modernen BI-Konzeption differenzieren. Sie werden im Folgenden hergeleitet und einzeln dargestellt.

2.3 Etymologische Annäherung an Business Intelligence – Taxonomie des State-of-the-Art

Nach der chronologischen Aufarbeitung der Veränderungen im Bedeutungsfeld der BI dient der folgenden Abschnitt der Untersuchung der prinzipiellen Verständnisebenen des BI-Begriffs auf Basis seines etymologischen Bedeutungsfelds. Der englische Terminus „Business Intelligence“ findet in der Fachliteratur in vielen verschiedenen Facetten Verwendung. Die Bedeutungsvarianten lassen sich vor allem auf den semantischen Polymorphismus des Begriffs der „intelligence“ zurückführen. Die folgenden Ausführungen greifen hierfür auf eine etymologische (nach der Wortherkunft) und lexikalische (nach der Wortverwendung) Analyse der beiden Bestandteile des Kompositums BI zurück. Auf Basis dieser Analysen werden anschließend Bedeutungsebenen eines umfassenden BI-Verständnisses differenziert dargestellt. So lässt sich gleichsam eine Taxonomie des State-of-the-Arts des BI-Begriffs skizzieren und die verschiedenen Verständnisebenen in Beziehung zueinander setzen.

Die etymologischen Begriffswurzeln der „intelligence“ liegen im Lateinischen „intellegentia“ begründet, wo sie als Bezeichnung für den Akt der Auswahl („legere“) zwischen („inter“) Alternativen fungieren. Über das Altfranzösische „intelligence“ wurde der Begriff im späten 14. Jahrhundert ins Englische übernommen. Etwa in der Mitte des 15. Jahrhunderts verschob sich der Bedeutungsinhalt. Zu dieser Zeit trat eine Verwendung für Nachrichten und Informationen im Allgemeinen und geheime Informationen von Spionen

im Speziellen an die Stelle der ursprünglichen Semantik (Online Etymology Dictionary 2012). Heute steht der Begriff für bestimmte Arten von Informationen oder Erkenntnissen, in der Regel solchen, die als vertraulich oder sehr wichtig klassifiziert werden („*gathering intelligence*“ Oxford Online Dictionary 2012). Man kann an dieser Stelle von einem resultativen Bedeutungsaspekt sprechen, da die „intelligence“ in diesem Fall als Bezeichnung für das Resultat einer Handlung Verwendung findet. In deutschen Erklärungen zum BI-Begriff wird aus diesem Verständnis heraus daher häufig das Bedeutungsfeld der Begriffe Erkenntnis (über etwas) oder Einsicht (in etwas) als Entsprechung angeführt.

Auf die angeführte Bedeutungsverschiebung gegen Ende des Mittelalters geht die Prägung des Begriffs als *Terminus technicus* für die Bezeichnung von militärischen und politischen Informationen zurück. Bis ins 19. Jahrhundert hinein war diese Verwendung des Intelligenzbegriffs in der Hochsprache gebildeter Kreise des deutschen Sprachraums gebräuchlich. So erschienen behördliche oder kommunale Bekanntmachungen in dieser Zeit häufig in einem „Amts- und Intelligenzblatt“, so zum Beispiel das „Salzburger Amts- und Intelligenzblatt“ oder das „Königlich bayrische Amts- und Intelligenzblatt für die Pfalz“ (Böning 2004). Auf Grund der Spezialisierung der Bedeutung auf politische und militärische Inhalte wurde der Begriff sukzessive auf die mit der Beschaffung dieser „intelligence“ betrauten (politischen oder militärischen) Organisationen übertragen („*people employed in the collection of military or political information*“ Oxford Online Dictionary 2012). Diese Begriffsverwendung lässt sich als institutionell charakterisieren. Im Deutschen findet dieser Bedeutungsaspekt gelegentlich in der Bezeichnung der sozialen Gruppe der Intellektuellen und/oder der wissenschaftlich Gebildeten in einer Gesellschaft als der „Intelligenz“ seinen Niederschlag (Duden online 2012). Im Allgemeinen wird der entsprechende Bedeutungsgehalt im politischen oder militärischen Bereich (zum Beispiel taktische Aufklärung) im Deutschen jedoch durch die Bezeichnungen „Nachrichten-“ beziehungsweise „Aufklärungsdienst“ transportiert. In dieser Form ist er bis heute in Verwendung (zum Beispiel beim „Bundesnachrichtendienst“).

Das spezifizierende Substantiv „business“ liefert den Kontext der „intelligence“. Damit wird eingeschränkt worüber Erkenntnisse vorliegen sollen. Der Begriff „business“ ist in diesem Zusammenhang am ehesten mit „Geschäftsbetrieb“ wiederzugeben. Damit kann zugleich sowohl der Betrieb als wirtschaftliche Einheit gemeint sein als auch die Gesamtheit der in ihm ablaufenden Aktivitäten (unter anderem Geschäftsprozesse), die zur Aufrechterhaltung des Betriebs notwendig sind. In diesem Sinne brauchen Entscheider in Unternehmen also Erkenntnisse über den Zustand der wirtschaftlichen Einheit ihres Verantwortungsbereichs zur Steuerung und Sicherstellung des zielgerichteten Ablaufs der „lebenserhaltenden“ Geschäfts- und Unterstützungsprozesse.

Über diese sprachliche Ableitung des Begriffs „intelligence“ lassen sich in der einschlägigen Literatur (insbesondere der deutschsprachigen) weitere Bedeutungsfelder des BI-Begriffs identifizieren. Dabei handelt es sich zumeist um Definitionen, die (zumindest implizit) den Methoden- oder den Prozesscharakter der BI betonen. Sie beschäftigen sich dann damit, welche Mittel und Techniken eingesetzt werden, um Erkenntnisse über einen (geschäfts-) relevanten Sachverhalt zu gewinnen und wie die Abfolge der Tätigkeiten ihrer Gewinnung

(zeitlich) geordnet ist (siehe zum Beispiel Foley und Guillemette 2010). Zusätzlich zu den oben differenzierten sprachlichen Betrachtungen lassen sich demnach im Wesentlichen vier Dimensionen der Begriffsverwendung abgrenzen. Sie werden im Folgenden zusammen mit den entsprechenden Literaturangaben ihrer Verwendung dargelegt (Schieder 2014):

1. Die BI im **institutionellen oder konfigurativen Sinn** entspricht einer Personengruppe innerhalb einer Organisation (häufig eine dedizierte Teileinheit eines Unternehmens), die mit der Sicherstellung der Realisierung von BI-Prozessen betraut ist (Dresner et al. 2002). In der Literatur werden konkrete Realisierungsformen dieses Bedeutungsaspektes zumeist mit den Konstrukten „BI Competency Center“ oder „BI-Team“ thematisiert (Totok 2010). Dieses Verständnis spiegelt sich in den frühen Arbeiten zur BI wieder, wo organisationale Architekturbetrachtungen im Vordergrund standen, die euch heute wieder stärker in den Fokus rücken.
2. Im **funktionalen oder prozessualen Sinn** verstanden, bezieht sich der Begriff auf die für die Generierung geschäftsrelevanter Informationen, Erkenntnisse und Wissen erforderlichen Tätigkeiten und Aufgaben („*Business Intelligence als analytischer Prozess [...]*“ Grothe und Gentsch 2000). Hier steht die Beherrschung eines Informationsversorgungsprozesses im Mittelpunkt der Betrachtung (Gabriel und Dittmar 2001). Dessen Aufgabe besteht darin, fragmentierte Unternehmens- und Wettbewerbsdaten in handlungsgerichtetes Wissen zu überführen (Golfarelli et al. 2004). Die Sicherstellung der zielgerichteten BI-Prozessausführung firmiert unter der Bezeichnung BI-Governance (Horakh et al. 2008). Foley und Guillemette (2010) schließen sich diesem Verständnis an, wenn sie BI definieren als „*a combination of processes, politics, culture, and technologies for gathering, manipulating, storing, and analyzing the data collected from internal and external sources in order to communicate information, create knowledge and inform decision making.*“
3. BI in **resultativem oder ergebnisorientiertem Sinn** gebraucht, bezeichnet das Ergebnis eines Erkenntnisprozesses. Dieses Ergebnis besteht in zweck- und zielorientiertem Wissen in Form von Berichten, Prognosen, Analysen, Mitteilungen etc., das Entscheidungsträgern in Unternehmen zu einem besseren Verständnis des eigenen Wertschöpfungsgefüges (Becker et al. 2010) und dessen Situierung im Marktgeschehen verhilft (Turban et al. 2010). Dieses Verständnis geht häufig mit einer prozessorientierten Sichtweise einher. So ergänzen (Foley und Guillemette 2010) ihre oben angeführte BI-Definition um folgende Aussage: „*BI helps report business performance, uncover new business opportunities and make better business decisions about competitors, suppliers, customers, financial issues, strategic issues, products, and services.*“ Dieses Verständnis kommt dem Bedeutungsfeld der Begriffswurzel aus etymologischer Perspektive am nächsten.
4. Der **technisch-methodische Sinn** des BI-Begriffs schlägt sich nieder in der Funktion des Terminus als „*begriffliche Klammer, die eine Vielzahl unterschiedlicher Ansätze zur Analyse geschäftsrelevanter Daten zu bündeln versucht*“ (Gluchowski 2001). Insbesondere umfasst diese Klammer eine Sammlung von (informationstechnischen)

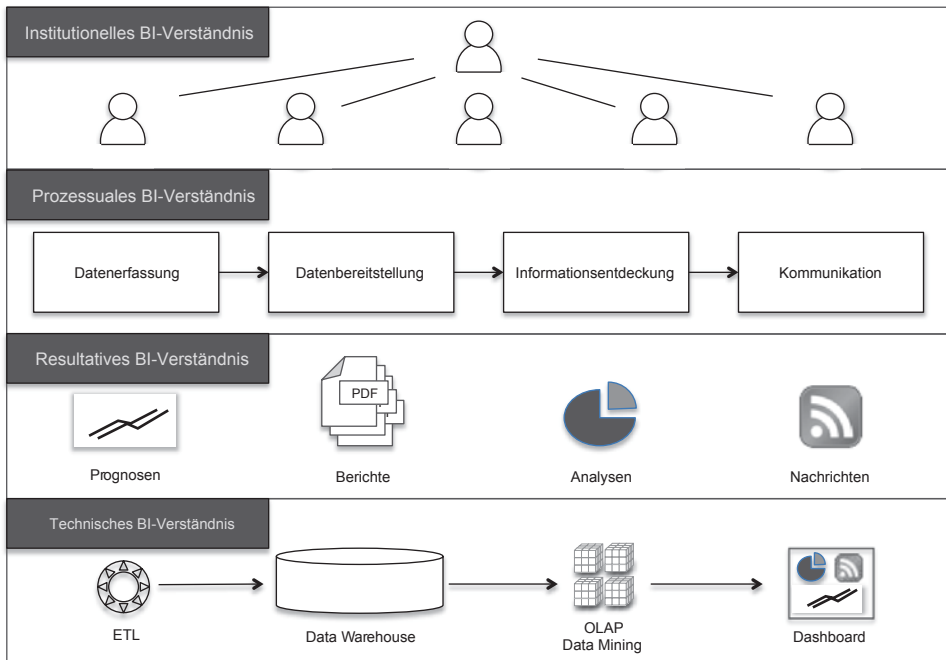


Abb. 2.3 Ebenen eines ganzheitlichen BI-Verständnisses. (Schieder 2014)

Werkzeugen, Architekturen, Systemen, Technologien und Konzepten zur Integration, Aufbereitung und Bereitstellung fragmentierter, geschäftsrelevanter interner und externer, strukturierter und unstrukturierter Daten zum Zweck der Informationsgewinnung (Turban et al. 2010).

Eine zusammenfassende Darstellung dieser vier Verständnisebenen der BI liefert die nachfolgende Abb. 2.3. Die Darstellung deutet zudem eine gewisse Hierarchie der Ordnungsprioritäten an, in der die institutionellen Aspekte vor Problemen der prozessualen Kategorie stehen. Gleichwohl steht außer Frage, dass moderne BI-Konzeptionen maßgeblich von ihrer technologischen Basis überhaupt erst ermöglicht werden.

Die Realisierung einer erfolgreichen BI (institutionell) und dem zielgerichteten Erzeugen von BI (resultativ) im Unternehmen ist Absicht des BI-Engineering. Mit dessen Hilfe etablieren die mit der Umsetzung der BI betrauten BI-Institutionen beziehungsweise Organisationseinheiten (häufig BI-Competency-Center, kurz BICC) abteilungs- und unternehmensweite BI-Prozesse und stellen die Konformität der Prozessergebnisse mit den Unternehmenszielen durch BI-Governance sicher. Die methodisch-technische Basis hierfür bilden BI-Systeme. Die BI-Architektur eines Unternehmens setzt sich in der Regel aus mehreren BI-Systemen beziehungsweise deren (informationstechnischen) BI-Komponenten zusammen, die unterschiedliche Aufgaben und Funktionen der Datensammlung, -integration, -bereitstellung, -speicherung und -auswertung sowie der Präsentation und Kommunikation der Analyseergebnisse unterstützen.

Allgemein gesprochen hat die BI das Ziel, Entscheider beim Erkenntnisprozess zu unterstützen und ihnen die hierfür erforderlichen Informationen zur richtigen Zeit in der gewünschten Qualität (Format, Darstellung, Aktualität, Güte, etc.) an den Ort zu liefern, wo sie benötigt werden. Aus kognitionswissenschaftlicher Sicht fokussiert die Unterstützung auf die Vorbereitung und Durchführung von situationsadäquaten Auswahlentscheidungen, wofür ein korrektes, mentales Situationsabbild beim Entscheider eine notwendige Bedingung darstellt. Der folgende Abschnitt befasst sich daher mit dem kognitionswissenschaftlichen Konstrukt der Situation Awareness, dass als Zielkonstrukt für die Gestaltung einer kognitions- und entscheidungsorientierten BI in Frage kommt.

2.4 Situation Awareness als Zielkonstrukt einer kognitionsorientierten Gestaltung der Business Intelligence

Mit der Zunahme des Zeitaspekts für das Treffen von Entscheidungen (Stichwort „Real-time Enterprise“) und einem beständig zunehmendem Informationsangebot (Stichwort „Big Data“) verschärft sich das Problem der situationsadäquaten Auswahl zwischen Entscheidungsalternativen. Insbesondere Handelnde in operativen Unternehmensprozessen sind herausgefordert, in kurzer Zeit und bei hoher Frequenz geschäftsrelevante Entscheidungen zu treffen. In der Vergangenheit haben dabei Konzepte aus militärischen Anwendungsbereichen (Operations Research, OODA-Loop) gewinnbringende Impulse für die unternehmenswichtige Anwendungssysteme geliefert. Aus einem ebensolchen Kontext stammt das Konstrukt der Situation Awareness (SA). Der nachfolgende Abschnitt führt daher in die Grundlagen des Konstrukts ein, mit dem Ziel die Eignung von SA als erwünschte Zielkategorie von ganzheitlichen BI-Systemen zu demonstrieren.

Einer populären Definition von (Endsley 1995) folgend, wird SA wie folgt verstanden: „the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future“. Extraktion, Verdichtung und Bereitstellung von Informationen über gewisse „elements in the environment“ (prozessbezogene Informationsobjekte, externe Wissensbausteine) beschreibt eine Kernaufgabe der BI. Primäres Ziel der IT-Unterstützung ist dabei, Entscheidern die Wahrnehmung (*perception*) relevanter Informationsbausteine zu ermöglichen. Weiterhin sollen Analysefunktionen Anwender in die Lage versetzen, vertieftes Verständnis (*comprehension*) in die Zusammenhänge und Abhängigkeiten der durch die Informationsobjekte repräsentierten Umweltzustände zu gewinnen. Der kombinierte Einsatz fortgeschrittener Methoden der BI (z. B. Data Mining) legt schließlich die Basis zur Projektion (*projection*) zukünftiger Systemzustände. SA hat sich als valide Basis für das Verständnis des Entscheidungshandelns operativer Akteure und eignet sich damit prinzipiell für die Gestaltung von Entscheidungsunterstützungssystemen in Risikobranchen. Da die Tendenz zur Operationalisierung auch im BI-Umfeld zu beobachten ist, existieren erste Ansätze zur Übertragung der SA auf die Domäne der BI (Schieder 2011).

Entsprechend obiger Definition, beschreibt das Konstrukt der SA den Zustand eines Akteurs bezüglich dreier, aufeinander aufbauender Zustände (Endsley 1995):

- seiner Wahrnehmung (Perception) der Elemente seiner Handlungsumwelt, in einem bestimmten zeitlichen und räumlichen Umfang, auch als Level 1 SA bezeichnet;
- seines Verständnisses (Comprehension) der Bedeutung der Elemente, ihrer Zustände und Interdependenzen für die eigene Situation (Level 2 SA) und
- seiner Prognose (Projection) des Zustandes der Umweltelemente in der nahen Zukunft (Level 3 SA).

SA bettet den (kognitiven) Zustand eines Akteurs in den Kontext der Auswirkungen seines Entscheidungshandelns auf den Zustand der Umwelt ein. Umweltveränderungen machen wiederum eine Anpassung der SA erforderlich. Die Fähigkeit eines Akteurs zur Adaption seiner SA wird nach diesem Modell durch aufgaben- bzw. systemimmanente (auch als „extrapersonale“ bezeichnete) Faktoren einerseits und intrapersonale Faktoren andererseits determiniert. Zu ersteren gehören alle Einflussfaktoren des Arbeitssystems (task/system factors): die Eignung des unterstützenden IT-Systems, insbesondere die Gestaltung der Mensch-Maschine-Schnittstelle, die Arbeitsbelastung, die Komplexität der zu bewältigenden Aufgabe sowie der Grad der Automatisierung des Arbeitsprozesses. Zu den personen-bezogenen Faktoren (individual factors) zählen unter anderen individuelle Ziele und Absichten sowie Erwartungen, die sich in Form von mentalen Modellen aus der Internalisierung von Erfahrungswissen ergeben. Daneben beeinflussen Erfahrungswissen, die individuelle Fähigkeit Informationen zu verarbeiten und Handlungsautomatismen die SA eines Akteurs. Die intrapersonalen Faktoren werden in entscheidendem Maße von den persönlichen Fertigkeiten, der Erfahrung und dem Ausbildungsstand des Handelnden geprägt. Alle diese Faktoren beeinflussen Entscheidungen und Handeln eines Akteurs sowohl in indirekter Weise als auch indirekt über den Zustand seiner SA. Abbildung 2.4 stellt das beschriebene Modell zum Verständnis des Phänomens der SA zusammenfassend dar.

Vor allem Handelnde in operativen Geschäftsprozessen sind in hohem Maße auf eine aktuelle, intuitiv erfassbare Situationsbeschreibung zur Auswahl möglicher Handlungsalternativen angewiesen (Artman 2000). Gleiches gilt allerdings auch mehr und mehr für strategische Entscheidungssituationen. Dabei muss vor dem beschriebenen Hintergrund sichergestellt sein, dass die in der jeweiligen Situation relevanten Informationen in der Flut der auf die Akteure einströmenden Signale wahrgenommen werden können (Wickens 2008). Dabei hat sich SA insbesondere in Forschungsarbeiten und Experimenten zu Entscheidungsprozessen in hochkompetitiven Entscheidungssituationen als dominanter Erfolgsfaktor erwiesen (Yang et al. 2009).

Die beschriebenen Teilzustände der SA können allesamt als Ergebnisartefakte des Einsatzes von BI-Werkzeugen aufgefasst werden. Aus Sicht der Kognitionswissenschaft unterstützt BI daher Entscheider in ihrer Situationswahrnehmung und -verarbeitung beziehungsweise beim Aufbau ihrer SA. Damit ist zwar nicht garantiert, dass Entscheider situationsadäquate Entscheidungen treffen. Allerdings ist eine notwendige Bedingung erfüllt, ohne die optimale Entscheidungen (vor allem, aber nicht nur unter dem Aspekt kurzer Latenzzeiten) nicht möglich sind (Wickens 2008). Aus diesem Verständnis heraus wäre es demnach Aufgabe der BI, Akteuren bei der Erlangung dieser Zustände bestmöglich

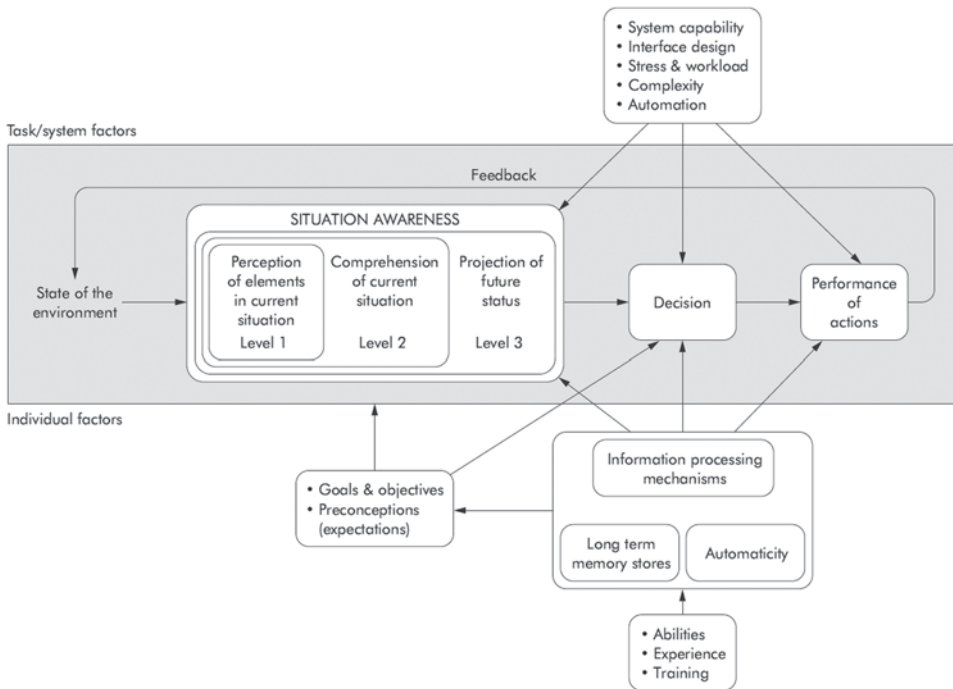


Abb. 2.4 Modell zur Situation Awareness in dynamischen Entscheidungssituationen von. (Endsley 1995)

behilflich zu sein. Dieser Argumentation folgend bietet sich das Konstrukt der SA als Ziel- oder Leitkonstrukt für die Gestaltung der BI an. BI-Konzeptionen könnten angelehnt an das Verständnis der SA auf ihre Tauglichkeit aus Perspektive der Kognitionswissenschaft untersucht und ausgerichtet werden.

Dass die Rolle einer zukünftigen BI möglicherweise noch darüber hinaus reicht, zeigt die Zusammenfassung der bisherigen Entwicklungen und die Darstellung aktueller Trends im Folgenden und letzten Abschnitt des Beitrags.

2.5 Zusammenfassung und Ausblick

Ziel des Beitrags war es, historische und taxonomische Fragmente der BI chronologisch und systematisch aufzubereiten und darzustellen. Dabei sollten insbesondere bisher unbekannte Aspekte der Entwicklungsgeschichte des Bedeutungsfelds der BI aufgezeigt werden, um die Grundlage für ein umfassendes und ausdifferenziertes Verständnis der BI zu legen. Der Einblick in die Geschichte der Begriffsverwendung hilft dieses Verständnis zu schärfen, erlaubt eine Standortbestimmung aktueller Entwicklungen und legt damit den Grundstein für die Einordnung zukünftiger Innovationen im Umfeld. Dieser letzte

Abschnitt soll die bisherigen Ausführungen zusammenfassen und auf Basis aktueller Entwicklungen einen Ausblick in die Zukunft der BI wagen.

Ausgehend von einer Analyse aktueller Beobachtungen zur BI formulieren (Baars et al. 2014) ein Leitbild der BI von morgen. Darin sehen sie die Rolle der BI als zentrales Vehikel in „*agil koordinierten Netzwerken aus flexibel kombinierbaren Informationsintegrations- und Analyse-Capabilities*“ und identifizieren fünf hauptsächliche Innovationsstränge für die Weiterentwicklung der BI:

1. Tiefgreifende Integration von BI und Geschäftsprozessmanagement
2. BI über Organisationsgrenzen sowie kooperativ entwickelte und betriebene BI-Lösungen
3. Neue Ansätze zur Verarbeitung semi- oder unstrukturierter Daten, insbesondere aus unternehmensexternen Quellen
4. Agilität und Bereitstellung von BI-Lösungen und benutzergesteuerte BI-Entwicklungsmodelle
5. Neue Konzepte für eine BI-Governance zur strategiegesteuerten Lenkung und Kontrolle der BI

Kam der BI historisch im Vergleich zu den Enterprise Systems häufig eine eher ergänzende Rolle zu, zum Teil als analytische Insellösungen, rücken zukünftige BI-Systeme ins operative Herz von Unternehmen („BI und Geschäftsprozessmanagement“). Anders ausgedrückt: hat die traditionelle BI bisher vor allem auf das „OO“ im „OODA“-Loop fokussiert (Minelli et al. 2013) ist davon auszugehen, dass die Entwicklung der BI zunehmend in Richtung einer „Decision Intelligence“ zur Entscheidungsautomatisierung läuft und die Situation Awareness als Zielkonstrukt in den Fokus rückt (Sharma et al. 2014). Der Schwerpunkt von BI-Services wandelt sich von reaktiv zu prädiktiv und selbst eine antizipative Ausprägung wird denkbar: BI-Services nehmen zukünftige Situationsentwicklungen vorweg und bieten Entscheidungshilfen für noch nicht eingetretene Umweltzustände an. Gleichzeitig sprengt BI bisher bestehende Organisationsgrenzen und schafft die analytischen Rahmenbedingungen für die unternehmensübergreifende Koordination von Wertschöpfungsketten („BI über Organisationsgrenzen“). Dabei werden neue Ansätze zu Verarbeitung und Auswertung polystrukturierter Daten aus heterogenen Datenquellen ein zentrales Element bilden. Zur Beschleunigung der Anpassungsgeschwindigkeit der BI werden zudem agilere Methoden für eine gleichermaßen benutzergesteuerte wie auch strategiekonforme Entwicklung, Lenkung und Kontrolle der BI gesucht.

Das diese Trends nicht nur reines Wunschdenken der BI-Community sind, zeigt die zunehmende Relevanz der BI in Wissenschaft und Praxis. Abbildung 2.5 vergleicht die Entwicklung der Häufigkeit von Suchanfragen zu BI-bezogenen Suchbegriffen auf Google im Zeitverlauf der letzten zehn Jahre. Ins Auge fällt dabei die rapide Abnahme von Suchanfragen zu „Data Mining“ und „OLAP“, die in Anfang der 2000er Jahre den Höhepunkt ihres „Hype Cycles“ durchschritten haben. Ersichtlich ist der stabile, wenngleich leicht rückläufige, Verlauf der Häufigkeit von Suchanfragen unter dem Term „Business

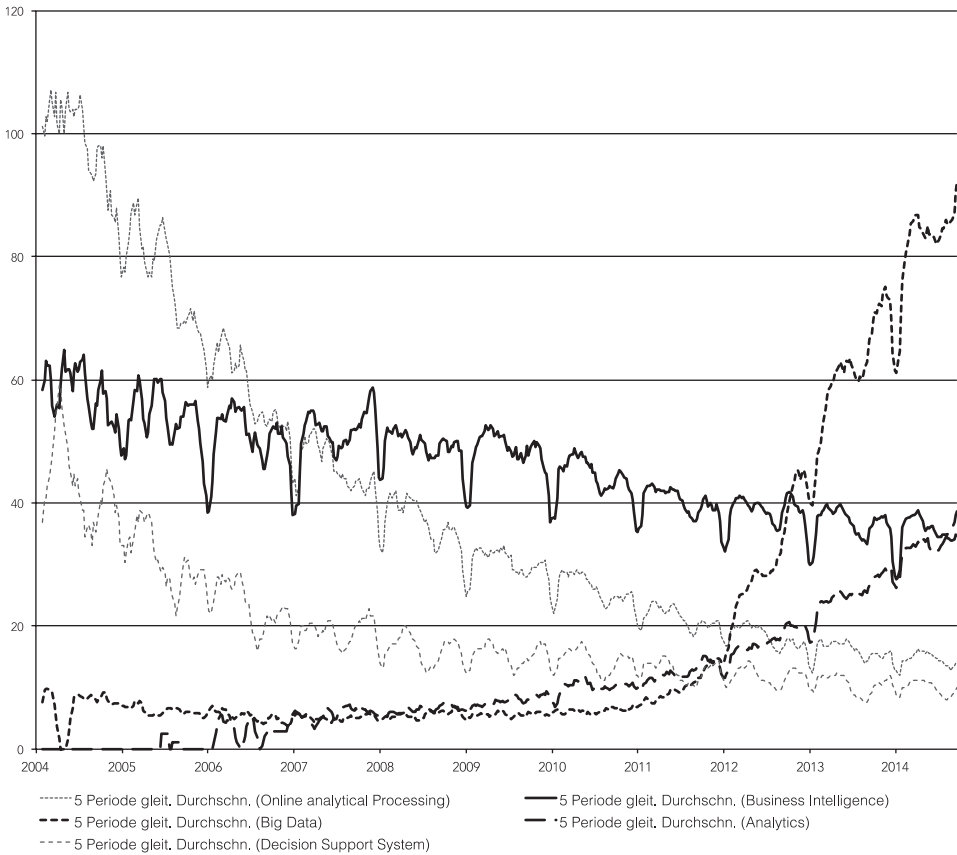


Abb. 2.5 Vergleich der Häufigkeiten von BI-bezogenen Suchbegriffen auf Google im Zeitverlauf. (Quelle: Google Trends, geglätteter Durchschnitt über 14 Wochen)

Intelligence“. In den vergangenen beiden Jahren vervielfachten sich dagegen die Anfragen zum Begriff „Big Data“. Diese Entwicklung hat augenscheinlich ihren Höhepunkt noch nicht erreicht. Suchanfragen zu „Business Analytics“ beziehungsweise „Data Analytics“ weisen in der Addition mittlerweile den selben Stellenwert auf, wie diejenigen zur „Business Intelligence“. Die Stabilität des BI-Trends gibt Anlass zu der Vermutung, dass BI auch weiterhin das führende und etablierte Konstrukt zur Bezeichnung analytischer Informationssysteme bleiben wird. Inwieweit es sich beim Big-Data-Begriff um mehr als eine Mode handelt, bleibt allerdings abzuwarten. Derzeit sind in der Praxis jedenfalls weiterhin BI-Kompetenz weitaus stärker nachgefragt als diejenigen, die in den Bereich von sogenannten Data Scientists fallen würden (Debortoli et al. 2014).

Der IT-Klimawandel im Zuge der Datafication (Lycett 2013) jedenfalls geht weiter. Der Daten-Meeresspiegel steigt und das geflügelte Wort vom „Drowning in Data – starving for information“ bleibt solange gültig wie Moore’s Law gilt. Aus diesem Grund können auch (Hosack et al. 2012) ohne weiteres einen Blick in die Zukunft der DSS-Forschung wagen

und bescheinigen ihr eine ungebrochene Vitalität, abseits aller Terminologie-Debatten und Moden. Dasselbe darf ohne Zweifel auch für die BI angenommen werden. Was vor allem ersichtlich ist, ist die Tatsache, dass sich die äußere (technologische) Gestalt der BI kontinuierlich verändert („*BI is constantly changing*“ Watson 2009), während die grundlegenden Prinzipien und Notwendigkeiten ihres Einsatzes stabil bleiben. Insbesondere wird auch unter dem Hype um Big Data die Notwendigkeit Entscheider beim Erkenntnisprozess zu unterstützen und ihnen die hierfür erforderlichen Informationen zur richtigen Zeit in der gewünschten Qualität (Format, Darstellung, Aktualität, Güte, etc.) an den richtigen Ort zu liefern, nicht ab- sondern zunehmen.

Literatur

- Adam, F.: 20 years of decision making and decision support research. *J. Decis. Syst.* **21**(2), 93–99 (2012)
- Artman, H.: Team situation assessment and information distribution. *Ergonomics* **43**(8), 1111–1128 (2000)
- Baars, H., Felden, C., Gluchowski, P., Hilbert, A., Kemper, H.-G., Olbrich, S.: Gestaltung der nächsten Inkarnation von Business Intelligence. *Wirtschaftsinformatik* **56**(1), 13–19 (2014)
- Becker, W., Kollacks, K., Ulrich, P.: ZP-Stichwort: Business Intelligence und Business Intelligence-Tools. *Z. Planung Unternehmenssteuerung* **21**(2), 223–232 (2010)
- Boyd, D., Crawford, K.: Critical questions for big data. *Inf. Commun. Soc.* **15**(5), 662–679 (2012)
- Böning, H.: Pressewesen und Aufklärung – Intelligenzblätter und Volksaufklärer. *Goethezeitportal* (2004)
- Chamoni, P., Gluchowski, P.: Analytische Informationssysteme – Business Intelligence-Technologien und -Anwendungen, 4. Aufl. Springer-Verlag Berlin (2010)
- Chaudhuri, S., Dayal, U.: An overview of data warehousing and OLAP technology. *ACM SIGMOD Rec.* **26**(1), 65–74 (1997)
- Codd, E.F., Codd, S.B., Salley, C.T.: Providing OLAP to user-analysts: An IT Mandate. Whitepaper (1993)
- Curts, R., Campbell, D.: Avoiding information overload through the understanding of OODA loops. *Proceedings of the Command & Control Research & Technology Symposium* (December 1986), S. 1–18 (2001)
- Debortoli, S., Müller, O., Brocke, J.V.: Vergleich von Kompetenzanforderungen an Business-Intelligence- und Big-Data-Spezialisten. *Wirtschaftsinformatik* **56**(5), 315–328 (2014)
- Dinter, B., Lahrman, G., Winter, R.: Information logistics as a conceptual foundation for enterprise-wide decision support. *J. Decis. Syst.* **19**(2), 175–200 (2010)
- Dittmar, C., Ossendoth, V.: Die organisatorische Dimension von Business Intelligence – Vorgehen und Erfahrungen bei der Gestaltung von Business Intelligence Organisationen. In: Chamoni, P., Gluchowski, P. (Hrsg.) *Analytische Informationssysteme*, 4. Aufl., S. 59–86. Springer-Verlag Berlin (2010)
- Dresner, H., Linden, A., Buytendijk, F., Strange, K., Knox, M., Camm, M.: The Business Intelligence Competency Center: An Essential Business Strategy. *Gart. Strateg. Anal. Rep.* Stamford (2002)
- Duden online: Intelligenz, die. Duden. <http://www.duden.de/zitieren/10020586/1.9> (2012). Zugriffen: 27. Jul. 2015
- El Sawy, O.A., Majchrzak, A.: Critical issues in research on real-time knowledge management in enterprises. *J. Knowl. Manag.* **8**(4), 21–37 (2004)

- Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **37**(1), 32–64 (1995)
- Foley, É., Guillemette, M.G.: What is business intelligence? *Int. J. Bus. Intell. Res.* **1**(4), 1–28 (2010)
- Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R.: From data mining to knowledge discovery in databases. *AI Mag. Am. Assoc. Artif. Intell.* **17**(3), 37–54 (1996)
- Gabriel, R., Dittmar, C.: Der Ansatz des Knowledge Managements im Rahmen des Business Intelligence. *HMD – Praxis Wirtschaftsinform.* **222**, 17–28 (2001)
- Ghoshal, S., Kim, S.K.: Building effective intelligence systems for competitive advantage. *Sloan Manag. Rev.* **28**(1), 49–58 (1986)
- Gilad, T., Gilad, B.: Business intelligence – the quiet revolution. *Sloan Manag. Rev.* **27**(4), 53–61 (1986)
- Gluchowski, P.: Business Intelligence – Konzepte, Technologien und Einsatzbereiche. *HMD – Praxis Wirtschaftsinform.* **222**, 5–15 (2001)
- Gluchowski, P., Kemper, H.-G.: Quo Vadis Business Intelligence? Aktuelle Konzepte und Entwicklungstrends. *BI-Spektrum* **2**(1), 12–19 (2006)
- Gluchowski, P., Gabriel, R., Dittmar, C.: Management Support Systeme und Business Intelligence, 2. Aufl. Springer, Berlin (2008)
- Gluchowski, P., Schieder, C., Böhringer, M.: Web 2.0-inspirierte Business-Intelligence-Lösungen für die Anwender der Zukunft. *HMD – Praxis Wirtschaftsinform.* **289**, 16–25 (2011)
- Golfarelli, M., Rizzi, S., Cella, I.: Beyond data warehousing: What's next in business intelligence? In: Proceedings of the 7th ACM international workshop on Data warehousing and OLAP, S. 1–6 (2004)
- Grothe, M., Gentsch, P.: Business Intelligence – Aus Informationen Wettbewerbsvorteile gewinnen. Addison-Wesley, München (2000)
- Heinrich, L.J.: Forschungsmethodik einer Integrationsdisziplin: Ein Beitrag zur Geschichte der Wirtschaftsinformatik. *NTM Int. J. Hist. Ethics Nat. Sci. Technol. Med.* **13**(2), 104–117 (2005)
- Horakh, T.A., Baars, H., Kemper, H.-G.: Mastering business intelligence complexity – a service-based approach as a prerequisite for BI governance. *AMCIS 2008 Proceedings*, S. 333 (2008)
- Hosack, B., Hall, D., Paradise, D., Courtney, J.F.: A look toward the future: decision support systems research is alive and well. *J. Assoc. Inf. Syst.* **13**(5), 315–340 (2012)
- Houghton, R., El Sawy, O.A., Gray, P., Donegan, C., Joshi, A.: Vigilant information systems for managing enterprises in dynamic supply chains: real-time dashboards at Western Digital. *MIS Q. Exec.* **3**(1), 19–35 (2004)
- Kozak-Holland, M.: Churchill's Adaptive Enterprise: lessons for business today. In: *Reihe: Lessons from History Multi-Media Publications*, Oshawa (2005)
- Lange, V.: Technologische Konkurrenzanalyse: zur Früherkennung von Wettbewerberinnovationen bei deutschen Großunternehmen. Springer, Wiesbaden (1994)
- Luftman, J., Ben-Zvi, T.: Key issues for IT executives 2010: Judicious IT investments continue post-recession. *MIS Q. Exec.* **9**(1), 49–59 (2011)
- Luhn, H.P.: A business intelligence system. *IBM J. Res. Dev.* **2**(4), 314–319 (1958)
- Lycett, M.: „Datafication“: making sense of (big) data in a complex world. *Eur. J. Inf. Syst.* **22**(4), 381–386 (2013)
- Meier, M.C., Mertens, P.: Competitive intelligence. *Wirtschaftsinformatik* **46**(5), 405–407 (2004)
- Minelli, M., Chambers, M., Dhiraj, A.: Big data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses. Wiley, Hoboken (2013)
- Online Etymology Dictionary: Intelligence (n.). <http://www.etymonline.com/index.php?term=intelligence> (2012). Zugriffen: 27. Jul. 2015
- Oxford Online Dictionary: Intelligence. <http://oxforddictionaries.com/definition/intelligence?q=intelligence>. Oxford University Press (2012). Zugriffen: 27. Jul. 2015
- Power, D.J.: A brief history of decision support systems. *DSSResources.com* (2003). Zugriffen: 27. Jul. 2015

- Power, D.J., Burstein, F., Sharda, R.: Reflections on the past and future of decision support systems: perspective of eleven pioneers. In: Schuff, D., Paradice, D., Burstein, F., Power, D.J., Sharda, R. (Hrsg.) *Decision support, annals of information systems* 14, S. 25–49. Springer-Verlag New York (2011)
- Schieder, C.: Kognitionsorientierte Gestaltung operativer Business Intelligence-Lösungen. In: Bayreuther Arbeitspapiere zur Wirtschaftsinformatik Tagungsband zum Doctoral Consortium der WI. Bayreuth, S. 153–162 (2011)
- Schieder, C.: Methodenentwurf zur Gestaltung der Pricing Intelligence – Ein Methoden-Engineering-Ansatz zur Verbesserung der IT-Unterstützung im Preismanagement, Hamburg (2014)
- Schieder, C., Dinter, B., Gluchowski, P.: Metadatenmanagement in der Business Intelligence – eine empirische Untersuchung unter Berücksichtigung der Stakeholder-Perspektiven. In: 12th international conference on Wirtschaftsinformatik (2015)
- Scott, M.M.S.: Computer-driven visual display devices – their impact on the management decision-making process (1967)
- Sharma, R., Mithas, S., Kankanhalli, A.: Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations. *Eur. J. Inform. Syst.* **23**(4), 433–441 (2014)
- Simon, H.A.: Perspektiven der Automation für Entscheider. Quickborner Team, Hamburg (1966)
- Steininger, K., Riedl, R., Roithmayr, F., Mertens, P.: Moden und Trends in Wirtschaftsinformatik und Information Systems – Eine vergleichende Literaturanalyse. *Wirtschaftsinformatik* **51**(6), 478–495 (2009)
- Totok, A.: Von der Strategie zum Business Intelligence Competency Center. Hanser Fachbuchverlag, München (2010)
- Turban, E., Sharda, R., Delen, D.: *Decision Support and Business Intelligence Systems*, 9. Aufl. Prentice Hall, New Jersey (2010)
- Wand, Y., Wang, R.Y.: Anchoring data quality dimensions in ontological foundations. *Commun. ACM* **39**(11), 86–95 (1996)
- Wang, R.Y., Strong, D.M.: Beyond accuracy: What data quality means to data consumers. *J. MIS* **12**(4), 5–33 (1996)
- Watson, H.J.: Tutorial : Business intelligence – past, present, and future. *Commun. AIS* **25**(1), 487–510 (2009)
- Weber, J., Schäffer, U.: *Einführung in das Controlling*, 11. Aufl. Schäffer-Poeschel, Stuttgart (2006)
- Wickens, C.D.D.: Situation awareness: review of Mica Endsley’s 1995 articles on situation awareness theory and measurement. *Hum Factors* **50**(3), 397–403 (2008)
- Winter, R.: Analytische Informationssysteme aus Managementsicht. In: Chamoni, P., Gluchowski, P. (Hrsg.) *Analytische Informationssysteme*, 4. Aufl., S. 87–114. Springer, Berlin (2010)
- Yang, S., Stotz, A., Holsopple, J., Sudit, M., Kuhl, M.: High level information fusion for tracking and projection of multistage cyber attacks. *Inf. Fusion* **10**(1), 107–121 (2009)

Von der Business-Intelligence-Strategie zum Business Intelligence Competency Center

3

Andreas Totok

Inhaltsverzeichnis

3.1	Einführung	34
3.2	Entwicklung einer BI-Strategie.	36
3.2.1	Business-Intelligence-Strategie.	36
3.2.2	Vorgehensmodell zur Entwicklung einer BI-Strategie	36
3.3	Aufbau eines BI Competency Center	45
3.4	Ausgewählte Aspekte einer BI-Strategie	46
3.4.1	Kosten und Nutzen bewerten	46
3.4.2	Auf Trends reagieren.	48
3.4.3	Architektur gestalten	48
3.4.4	Softwareportfolio festlegen.	49
3.4.5	Spannungsfeld erkennen.	50
3.5	Fazit	52
	Literatur	53

Zusammenfassung

Um das Thema Business Intelligence (BI) professionell zu steuern, bedarf es einer klaren Strategie sowie einer übergreifenden organisatorischen Verankerung. Zwar haben Unternehmen seit langer Zeit umfangreiche Erfahrung im Einsatz von Analytischen Informationssystemen, dennoch werden weiterhin Probleme bzgl. Standardisierung, Datenqualität oder Komplexität genannt. Hinzu gekommen sind in den letzten Jahren

A. Totok (✉)
Enterprise Information Management, Finanz Informatik Solutions Plus GmbH,
Frankfurt am Main, Deutschland
E-Mail: andreas.totok@f-i-solutions-plus.de

Trends wie Big Data, Mobile BI oder Self Service BI. Um den besonderen Anforderungen zur erfolgreichen Umsetzung von Analytischen Informationssystemen gerecht zu werden, sollten Unternehmen daher eine durchgängige BI-Strategie verfolgen. Berücksichtigt werden müssen in diesem Kontext die jeweilige Unternehmenssituation, wie Art des Geschäftsmodells, Unternehmensorganisation, Geschäftsprozesse, vorhandene IT-Infrastruktur und natürlich Kosten-/Nutzenaspekte. Gesteuert werden sollte die Umsetzung der Strategie durch eine übergreifende Organisationseinheit. Als Best Practice hierzu hat sich das Business Intelligence Competency Center (BICC) herausgebildet, das Verantwortung, Rollen und Prozesse für die Anwendung der AIS bündelt.

3.1 Einführung

Dieser Beitrag erscheint als Bestandteil dieses Sammelwerks zum dritten Mal und wurde erneut an die aktuelle Situation angepasst. Wurde in der dritten Auflage 2006 die Notwendigkeit einer BI-Strategie noch begründet, so erschien sie mit der vierten Auflage 2010 als fest gesetzt. Zwar gilt diese Aussage auch 2015 weiterhin, doch scheint die BI-Strategie aktuell zugunsten einer „Big Data“-Strategie in manchen Unternehmen in den Hintergrund zu treten. Dies geht einher mit der provokanten Frage, ob der klassische Data-Warehouse-Ansatz nicht sogar überholt sei. An markige Vertriebskampagnen von IT-Dienstleistern, Soft- und Hardwareanbietern ist man ja seit Jahren gewöhnt, dennoch hat die aktuelle Diskussion manchen BI-Spezialisten nachdenklich gemacht zu überlegen, ob in der Vergangenheit nicht vielleicht gut gemeinte Ratschläge in die falsche Richtung gelaufen sind. So ist beispielsweise die Diskussion um Reifegradmodelle in den letzten Jahren deutlich abgeebbt und das Ziel der Erreichung einer höchsten Reifegradstufe für AIS vielen Unternehmen ins Stocken gekommen und ggf. sogar verworfen worden. An die Stelle einer Einordnung der eigenen Situation in einem Reifegradmodell mit einem allgemeinen Anspruch ist eher ein Vergleich zu branchenbezogenen oder übergreifenden Best Practices getreten.

Man kann feststellen, dass im Themenbereich Business Intelligence in den letzten fünf Jahren zahlreiche neue, wichtige Trends entstanden. „Gefühl“ ist dabei die Anzahl der Trends in der langen Geschichte der AIS so hoch wie kaum jemals zuvor. Zu den Trends zählen unter anderem:

- Big Data: Integration und Analyse von großen Mengen strukturierter und semi-strukturierter Daten auf Basis von Datei-basierten Frameworks (prominentester Vertreter ist Hadoop).
- Neue Speicherkonzepte für analytische Daten: hauptspeicher- und spaltenbasierte Speicheransätze, die zu außergewöhnlich kurzen Antwortzeiten führen.
- Cloud BI: Bezug von BI-Leistungen bzw. BI-Funktionalität aus der Cloud. Dabei müssen insbesondere Aspekte wie Vertraulichkeit und Sicherheit berücksichtigt werden.

- Mobile BI: Analyse- und Reportingmöglichkeiten auf mobilen Geräten unter Nutzung entsprechender Bedienkonzepte wie Finger- bzw. Gestensteuerung.
- Self Service BI: einfach zu nutzende Funktionen, die Fachbereichsbenutzern mehr Eigenständigkeit in BI-Anwendungen ermöglichen, Beispiel: Anpassungen im Datenmodell durch den Power User.
- Agile BI: Wechsel vom klassischen sequentiellen Vorgehen hin zum iterativen Projektvorgehen in kurzen, mehrwöchigen Entwicklungszyklen.

Trotz dieser Fülle von Trends und der zunehmenden Erfahrung im Umgang mit BI, stellen sich viele Unternehmen die Frage, ob sie eine BI-Strategie benötigen. Man hat doch schon eine Unternehmens- und eine IT-Strategie, wozu also noch eine zusätzliche BI-Strategie? Die Frage, ob ein Unternehmen eine BI-Strategie benötigt, lässt sich nicht pauschal beantworten. Gut aufeinander abgestimmte Unternehmens- und IT-Strategien könnten theoretisch ausreichen, um die reibungslose Anwendung von BI zu gewährleisten. Dennoch gibt es in der Praxis eine Reihe von immer wiederkehrenden praktischen Problemstellungen, die ein starkes Indiz für die Notwendigkeit einer BI-Strategie sind. Die im Folgenden aufgeführten Punkte können einzeln oder in Kombination ein Zeichen hierfür sein:

- Das Management wird mit einer Reihe unterschiedlicher Berichte aus unterschiedlichen Quellen versorgt. Ähnliche oder gleichlautende Kennzahlen haben unterschiedliche Inhalte und weichen voneinander ab.
- Die Datenqualität der BI-Systeme wird von den Anwendern als unzureichend bewertet.
- Es gibt mehrere unterschiedliche – ggf. konkurrierende – Informationssysteme im Unternehmen, die den Anspruch erheben, den Status eines Data Warehouse zu besitzen.
- Es existieren aus historischen Gründen eine Reihe unterschiedlicher Frontend-Werkzeuge für Analyse und Reporting, die ähnliche Funktionen besitzen und von verschiedenen verantwortlichen Stellen gepflegt werden müssen.
- Das Know-how für die Benutzung und Weiterentwicklung der vorhandenen BI-Systeme ist nur auf wenige Personen beschränkt, die in Spitzenzeiten regelmäßig einen Kapazitäts-Engpass bilden.
- Die Know-how-Träger für BI sind an verschiedenen Stellen der Organisation angesiedelt. Eine gezielte Kommunikation erfolgt nicht.
- Das Antwortzeitenverhalten der BI-Systeme ist aus Sicht der Anwender unbefriedigend. Rein technologische Verbesserungsversuche waren bisher nicht erfolgreich.
- Die Realisierungszeiten für die Umsetzung neuer Anforderungen in existierenden BI-Systemen haben sich kontinuierlich verlängert bzw. werden von den Anwendern als zu langsam bewertet.
- Die Unternehmensorganisation ist nicht in der Lage, kurzfristig auf BI-Trendthemen (siehe vorherige Beispiele) zu reagieren.

Sollten mehrerer dieser oder ähnliche Fragestellungen zutreffen, empfiehlt sich die Entwicklung einer BI-Strategie.

3.2 Entwicklung einer BI-Strategie

Eine BI-Strategie ist die zukunftsorientierte Gesamtplanung aller BI-Initiativen und -Maßnahmen abgeleitet aus der Geschäfts- und der IT-Strategie eines Unternehmens.

3.2.1 Business-Intelligence-Strategie

Durch eine BI-Strategie wird die grundsätzliche Ausrichtung der BI-Landschaft eines Unternehmens bestimmt. Ihre Ziele sollten aus einer übergeordneten und langfristigen Perspektive definiert werden. Das strategische Management der BI-Strategie sorgt dafür, die notwendigen Entscheidungen und Maßnahmen durchzusetzen, um damit den langfristigen Erfolg von Business Intelligence im Unternehmen sicherzustellen.

Generell hat sich eine BI-Strategie konsequent an den Unternehmens- und IT-Zielen auszurichten und muss letztendlich alle Maßnahmen beschreiben, die durch eine optimale Informationsversorgung den Unternehmenserfolg ermöglichen. Daher sind die Unternehmens- und IT-Strategie Ausgangsbasis für die Entwicklung einer BI-Strategie. Im Sinne eines zielgerichteten Corporate Performance Management (Hoffmann 2002, S. 88) sollte die Ableitung der Ziele der BI-Strategie top down erfolgen. Neben den Unternehmens- und IT-Zielen bilden die kritischen Erfolgsfaktoren und die strategischen Maßnahmen den Ausgangspunkt für die Herleitung einer Kausalkette. Diese mündet über verschiedene Zwischenschritte schließlich in ein Konzept für die Steuerung/Informationsversorgung, der Architektur/Technologie sowie der Aufbau- und Ablauforganisation. In einer Roadmap wird der zukünftige Entwicklungspfad in Form eines Phasenmodells definiert.

3.2.2 Vorgehensmodell zur Entwicklung einer BI-Strategie

Vorgehensmodelle beschreiben die formale Abfolge von Prozessschritten, die zu einem gewünschten Ergebnis führen, z. B. in Form eines Produktes oder einer Leistung. Inhaltlich kommen innerhalb der Modellphasen meist unterschiedliche Methoden und Techniken zum Einsatz. Die Schritte werden in der Regel so allgemein definiert, dass sie übertragbar sind auf vergleichbare Situationen, also wieder verwendet werden können. Vorgehensmodelle finden sowohl im fachlichen (z. B. für die Entwicklung von Unternehmensstrategien) als auch im technologischen Umfeld (z. B. in der Softwareentwicklung) Anwendung. Aufgrund des zunehmenden Bedarfs gibt es inzwischen eine Reihe von Vorgehensmodellen für die Entwicklung von BI-Strategien. Eine ausführliche Übersicht über solche Modelle findet sich bei (Lahrman 2011, S. 195). Darüber hinaus gibt es eine Reihe von Konzepten zur Entwicklung von BI-Teilstrategien wie z. B. für die Erreichung einer hohen Datenqualität oder für die strategische Ausrichtung von Data Warehouses.

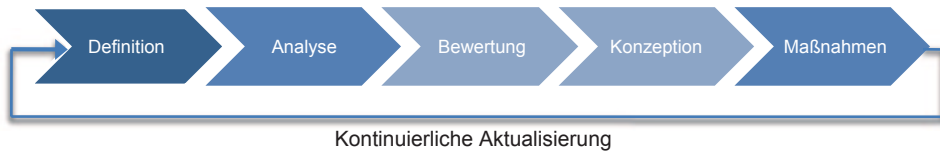


Abb. 3.1 Phasen des Strategieentwicklungsprozesses

Die Entwicklung einer Strategie ist grundsätzlich ein kontinuierlicher, zyklischer Prozess, der aufgrund der langfristigen Ausrichtung über mehrere Jahre wirkt. Typischerweise wird eine BI-Strategie initial in einem mehrphasigen Vorgehen in ca. 3–6 Monaten entwickelt. Die Weiterentwicklung der Bestandteile, wie z. B. der Maßnahmenplanung, erfolgt danach fortlaufend im Rahmen des Anforderungsmanagements in beispielsweise monatlichen Abständen und/oder jährlich als Grundlage für den Budgetierungsprozess. In mehrjährigen Abständen sollte die Strategie komplett überprüft und an die Ausrichtung des Unternehmens angepasst werden.

Der Phasenablauf für die Entwicklung einer BI-Strategie soll im Folgenden analog zum strategischen Entwicklungsprozess eines Unternehmens definiert werden (siehe Abb. 3.1). Nicht Bestandteil der Entwicklung und daher auch nicht in der Abbildung dargestellt sind die an die Entwicklung anschließenden Phasen der eigentlichen Umsetzung sowie das Controlling der Umsetzung.

Eine BI-Strategie wird typischerweise aus mehreren Betrachtungsperspektiven entwickelt. Die wesentlichen drei Perspektiven sind Fachlichkeit, Architektur/Technologie und Organisation, wie aus Abb. 3.2 ersichtlich wird. Strategische Projekte innerhalb

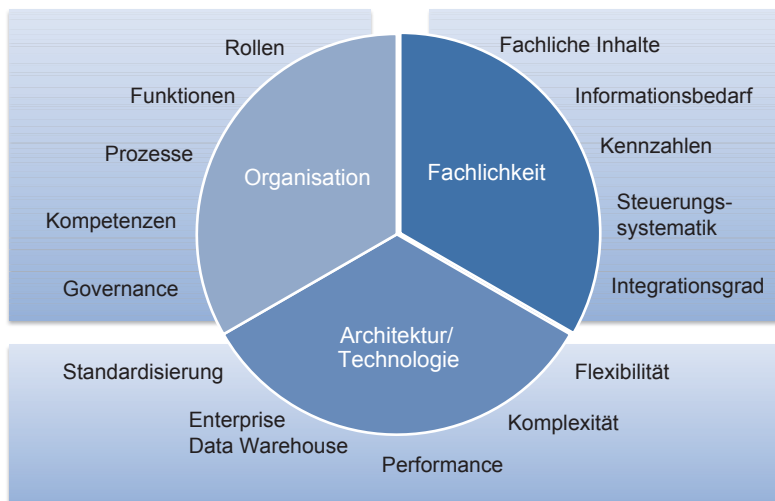


Abb. 3.2 Perspektiven einer BI-Strategie (Chamoni und Gluchowski 2004, S. 120; Vierkorn 2008, o. S.)

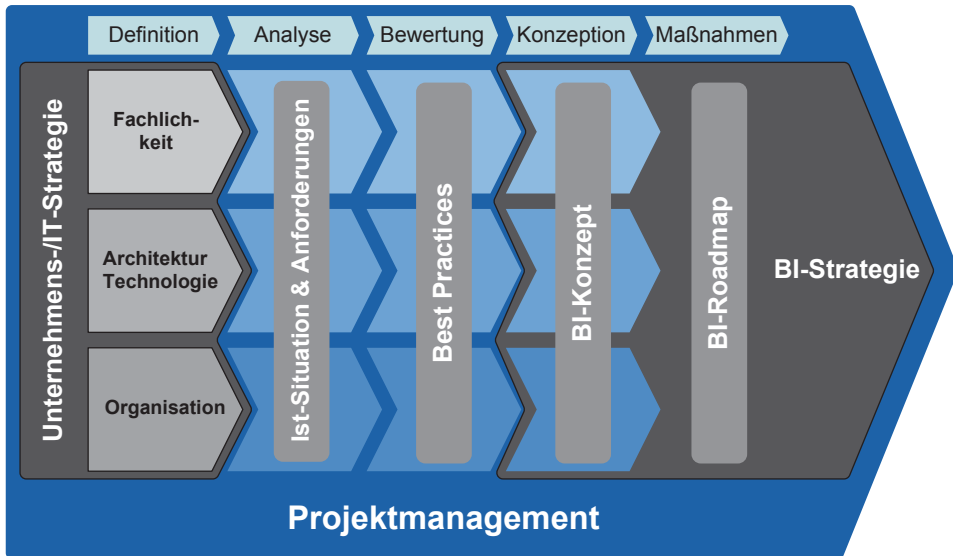


Abb. 3.3 Vorgehensmodell für eine BI-Strategie (Gansor et al. 2010, S. 75)

eines Unternehmens können sich prinzipiell auch jeweils nur mit einer der drei Perspektiven beschäftigen, z. B. nur mit Architektur/Technologie. Dies geschieht beispielsweise im Rahmen der Entwicklung einer Technologie-Strategie. Ein anderes Vorgehen ist die Konzentration auf die Perspektive der Fachlichkeit, wenn sich das Unternehmen mit der Entwicklung einer neuen Steuerungssystematik, wie nach der Balanced Scorecard, beschäftigt. Ein ganzheitlicher (holistischer) Ansatz zur Entwicklung einer BI-Strategie im Sinne des Autors beinhaltet allerdings die Berücksichtigung aller drei Perspektiven. Die Schwerpunkte können bei einer konkreten Strategieentwicklung jedoch zwischen den unterschiedlichen Perspektiven variieren – je nachdem, welche Priorität die verschiedenen Anforderungen aus den drei Bereichen für das jeweilige Unternehmen haben und auf welcher ggf. bereits vorhandenen Basis aufgebaut werden kann.

Kombiniert man den Phasenablauf der Strategieentwicklung mit den Perspektiven, so entsteht eine zweidimensionale Struktur (siehe Abb. 3.3), die sich in 10 Prozessschritte gliedern lässt. Die Phasen geben die Reihenfolge der Entwicklungsschritte vor, die Perspektiven die inhaltliche Differenzierung der zu betrachtenden Themen. Die Darstellung sollte jedoch nicht so verstanden werden, dass zwingend alle drei Perspektiven gleichzeitig abgearbeitet werden müssten. Die Perspektiven können parallel, nacheinander (Fachlichkeit → Architektur/Technologie → Organisation) oder auch in Kombination ausgearbeitet werden, und dies jeweils in den Phasen Analyse, Bewertung, Konzeption und Maßnahmen. Sinnvollerweise steht jedoch meist die Untersuchung der Fachlichkeit vor Architektur und Technologie. Oftmals müssen auch zunächst informatorische und architektonische Details konzipiert werden, bevor Aussagen zur Organisation gemacht werden können.

Die zu liefernden Ergebnisse (Deliverables) des Strategieentwicklungsprozesses sind unter anderem:

- Ist-Situation und Anforderungen: Beschreibung der Ist-Dokumentation und Erhebung der Anforderungen (Methode: z. B. Informationsbedarfsanalyse)
- Best Practices: Bewertung der Ist-Situation im Vergleich zu den Anforderungen und zu Best Practices aus anderen Unternehmen (Methoden: z. B. Gap-/SWOT-Analysen, Benchmarking)
- BI-Konzept: Beschreibung von Vision, Zielen, Informationsangebot, Architektur, Softwareportfolio, Technologie etc.
- BI-Roadmap: Entwicklung eines stufenweisen Maßnahmenplans und des Projektportfolios für die sukzessive Umsetzung des Konzepts

In den folgenden drei Abschnitten werden die drei Perspektiven einer BI-Strategie vorgestellt. In den letzten beiden Abschnitten dieses Kapitels wird auf Methoden und Prozesse sowie auf Bewertung und Dokumentation zur Entwicklung einer BI-Strategie eingegangen.

Fachlichkeit

Die fachliche Anwendung ist der eigentliche Sinn von Business Intelligence. Die angewandten betriebswirtschaftlichen Verfahren und Methoden werden hinsichtlich ihrer Auswirkungen auf die BI-Anwendungslandschaft untersucht. Hierzu zählen unter anderem die Art der Planung und der Hochrechnung, der Kalkulation und der Deckungsbeitragsrechnung, die auf einem mittleren Detaillierungsgrad erfasst und systematisiert werden. Mit einher geht die Identifizierung der Entscheidungsobjekte, wie Kunden, Produkte, Filialen oder interner Organisation. Es handelt sich dabei meist um die betriebswirtschaftlichen Dimensionen in einem multidimensionalen Modell, wie sie für den Aufbau von Analytischen Informationssystemen charakteristisch sind (Totok 2000, S. 87). Verfahren, Methoden und auch Entscheidungsobjekte werden schließlich hinsichtlich ihrer Wirkung im Managementprozess beurteilt.

Als zentraler Bestandteil der fachlichen Analyse werden in „klassischer“ Form Informationsbedarf und -angebot abgeglichen, um eine Deckung zu erreichen. Grundlage hierfür bilden Performance-Metriken, die sich aus kritischen Erfolgsfaktoren und den betriebswirtschaftlichen Verfahren und Methoden ableiten lassen (s. Abb. 3.6). Es wird bestimmt, welche Informationen zur Unterstützung des Managementprozesses notwendig sind. Dabei wird bis auf die operativen Geschäftsprozesse heruntergebrochen, da BI-Systeme einerseits auf den Daten operativer Geschäftsprozesse basieren, andererseits im Sinne eines Regelkreises („Closed Loop“), operative Geschäftsprozesse durch BI-Systeme gesteuert werden. Daher wird eine Informationslogistik definiert, die das Zusammenwirken der Informationen von Managementprozessen und operativen Geschäftsprozessen beschreibt.

In großen BI-Strategieprojekten kann die Anzahl der ermittelten Informationsbedarfe, abhängig von vom vorher festgelegten Detaillierungsgrad, bis in die Hunderte gehen. Eine

Tab. 3.1 Erhebung Informationsbedarfe (vereinfachte Darstellung)

ID	Informations- bedarf	Beschreibung	Nutzen	Typ	Priori- tät ^a
IB 5.1.1	Reklama-tions- quote	Detaillierte Statistik nach Reklamations- gründen pro Produkt	Aufdeckung von Qualitätsmängeln	Qualitäts- management	4
IB 5.3.1	Kundenprofil	Kunden nach Eigen- schaften segmentieren und Profile erstellen	Deutliche Steige- rung Kundenzu- friedenheit durch individuelle Ansprache	CRM	5
IB 7.1.1	Profit-Center- Rechnung nach Produkten	Detaillierte Ergebnis- rechnung pro PC nach Produkten detaillierbar	Verbesserte Produktplanung	Ergebnis- rechnung	3
IB 7.3.1	Profit-Center- Rechnung nach Vertriebswegen	Ergebnisrechnung pro PC nach Vertriebswe- gen detaillierbar	Verbesserte Ver- triebssteuerung	Ergebnis- rechnung	5

^a 5: entscheidend, 4: sehr wichtig, 3: wichtig, 2: relevant, 1: wünschenswert

sehr anspruchsvolle Aufgabe ist die sich anschließende Clusterung bzw. Typisierung der Informationsbedarfe zu zusammenhängenden Themenblöcken. Diese Aufgabe kann meist nur mit Unterstützung von Schlüsselpersonen des Unternehmens erfolgen, die über sehr gutes internes Querschnitts-Know-how verfügen. In Tab. 3.1 wird die Dokumentation von Informationsbedarfen in vereinfachter Form dargestellt. Jeder Informationsnachfrager sollte im Rahmen des Strategieprozesses seinen Bedarf anhand der Unternehmensziele erläutern und den Nutzen darstellen können. Informationsbedarfe, die sich nicht in dieser Form herleiten lassen, sind kritisch zu hinterfragen und normalerweise eher niedrig zu priorisieren.

Architektur/Technologie

Die Bestimmung der geeigneten Architektur für eine strategische Anwendungslandschaft für Business Intelligence ist die Kernaufgabe einer BI-Strategie. Dabei müssen die spezifischen Eigenschaften von Business Intelligence berücksichtigt werden. Abbildung 3.4 zeigt die wichtigsten Themengebiete als Dimensionen, die bei der Architekturgestaltung berücksichtigt werden sollten. Alle bestehenden und auch alle neuen, im Rahmen der Strategie ermittelten BI-Lösungen sollten nach den Dimensionen geordnet und beschrieben werden. Dadurch wird Vergleichbarkeit erreicht und ähnliche Lösungen können identifiziert und gruppiert werden. Dies ist wiederum die Ausgangsbasis für die Ermittlung von Synergiepotenzialen und die Vereinheitlichung der Architektur.

Organisation

Viele Probleme in der praktischen Anwendung von Business Intelligence lassen sich auf Unstimmigkeiten in der Aufbau- und Ablauforganisation zurückführen. Daher müssen für

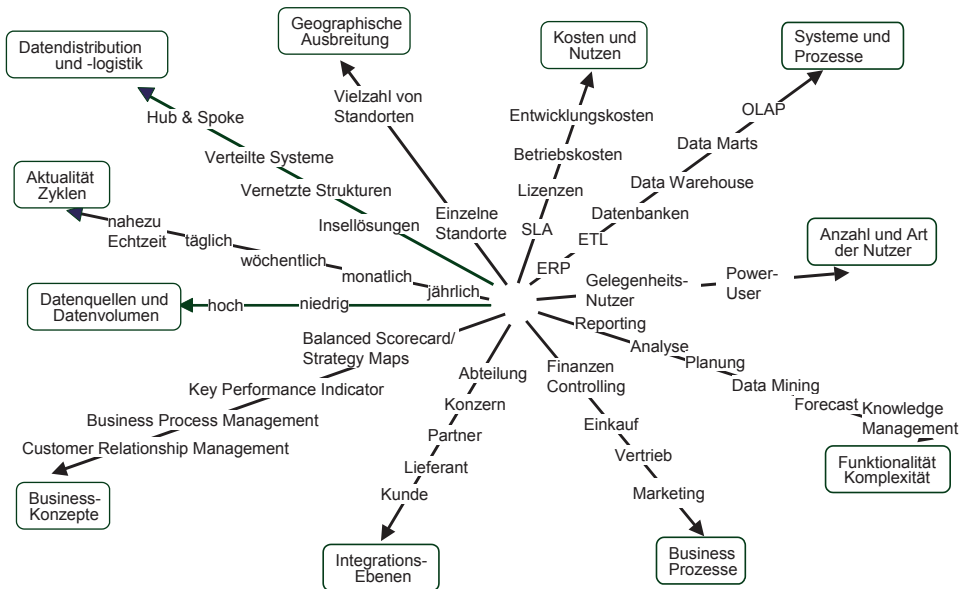


Abb. 3.4 Dimensionen einer BI-Architektur

eine erfolgreiche Nutzung von Business Intelligence die übergreifende Zusammenarbeit der Fachabteilung und das Zusammenspiel zwischen Fachbereichen und IT funktionieren. Ein Prozess, der häufigen Anpassungen unterliegt, ist beispielsweise der Planungsprozess. Ist sichergestellt, dass jede Änderung ausreichend an alle Beteiligten kommuniziert und im System abgebildet wird? Reagiert die Planungsadministration schnell genug auf Prozessanpassungen? Werden Entscheidungen im Planungsprozess ausreichend dokumentiert, sodass diese später auch für Kontrollinstanzen plausibel sind? Das Ziel lautet, alle BI-Prozesse verlässlich und nachvollziehbar zu gestalten. Zur Effizienzverbesserung der BI-Organisation empfiehlt sich der Aufbau eines BICC (siehe Abschn. 3).

Ein organisatorischer Aspekt, der inzwischen auch für BI relevant ist, ist das Outsourcing von Dienstleistungen. Standen in der Vergangenheit eher infrastrukturelle Themen im Vordergrund, so wird heute die Auslagerung von Entwicklung und Betrieb von BI-Applikationen vorangetrieben. Hierbei sollte beachtet werden, dass für BI ein tiefgehendes Verständnis der im Unternehmen ablaufenden Prozesse und deren Zusammenwirkens mit zahlreichen Systemen notwendig ist. Eine genaue Kenntnis über die ablaufenden BI-Prozesse und deren Dokumentation ist die Voraussetzung für ein erfolgreiches Outsourcing. Die Auslagerung von BI-Prozessen sollte überhaupt nur dann erwogen werden, wenn im Unternehmen bereits entsprechende Erfahrungen gesammelt wurden (Philippi 2004, S. 104). Sollte es zur externen Vergabe kommen, so sollte ein schrittweises Vorgehen zunächst mit BI-Prozessen hohe Standardisierbarkeit gewählt werden. Hierzu können z. B. das Monitoring der Ladeprozesse oder der First Level Support zählen.

Zur Sicherstellung der „Spielregeln“ für den Einsatz von Business Intelligence empfiehlt sich die Definition einer BI-Governance, die den Rahmen für die BI-Nutzung

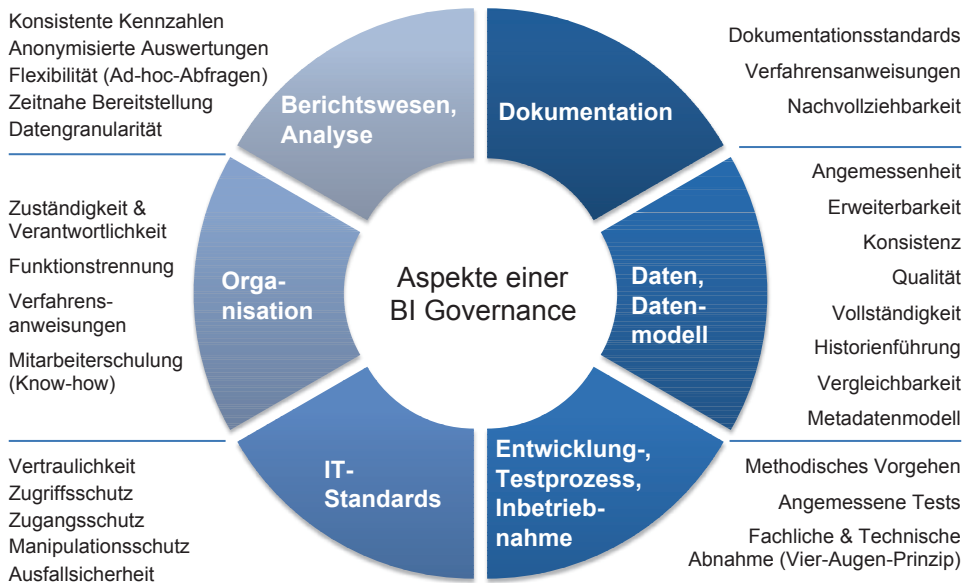


Abb. 3.5 Aspekte einer BI Governance

beschreibt. Dadurch soll sichergestellt werden, dass die Anwendung von Business Intelligence kontinuierlich auf Unternehmensziele und -prozesse ausgerichtet ist. Insofern kann diese aus einer wie hier beschriebenen BI-Strategie direkt abgeleitet werden. Weiterhin soll durch eine BI-Governance bewirkt werden, dass die Ressourcen für BI verantwortungsvoll eingesetzt und Risiken minimiert werden. Der Regelungsumfang sollten mit „Fingerspitzengefühl“ gewählt werden. Zuviel Regelung widerspricht dem eigentlich gewünschten spontanen Umgang mit dispositiven Informationen. Zuwenig Governance kann demgegenüber zu einem Vertrauensverlust der Dateninhalte führen. Aber auch regulatorische, externe Aspekte können zu einem Zwang zu Governance führen. So fordern aktuelle regulatorische Vorgaben im Bankenumfeld explizit den Nachweis einer nachvollziehbaren Herleitung von Risikokennzahlen bei hoher Datenqualität. Abbildung 3.5 gibt eine Übersicht möglicher durch eine Governance zu regelnde Themenbereiche.

BI-Konzept und Roadmap

Die Ergebnisse der BI-Strategie werden in BI-Konzept und Roadmap festgehalten. In der Roadmap werden alle Initiativen und deren Abhängigkeiten beschrieben, die im Strategieprozess definiert wurden. Eine abschließende Priorisierung aller – auch konkurrierenden – Anforderungen ermöglicht die Aufstellung eines Projektportfolios. Das Spektrum der Projekte kann dabei vom Aufbau neuer BI-Systeme, über die unternehmensweite Stammdaten-Standardisierung bis hin zum Redesign von Geschäftsprozessen reichen. Wichtig ist, dass durch die Beschreibung der Abhängigkeiten und die übergreifende Priorisierung nachvollziehbar ist, welche Projekte unmittelbare Voraussetzung für Folgeprojekte sind und von welchen Projekten der höchste Nutzen für das Unternehmen erwartet wird.

Methoden und Prozesse

Für die Entwicklung einer BI-Strategie wird eine Auswahl geeigneter Methoden und Prozesse benötigt. Die initiale Entwicklung einer BI-Strategie erfolgt in der Regel in Form eines Projektes. Das Kern-Projektteam sollte sich aus Schlüsselpersonen der einzelnen Fachbereiche und der IT zusammensetzen. Eine externe Unterstützung ist für die Moderation von Workshops sinnvoll oder wenn intern nicht genügend Ressourcen oder Know-how vorhanden sind. In einem erweiterten Projektteam sollten Vertreter von Unternehmensteilen aufgenommen werden, die BI-Lösungen intensiv einsetzen und ggf. als Multiplikatoren in ihren Bereichen agieren. Bei der Aufstellung des erweiterten Teams muss auf die internationale Vernetzung geachtet werden. Gerade im internen und internationalen Vergleich lassen sich kreative und vielleicht sogar unkonventionelle Lösungen für Business Intelligence identifizieren und gewinnbringend auf andere Bereiche übertragen. Das höhere Management sollte in die Strategiefindung im Rahmen eines Lenkungsausschusses und natürlich als Input-Geber in Workshops und Interviews eingebunden werden. Auf eine, in Veröffentlichungen oftmals geforderte, herausragende Rolle und ein starkes Engagement des Top Managements in BI-Projekten wird hier nicht bestanden. Vielmehr ist die Benennung eines Sponsors auf Top-Ebene in Verbindung mit der Teilnahme im Lenkungsausschuss die realistischere Form der Beteiligung.

Im Kick-off-Meeting für das BI-Strategieprojekt trifft das erweiterte Projektteam zum ersten Mal zusammen und die Projektziele und die Vorgehensweise werden konkretisiert. Danach folgen Workshops und Interviews zur Ermittlung und Systematisierung der Informationsbedarfe und der relevanten Performance-Metriken (siehe Abb. 3.6). Die Systematisierung kann durch Einbettung der Metriken in ein Kennzahlensystem, z. B. nach den Perspektiven einer Balanced Scorecard, erfolgen. In der folgenden Phase werden die Auswirkungen auf Reporting und Analyse untersucht und die Anpassung von bestehenden Applikationen und IT-Systemen überprüft. Hierbei wird das Unternehmen von seinem Reifegrad hinsichtlich des Lebenszyklus der vorhandenen BI-Landschaft her beurteilt.

Die Workshops und Interviews müssen intensiv vorbereitet werden. Durch eine zielführende Moderation wird sichergestellt, dass Ergebnisse im vorgegebenen Zeitrahmen erzielt werden. Die Tiefe der Diskussion sollte im Kern-Projektteam möglichst vorher abgestimmt werden, um sich nicht in Details zu verlieren. Vereinzelt neigen Teilnehmer an

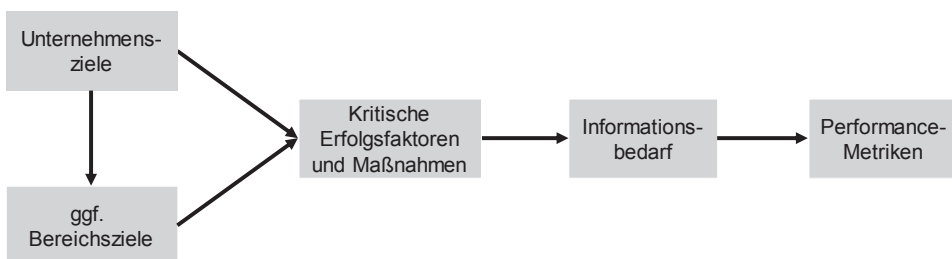


Abb. 3.6 Prozess zur Ableitung von Informationsbedarf und Performance- Metriken

Workshops in Strategieprojekten dazu, die Vergangenheit sehr ausführlich aufzuarbeiten. Hier muss die Moderation die Diskussionsthemen geschickt in Richtung Zukunft steuern.

Geeignete Instrumente für die Workshops sind z. B. das Sammeln von Beiträgen mit Metaplan-Karten oder die Dokumentation in Form von Mindmaps. Beim Einsatz von Metaplan-Karten sollte vorher die Farbcodierung festgelegt sein. Werden die Themen z. B. nach den Perspektiven der Balanced Scorecard systematisiert, so ist es sinnvoll, wenn jede Perspektive ihre eigene Farbe besitzt.

In den Workshops finden sich unterschiedliche Interessensgruppen zusammen, so dass die Ergebnisse verschiedener Workshops zu konkurrierenden Anforderungen führen können. Die an den Workshops teilnehmenden Mitglieder des Projektteams sollten Gegensätze möglichst schon in den Workshops aufzeigen und frühzeitig klären. Dennoch treten bei der Konsolidierung der Workshop-Ergebnisse häufig Zielkonflikte auf. Deren Auflösung kann z. B. mit Hilfe einer Konfliktmatrix erfolgen (Füting 2000, S. 276). Eine letztendliche Klärung erfolgt in der Regel im Projekt-Lenkungsausschuss unter Beteiligung der Entscheider.

Bewertung und Dokumentation

Die Anzahl der zu behandelnden Metriken, Entscheidungsobjekte und Informationsbedarfe kann im Rahmen einer BI-Strategiediskussion sehr umfangreich werden. Damit hier nicht der Überblick verloren geht, ist eine durchgängige Bewertung und strukturierte Dokumentation aller genannten Themen unerlässlich. Die Dokumentation sollte dabei die notwendige Flexibilität besitzen, um auf veränderte Anforderungen reagieren zu können. Verändern sich z. B. Prioritäten von Zielen, sollte der Einfluss auf Informationsbedarfe und Performance Metriken und damit natürlich auch auf den Masterplan sehr leicht nachvollziehbar sein. Weiterhin sind Auswertungsmöglichkeiten nach verschiedenen Informationsbedarfen oder Performance-Metriken sehr hilfreich, wenn z. B. die Frage gestellt wird, welche Metriken voneinander abhängen oder welche Datenquellen „angezapft“ werden müssen, um die gewünschte Information zu erhalten. Die traditionelle Art der Dokumentation basierend auf Textverarbeitungs- oder Präsentationsprogrammen hat sich für diese Zwecke als zu statisch herausgestellt. Daher empfiehlt sich die Dokumentation in einem Tabellenkalkulationsprogramm oder sogar als Datenbankanwendung.

Die Dokumentation einer BI-Strategie in Form voneinander abhängiger Tabellenblättern in einer Tabellenkalkulation wie Excel stellt einen guten Kompromiss zwischen Anwenderfreundlichkeit für nicht IT-Fachleute und Funktionalität dar. Selektionen können z. B. über die Autofilter-Funktion erfolgen und Abhängigkeiten zwischen Zielen, Informationsbedarfen etc. können über „S-Verweise“ modelliert werden. Jeder Schritt des Prozesses aus Abb. 3.6 wird dabei als eigene Tabelle angelegt. Unternehmensziele, Erfolgsfaktoren etc. werden über eine eindeutige Nummerierung mit einander in Beziehung gesetzt. Die Prioritäten werden top down durchgerechnet und ermöglichen schließlich eine Gesamtbewertung. In Tab. 3.2 wird in stark vereinfachter Form die Erfassung der Unternehmensziele in einer Tabelle gezeigt. Idealerweise liegt das Zielsystem des Unter-

Tab. 3.2 Dokumentation der Unternehmensziele

Unternehmensziele		Priorität ^a
Z1	Kostenstruktur optimieren	4
Z2	ROI von 12% erreichen	5
Z3	Marktrelevante Markenpositionierung vorantreiben	3
Z4	Marktführende Stellung in den wichtigsten, profitablen Marktsegmenten in Europa erreichen	3
Z5	Kundenzufriedenheit und Qualitätsbewusstsein erhöhen	4

^a 5: entscheidend, 4: sehr wichtig, 3: wichtig, 2: relevant, 1: wünschenswert

nehmens fertig dokumentiert, z. B. in Form einer Strategy Map vor und muss nur in die Projektdokumentation übernommen werden.

3.3 Aufbau eines BI Competency Center

Eine Grundlage für erfolgreiche BI ist die richtige organisatorische Einbettung in das Unternehmen. In der Praxis findet man jedoch häufig stark verteilte Verantwortlichkeiten für die entscheidungsorientierte Informationsversorgung vor. So existiert meist die klassische Trennung zwischen Fach- und IT-Bereich mit zusätzlichen Differenzierungen, wie Controlling, Vertrieb oder Marketing auf der fachlichen Seite. Im IT-Bereich wird häufig zwischen Datenbankadministration, Entwicklung und Architektur unterschieden (Adelman und Moss 2004, S. 183). Eine gemeinsame Benutzer- oder IT-Organisation für BI ist nur selten anzutreffen.

Unternehmen berichten aber, dass ihre BI-Aktivitäten insbesondere dann erfolgreich sind, wenn eine starke Brücke zwischen den verschiedenen Betreiber- und Nutzereinheiten geschlagen wird. Dies lässt sich durch die enge Verknüpfung von Prozessen, Inhalten und Funktionen erklären, die für BI-Anwendungen typisch sind. So existiert z. B. bei einem großen deutschen Finanzinstitut eine gemeinsame BI-Einheit, die sowohl Services für die Fachabteilungen wie Analysen, Berichtswesen oder Beratung anbietet, als auch die fachliche und technologische Entwicklung der Systeme verantwortet und zum großen Teil selbst vornimmt. Die Betreuung der technologischen Infrastruktur, wie Aufbau, Installation und Betrieb der Hardware, aber auch die Basiskonfiguration der Datenbanksysteme wird hingegen vom klassischen IT-Dienstleister vorgenommen. Der Vorteil dieser Konstellation liegt im tiefgehenden ganzheitlichen Know-how, über das diese BI-Einheit verfügt. Sie ist in der Lage, äußerst kurzfristig auf veränderte oder neue Anforderungen zu reagieren und verfügt darüber hinaus über eine hohe Kreativität bei der Schaffung nutzenbringender BI-Lösungen.

Eine sinnvolle Möglichkeit ist die Gründung eines BI Comptency Centers, das das Verbindungsglied zwischen Endanwendern, Software-Entwicklung und Betrieb darstellt (Miller et al. 2006, S. 9). Damit ist das Comptency Center für Fachbereiche und IT der

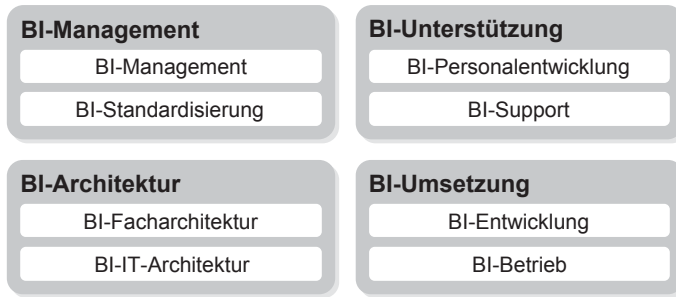


Abb. 3.7 Funktionen eines BI Competency Center (Gansor et al. 2010, S. 118)

zentrale Ansprechpartner für Business Intelligence. Das Competency Center muss dabei das Know-how sowohl für alle relevanten Geschäftsprozesse auf der fachlichen Seite, als auch für Datenmodell und -integration auf der technischen Seite besitzen. Es berät alle Seiten konzeptionell und gewährleistet die strategische Entwicklung und Nutzung der Systeme. Abbildung 3.7 zeigt typische Funktionen eines BICC.

Auch wenn keine gemeinsame Einheit für Business Intelligence eingerichtet wird, so sollte im Unternehmen zumindest eine reguläre Basis für den Erfahrungsaustausch und für die abgestimmte Definition neuer Anforderungen vorhanden sein. Hierfür bieten sich z. B. monatlich Meetings der Key User und Projektleiter auf nationaler und (zumindest) halbjährliche Workshops der Verantwortlichen auf internationaler Ebene an.

3.4 Ausgewählte Aspekte einer BI-Strategie

3.4.1 Kosten und Nutzen bewerten

Die Frage nach Kosten und Nutzen von BI-Lösungen ist in Unternehmen oftmals nicht ausreichend geklärt. In vielen Unternehmen findet zwar eine kontinuierliche (30%) oder einmalige (47%) Kostenkontrolle statt, es fehlt demgegenüber jedoch die Transparenz des Nutzens (36%) (Chamoni und Gluchowski 2004, S. 125). Eine unternehmensweite Übersicht über die Kosten von Business Intelligence zu erstellen, fällt vielen Unternehmen schwer, da die Kosten entweder in den verschiedenen Bereichen und Unternehmensteilen separat budgetiert oder viele Kosten gar nicht auf Business Intelligence bezogen erfasst werden. Ein Controlling findet eher projektbezogen auf einzelne Realisierungsschritte, als ganzheitlich über den Lebenszyklus statt. Kosten und Aufwände lassen sich zwar über geeignete Maßnahmen erfassen, schwieriger wird es aber bei der Nutzenquantifizierung. Wie hoch ist z. B. der monetäre Wert eines wöchentlichen Berichtes? Auf Fragen dieser Art wird in der Praxis gerne ausweichend mit der Gegenfrage nach den Kosten des Berichtes reagiert. Generell fällt es den meisten Informationsnutzern relativ schwer zu beurteilen, was eine Information für das Unternehmen wert ist.

Die Ermittlung der Ist-Kosten stellt einen ersten Schritt für die Ermittlung der Wirtschaftlichkeit von Business Intelligence dar. Existieren unterschiedliche BI-Lösungen an verschiedenen Stellen im Unternehmen, so werden die zugeordneten Kosten bzw. Aufwände meist nicht zusammenhängend erfasst. Die Ermittlung der Ist-Kosten sollte sich übrigens nicht nur auf den IT-Bereich beschränken, sondern sich bis zu den Fachanwendern erstrecken.

Aufgrund dieser sich wechselseitig bedingenden Faktoren lässt sich die eigentliche Wirtschaftlichkeit von BI nur sehr schwer ermitteln. Grundsätzlich kann man das Problem auf den Wert einer einzelnen Information zurückführen: je operativer diese ist, d. h. je größer ihr Einfluss auf den Verlauf eines Geschäftsprozesses ist, desto einfacher lässt sich ihr Wert bestimmen. Kommt z. B. der Verkauf eines Artikels aufgrund einer fehlenden oder falschen Information nicht zustande, so entspricht der Wert einer Information im einfachsten Fall dem Deckungsbeitrag des Artikels. Hingegen kann die Nicht-Entscheidung eines Managers aufgrund falscher oder fehlender Information das Unternehmen fast in beliebiger Höhe belasten, und die Auswirkungen werden trotzdem im Einzelfall erst nach Monaten deutlich.

Eine oft angewendete Vorgehensweise zur Bewertung der Wirtschaftlichkeit von BI liegt in der qualitativen Nutzenbetrachtung von geplanten Projekten oder Systemen. Der Informationsnutzer beschreibt qualitativ, wie ihm die Information hilft, z. B., dass die Qualität seiner Entscheidungen steigt. Besser ist es jedoch, wenn sich der Wert aus einem konsistenten und durchgängigen Zielsystem im Unternehmen ableitet. Aktuelle Informationen über wichtige Performance-Metriken besitzen bei dieser Vorgehensweise einen hohen Wert. Metriken von niedriger Priorität dürfen nur wenig kosten, da sie einen geringeren Wert besitzen. Der Gesamtwertansatz von BI wird über einen Lebenszyklus betrachtet und dabei an die geplante Entwicklung der Unternehmensziele über einen längeren Zeitraum gekoppelt. Der ROI der BI-Anwendungslandschaft ist damit von den Unternehmenszielen abhängig. Es würde nicht länger der Informationsnutzer fragen, was die Information kostet, sondern das Unternehmen würde insgesamt bestimmen, was die Information wert ist. Leider kann an dieser Stelle keine allgemeingültige Berechnungsvorschrift für den ROI von Business Intelligence gegeben werden, da die Anwendungssituationen zu unterschiedlich sind. Jedes Unternehmen benötigt einen individuellen Ansatz, der der jeweiligen Ausprägung von Business Intelligence gerecht wird (Watson et al. 2004, S. 16).

Eine vollständige Kostenermittlung im Rahmen eines Strategieprojektes kann allerdings auch einen ungewollten Effekt haben: Entscheidungsträger, die zum ersten Mal mit den tatsächlichen Gesamtkosten von BI konfrontiert werden, könnten hier spontan zu Kosteneinsparungen neigen, da ihnen die Beträge auf den ersten Blick als zu hoch erscheinen. In einem DAX-notierten deutschen Unternehmen beispielsweise stellt das Enterprise Data Warehouse die höchste Einzelposition im IT-Budget dar. Dies wird vom Management allerdings auch bewusst akzeptiert.

3.4.2 Auf Trends reagieren

Trendthemen sind Themen, die kurzfristig auf die Unternehmen und damit auf die BI-Strategie oder das BICC wirken, da sie im Markt eine gewisse Zeit lang eine große Aufmerksamkeit genießen. Trends wie Big Data oder Self Service BI sind einleitend ja auch bereits genannt worden. Fachanwender werden in solchen Hype-Phasen geradezu mit Marketing- und Vertriebsaktionen der Softwareanbieter und Dienstleister bombardiert und erzeugen dann unternehmensintern einen nicht zu unterschätzenden Druck des „Haben-Wollens“. Dieser Druck kann in übereilten Entscheidungen münden, neue Technologien auf die Schnelle einführen zu wollen, ohne der passenden Integration in die bestehende Landschaft Rechnung zu tragen. Business Cases werden in solchen Fällen – wenn überhaupt – gerne sehr optimistisch gerechnet, so dass sich der langfristige fachliche Nutzen evtl. nicht wirklich einstellt.

Ein BICC sollte es als originäre Aufgabe verstehen, nicht etwa Trends abzuwehren, sondern diese als Innovator zu treiben. Leider mussten gerade neu gegründete BICCs in den letzten Jahren oftmals die Rolle von Konsolidierern, internen Bremsern oder „Regulatoren“ annehmen, da einfach zu viele Probleme wie z. B. bzgl. mangelnder Datenqualität oder sanierungsbedürftiger Architekturelemente auf Lösungen warteten und ganz oben in der Prioritätenliste der BI-Strategie standen. In solchen Situationen besteht permanent die Gefahr, dass Trendthemen dann einfach am BICC vorbei realisiert werden. Dieser Gefahr muss das BICC durch einen guten Mix von Sanierungs- und Trend- bzw. Zukunftsthemen in der Prioritätenliste begegnen. Sollte beispielsweise ein Trend wie die Self Service durch Fachbereiche als wichtiges Thema genannt werden, so sollte das BICC bzw. die für BI zuständige Organisationseinheit den Ball aufnehmen und das weitere Vorgehen steuern. Hierzu zählen die Aufnahme der Anforderungen, ggf. die Softwareauswahl, die Anpassungen von Prozessen und Architekturelementen bis hin zum Aufsetzen eines Einführungsprojektes. Das BICC sollte dabei dann aber auch gemeinsam mit den beteiligten Fachbereichen die Spielregeln zur Nutzung von mehr Self-Service-Möglichkeiten festlegen, so dass diese strategiekonform zum Einsatz kommen.

3.4.3 Architektur gestalten

So vielfältig die BI-Anwendungen, so verschieden sind auch die architektonischen Ansätze, die in Unternehmen nebeneinander existieren. Das große, integrierte, zentrale Data Warehouse ist zwar oftmals vorhanden, deckt häufig aber nicht alle Themen im Detail ab. Vor allem besteht weiterhin die Trennung zwischen produkt-/kundenorientierter vs. finanzorientierter Data Warehouses. Eine hundertprozentige Integration der entscheidungsrelevanten Daten ist jedoch andererseits unter einer Kosten- und Nutzenbetrachtung nicht auf jeder Ebene sinnvoll, sodass die getrennte Behandlung stark unterschiedlicher Themenbereiche durchaus richtig sein kann. Dennoch sollen die Entscheidungsträger ihre Informationen vereinheitlicht und zusammenhängend erhalten – und das weltweit. Die

gewählte BI-Architektur sollte daher einen Bogen über alle Einsatzzwecke spannen und dabei ausreichend flexibel sein.

Die Konsolidierung und die damit einhergehende Standardisierung der BI-Anwendungslandschaft ist meist eine der am höchsten priorisierten architektonischen Anforderungen an eine BI-Strategie. Erwartet werden dabei vor allem Kosteneinsparungen aber auch eine Verbesserung der informationsverarbeitenden Prozesse und der Qualität. Die Konsolidierung bzw. Ablösung von bestehenden Teilsystemen kann sehr genau aus Kosten- und Nutzensicht beurteilt werden. Entwicklungsaufwände, Soft- und Hardwarekosten können meist gut geschätzt und damit die Amortisationsdauer bestimmt werden. Unternehmen, die bereits entsprechende Projekte durchgeführt haben, berichten, dass sich die Konsolidierung von Data Marts nach durchschnittlich 1–2 Jahren amortisiert hat (Eckerson 2004, S. 14). Konsolidiert und konsequent standardisiert werden sollten möglichst homogene Themengebiete, wie das Finanzreporting. Der Prozess beginnt mit der Vereinheitlichung der fachlichen Inhalte, wie der Bildung und Anwendung eines unternehmensweiten Kontenrahmens, der Einigung auf ein Softwareprodukt und der Ableitung einer gemeinsamen Hardwareplattform.

Als Best Practice für den Aufbau einer Data-Warehouse-Architektur wurde in der Vergangenheit die Etablierung eines Single Point of Truth empfohlen, in dem Unternehmensdaten bereinigt und integriert abgelegt werden. Bedingt durch neue Technologien wie Big-Data-Frameworks, die in der Lage sind, große Datenmengen zu einem sehr guten Preis-Leistungsverhältnis zu verwalten, ist in letzter Zeit eine neue Diskussion um die beste BI-Architektur entstanden. So ist beispielsweise das Konzept des Raw Data Lake bzw. Rohdatenpools aufgekommen, der dazu dient, Daten unterschiedlichster Herkunft im Originalformat zu speichern. Über Vereinheitlichung oder mögliche Transformationsschritte soll sich der Anwender erst zum Zeitpunkt der Datenverwendung Gedanken machen. Dies ist ein zum klassischen Modell fast schon radikal gegensätzlicher Ansatz. Es wird in den kommenden Jahren spannend werden zu beobachten, wie diese neuen Konzepte tatsächlich Einfluss auf BI-Architekturen erlangen werden.

3.4.4 Softwareportfolio festlegen

Die Vielfalt der am Markt erhältlichen Softwareprodukte für Business Intelligence hat sich trotz zahlreicher Übernahmen in den vergangenen zehn Jahren nicht verringert. Wurden im Jahr 2004 in der jährlichen Umfrage des OLAP Survey die Teilnehmer zu 40 unterschiedlichen OLAP-Produkten befragt (Pendse 2004), so war die Anzahl der ernsthaften Nennungen im BI Survey im Jahr 2014 mit 34 (BARC 2014, S. 14) nur unwesentlich geringer. Hinzu kommt eine Reihe von Werkzeugen in der Rubrik „Sonstige“. Viele der sogenannten Plattformanbieter haben es nicht geschafft, die Anzahl der selbst entwickelten oder hinzu gekauften Produkte zu reduzieren, sondern haben aufgrund neuer Trends die Anzahl unterschiedlicher Produkte sogar noch erhöht. Dies wirkt sich leider nicht immer positiv für den Kunden aus. Anwendungsunternehmen, die ihr Softwareportfolio strate-

gisch ausrichten möchten, sehen sich einerseits immer größer werdenden Anbietern mit nicht konsolidierten Softwareangeboten gegenüber. Im Extremfall werden von diesen Anbietern sogar von heute auf morgen strategisch sicher geglaubte Produktlinien abgekündigt. Auf der anderen Seite bieten traditionelle aber auch neue BI-Anbieter spezialisierte Lösungen an, die ggf. nicht optimal ins Softwareportfolio passen, dafür aber eine hohe Anwenderzufriedenheit versprechen. Daneben gibt es seit einigen Jahren ernstzunehmende Produktalternativen aus dem Open-Source-Umfeld.

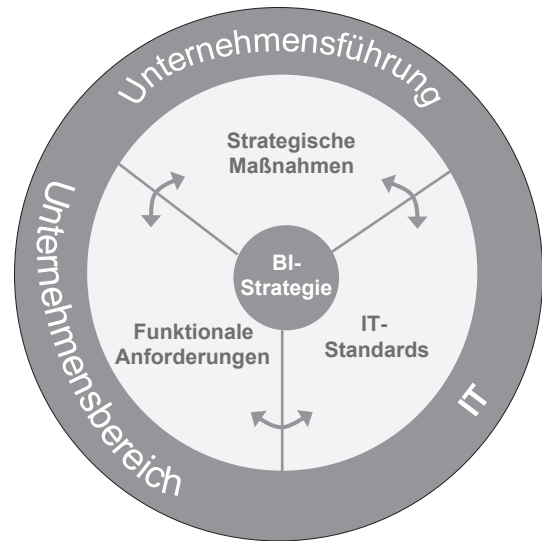
Vor diesem Hintergrund sollte man die Portfolioauswahl objektiv und nachvollziehbar gestalten. Grundlage für die Erstellung eines Softwareportfolios im Rahmen einer BI-Strategie sind die fachlich, technisch und architektonisch priorisierten Anforderungen, die in Form von Kriterienkatalogen zusammengestellt werden. Das Softwareportfolio sollte nach der Art der unterstützten Funktionalität (z. B. Planung, Analyse, Berichtswesen, Datenintegration) und der Tragweite (strategisch, taktisch, operativ) bestimmt werden. Einzelne, für das Unternehmen relevante Themenbereiche (z. B. Unternehmensreporting oder Vertriebsplanung) werden in das Portfolio eingeordnet und die bereits im Unternehmen eingesetzte Software entsprechend zugewiesen (Ist). Dabei sollte auch der Lebenszyklus für den Themenbereich im Unternehmenskontext berücksichtigt werden. Danach wird das Plan-Portfolio auf Basis der im Rahmen der Strategiefindung genannten Anforderungen aufgestellt. Das Ziel ist, für jeden Themenbereich ein führendes Softwarewerkzeug zu ermitteln oder sogar eine dominierende BI-Suite über alle Themenbereiche zu bestimmen. Die Anzahl der eingesetzten Werkzeuge sollte dabei systematisch reduziert werden. Ein Fahrplan bestimmt die zeitliche Abfolge, um den Plan zu erreichen.

Eine häufig angewandte Strategie zur Vereinheitlichung des BI-Softwareportfolios ist, alle BI-Werkzeuge von demselben Hersteller zu beziehen. Wird allerdings bei einer Vereinheitlichung kein Wert auf die Erhaltung einer größtmöglichen Funktionalität gelegt, sondern steht die Herstellerbindung im Vordergrund, so kann dies zu schwerwiegenden Akzeptanzproblemen führen.

3.4.5 Spannungsfeld erkennen

Bei der Entwicklung von unternehmensweiten BI-Strategien müssen in der Praxis immer wieder Hürden genommen und Widerstände überwunden werden. Zwar ist den meisten Beteiligten bewusst, dass eine effiziente Unternehmenssteuerung nur auf Basis einheitlicher und konsistenter Informationen möglich ist, dennoch stößt man immer wieder auf divergierende Bereichsinteressen, die eine gemeinsame Lösung erschweren. Die Erkenntnis, dass die verschiedenen Anwendungen durch die sie verbindenden Geschäftsprozesse zusammenhängend zu betrachten sind, erlangen manche Unternehmensvertreter erst im Rahmen von Strategiewerkshops. Hier liegt es am Projektteam, die Zusammenhänge ausreichend zu verdeutlichen und alle Prozessbeteiligten an einen Tisch zu holen.

Abb. 3.8 BI-Strategie im Spannungsfeld



Ein Thema, das in BI-Strategieprojekten häufig genannt, im betrieblichen Alltag allerdings gerne vernachlässigt wird, ist die mangelnde Vereinheitlichung von Stammdaten. Die Verantwortlichen eines operativen Systems pflegen oft die Stammdaten nur so gut, wie sie sie für ihre tägliche Arbeit benötigen. Abhängigkeiten zu Folgesystemen wie Data Warehouses sind entweder unbekannt oder ihr Stellenwert nicht nachvollziehbar. In der Strategieentwicklung müssen solche Probleme aufgedeckt und Maßnahmen beschrieben werden, um sie zu lösen. Die Lösungen liegen dabei nicht nur im technischen sondern auch im fachlichen oder sogar im zwischenmenschlichen Bereich.

In Abb. 3.8 wird das Spannungsfeld, in der sich eine BI-Strategie bewegt, vereinfacht dargestellt. Die Unternehmensführung gibt die strategischen Maßnahmen und Ziele vor, die durch eine BI-Strategie erreicht werden sollen. Die Unternehmensbereiche wiederum sind vor allem an einer individuellen, ihre Geschäftsprozesse optimierenden Funktionalität interessiert, die sie im Rahmen der Strategiefindung nennen. Die IT legt demgegenüber Wert auf die Standardisierung von Prozessen und Systemen. Tabelle 3.3 gibt exemplarisch eine genauere Darstellung von häufig genannten Anforderungen wieder.

Im Rahmen des Strategieprojektes müssen diese Anforderungen in Einklang gebracht werden. Voraussetzung für den Erfolg einer unternehmensweiten BI-Strategie ist, dass diese durchgehend kommuniziert, umgesetzt und damit gelebt wird. Sie muss den Anforderungen von Unternehmensführung/Holding und Unternehmensteilen/-bereichen gerecht werden. Nur dann, wenn kein oder nur noch geringer Bedarf nach individuellen Lösungen für die verschiedenen Teilbereiche existiert, macht eine gemeinsame BI-Strategie wirklich Sinn.

Tab. 3.3 Beispiele für Anforderungen im Rahmen einer BI-Strategie

Unternehmensführung
<ul style="list-style-type: none"> • Bessere Qualität der angebotenen Informationen • Erschöpfende Betrachtung des Unternehmens aus allen Perspektiven • Verbesserung der Kosten-/Nutzenrelation für Business Intelligence
Unternehmensteile/-bereiche
<ul style="list-style-type: none"> • Neue Metriken und Entscheidungsobjekte • Erweiterung der Funktionalität für Analyse und Reporting (auch beeinflusst durch Trendthemen) • Schnellere und genauere Verfügbarkeit von Informationen
Informationstechnologie
<ul style="list-style-type: none"> • Reduzierung des Umfangs der eingesetzten Softwareprodukte • Konsolidierung von Data Marts und Data Warehouses • Standardisierung der Prozesse

3.5 Fazit

Das Ergebnis einer BI-Strategie ist ein abgestimmtes Konzept für die Informationsversorgung im Unternehmen, das die strategische Ausrichtung beschreibt. Das Konzept legt unternehmensweite Standards fest und ist Basis für die BI-Roadmap, die ein durchgehend priorisiertes Projektportfolio mit einem Fahrplan für die kommenden Jahre enthält. Eine BI-Strategie ist jedoch keinesfalls statisch. Vielmehr ist sie analog zur Unternehmensstrategie kontinuierlich Veränderungen und Anpassungen unterworfen. Daher sollte eine BI-Strategie von vornherein als kontinuierlicher Prozess verstanden werden, der immer wieder Rückwirkung auf Masterplan und Roadmap hat. Das Unternehmen verschafft sich durch das Aufsetzen einer BI-Strategie Raum, um sich zukunftsorientierten Projekten zuzuwenden.

Hinweise für eine erfolgreiche BI-Strategie

- Leiten Sie die BI-Strategie klar aus der Unternehmensstrategie ab.
- Schaffen Sie eine durchgängige Priorisierung aller Anforderungen.
- Erstellen Sie eine umfassende Übersicht aller Ist-Kosten für Business Intelligence. Berücksichtigen Sie dabei auch Aufwände der Fachabteilungen.
- Verlangen Sie, dass der Nutzen jeder Information zumindest qualitativ definiert wird.
- Stimmen Sie ein Software-Portfolio ab, das im Unternehmen verpflichtend einzuhalten ist.
- Wählen Sie eine BI-Architektur, die integriert ist, aber auch Luft für außerordentliche Lösungen lässt, wie neue Big-Data-Ansätze.

- Schaffen Sie eine Betreiberorganisation, die fachliche Nutzer und IT eng miteinander verzahnt.
- Definieren Sie Ihre BI-Prozesse und leiten Sie sich eine passende BI-Governance für Ihr Unternehmen her.
- Das Outsourcing von BI-Aktivitäten ist äußerst anspruchsvoll. Werden Sie sich erst Ihrer BI-Prozesse voll bewusst, ehe Sie in diese Richtung vorgehen.

Literatur

- Adelman, S., Moss, L.T.: Data warehouse project management, 4. Aufl., Pearson Addison Wesley, Boston (2004)
- BARC: The BI survey 14. BARC, Würzburg (2014)
- Fütting, U.C.: Projektmanagement und _controlling von Data Warehouse-Projekten. In: Mucksch, H., Behme, W. (Hrsg.): Das Data Warehouse-Konzept, 4. Aufl. Gabler, Wiesbaden, S. 269–289 (2000)
- Gansor, T., Totok, A., Stock, S.: Von der Strategie zum Business Intelligence Competency Center. dpunkt, München (2010)
- Hoffmann, O.: Performance Management – Systeme und Implementierungsansätze, 3. Aufl., Haupt Verlag, Bern (2002)
- Miller, G.J., Bräutigam, D., Gerlach, S.: Business intelligence competency centers – a team approach to maximizing competitive advantage. Wiley, New Jersey (USA) (2006)
- Pendse, N.: The OLAP Survey 4 Preview. <http://www.olapreport.com/> (2004). Zugegriffen: 1.8.2015
- Philippi, J.: Outsourcing und Offshoring von Business Intelligence-Lösungen – Empirische Studien und Praxiserfahrung, in: Schelp, J., Winter, R. (Hrsg.): Auf dem Weg zur Integration Factory – Proceedings der DW2004 – Data Warehousing und EAI. Springer, Heidelberg S. 73–106 (2004)
- Totok, A.: Modellierung von OLAP- und Data-Warehouse-Systemen. DUV, Wiesbaden (2000)
- Vierkorn, S.: BI-Strategie – Facetten und Herausforderungen, Vortrag, 5. Europäische TDWI-Konferenz. Sigs Datacom, München (2008)
- Watson, H.J., Abraham, D., et al.: Data warehousing ROI: Justifying and assessing a data warehouse, in: Bus. Intell. J. 6–17 (2004)

Die nächste Evolutionsstufe von AIS: Big Data

4

Erweiterung klassischer BI-Architekturen mit neuen
Big Data Technologien

Carsten Dittmar

Inhaltsverzeichnis

4.1	Big Data und Business Intelligence	56
4.2	Erweiterung der analytischen Anwendungsfelder durch Big Data	58
4.3	Überblick über Technologien für Big Data	61
4.4	Erweiterung klassischer BI-Architekturen durch Big Data	62
4.5	Fazit	64
	Literatur	65

Zusammenfassung

Der Begriff „Big Data“ ist derzeit als einer der sogenannten Megatrends in aller Munde. In der breiten öffentlichen Diskussion zu Big Data löst die damit verbundene Idee, aus der Vielzahl und Vielfalt der verfügbaren Daten schnell werthaltige Informationsschätze zu heben, sehr viele positive aber auch negative Reaktionen aus: Während sich auf der einen Seite ein schier unbegrenztes Spektrum an Anwendungsfeldern mit zum Teil komplett neuen Geschäftsmodellen eröffnet, wird auf der anderen Seite eindringlich auf potenzielle Gefahren hingewiesen, die sich aus der breiten Verknüpfung von Dateninhalten und der damit einhergehenden, vollständigen Transparenz über das Verhalten und Vorlieben Einzelner ergeben.

C. Dittmar (✉)
Business Development Big Data & Analytics,
NTT DATA Deutschland GmbH, Köln, Deutschland
E-Mail: carsten.dittmar@tdwi.eu

Im vorliegenden Beitrag wird Big Data aus dem Blickwinkel der Analyseorientierten Informationssysteme eingeordnet und anhand der wesentlichen Anwendungsfelder und Technologien konkretisiert. Anschließend wird der resultierende Einfluss von Big Data auf etablierte Architekturparadigmen skizziert.

4.1 Big Data und Business Intelligence

Die aktuelle Diskussion um IT relevante Themen wird derzeit sehr intensiv durch den Begriff „Big Data“ geprägt. Der Begriff Big Data lässt sich dabei nicht nur in den traditionellen IT affinen Medien und Kommunikationskanälen finden, sondern als Megatrend hat es Big Data geschafft, auch in der allgemeinen Tagespresse aufgegriffen zu werden. Big Data wird häufig stellvertretend für das Phänomen der zunehmenden Digitalisierung aller Lebensbereiche angeführt, die tiefgreifend auch Struktur und Wertschöpfung der Unternehmen verändert und Information zu einer dominierenden Ressource kürt.

In frühen Veröffentlichungen wird das Phänomen Big Data häufig anhand einer Negativabgrenzung anhand der Leistungsfähigkeit klassischer IT vorgenommen. Mit Big Data werden Daten beschrieben, die die Möglichkeiten der klassischen Datenhaltung, -verarbeitung und -analyse auf konventioneller Hardware übersteigern (Merv 2011; Plattner 2013).

Mittlerweile wird Big Data genauer anhand von drei charakteristischen Eigenschaften beschrieben. Die erste Eigenschaft referenziert auf die extrem gestiegene Menge auswertungsrelevanter Daten. Das exponentiell wachsende Datenvolumen entsteht u. a. durch den zunehmenden Einsatz von permanent datenerzeugender Sensorik, die zunehmende Protokollierung von IT-Nutzung durch Logs sowie dem ständigen Zuwachs an Transaktionsdaten. Die steigende Informationsintensität in der Wertschöpfung, aber auch in den Produkten/Dienstleistungen treibt die Menge an auswertungsrelevanten Daten aus Sicht der Unternehmen deutlich stärker als die häufig kolportierte Datenflut durch jegliche Form von (Multimedia-) Content in Social Media Kanälen.

Auch wenn der zusammengesetzte Begriff Big Data den Eindruck suggeriert, dass ausschließlich das hohe Datenvolumen Big Data kennzeichnet, stellt sich in Zeiten des Wandels und mannigfaltiger Anforderungen für Unternehmen auch die Frage nach der Komplexität und Schnelligkeit der Verknüpfung, Integration und Analyse von Daten unterschiedlichster Struktur und Herkunft. Als weitere Eigenschaften von Big Data werden daher neben dem immensen Datenvolumen (Volume) die erhebliche Vielfalt an Datenformaten (Variety) als auch die Geschwindigkeit (Velocity), mit der neue Daten entstehen und auszuwerten sind, angeführt. In der folgenden Tab. 4.1 sind die drei charakterisierenden „Vs“ nochmal zusammenfassend dargestellt.

Unter Berücksichtigung der skizzierten Eigenschaften bezeichnet Big Data dann insofern Methoden und Technologien für die hochskalierbare Integration, Speicherung und Analyse polystrukturierter Daten. Gemäß der Definition wird deutlich, dass es sich bei Big Data insofern nicht ausschließlich um rein technische Fragestellungen handelt, sondern auch methodische Aspekte eine wichtige Rolle spielen. Zudem bezieht sich der

Tab. 4.1 Charakterisierende Eigenschaften von Big Data

Eigenschaft	Erläuterung
<i>Datenvolumen (Volume)</i>	Unternehmen verfügen über gigantische Datenberge, die von einigen Terabytes bis hin zu Größenordnungen von Petabytes führen, da IT-gestützte Geschäftsprozesse permanent Daten produzieren. Neben den internen Daten stehen eine Vielzahl von externen Datenquellen und Devices/Maschinen zur Verfügung, die einen konstanten Datenstrom erzeugen
<i>Datenvielfalt (Variety)</i>	Mit der Vielzahl an Datenquellen wächst auch die Vielfalt von Datenformaten. Aus immer mehr Quellen liegen Daten in unterschiedlicher Art vor, die sich grob in unstrukturierte, semistrukturierte und strukturierte Daten gruppieren lassen. Die technologische Auswertbarkeit von diesen polystrukturierten Daten hat dabei extrem zugenommen
<i>Daten-geschwindigkeit (Velocity)</i>	Die Prozessintegration bedingt, dass Unternehmen die Datensammlung, -integration und -analyse immer schneller – nicht selten in Echtzeit – durchführen müssen, um relevante Erkenntnisse zu erzielen bzw. Aktionen einzuleiten. Daneben unterliegen auch Datenstrukturen, Datenquellen und Schnittstellen einer immer höheren Änderungsdynamik

In der Diskussion treten gelegentlich auch noch weitere charakterisierende „V’s“ auf, z.B. Variability, Value etc. Diese werden allerdings eher aus Marketing Gründen eingeführt und liefern keinen weiteren Beitrag zur Begriffsabgrenzung von Big Data.

Skalierungsanspruch nicht ausschließlich auf ein hohes Datenvolumen (Volume), sondern auch auf eine hohe zeitliche Skalierung hinsichtlich hoher Datenaktualität und hoher Performanceansprüche bei der Durchführung komplexer Analysen (Velocity) auf Basis von Daten, die in vielfältiger Strukturierung vorliegen können (Variety). Weiterhin werden sämtliche Ebenen einer Datenverarbeitungsarchitektur von der Integration über die Speicherung bis hin zur Analyse erfasst (Eaton et al. 2012; Business Application Research Center (BARC) 2014).

Die Integration, Speicherung und Analyse von Daten aus unterschiedlichen Quellen mit dem Ziel ein tieferes Geschäftsverständnis abzuleiten, ist dabei keinesfalls neu. Mit Business Intelligence (BI) hat sich längst ein Konzept im Umfeld der entscheidungsunterstützenden Systeme in Theorie und Praxis etabliert, hinter dem sich weniger eine einzelne Lösungsmethode oder technische Systemklasse als vielmehr eine begriffliche Klammer für eine Vielzahl von unterschiedlichen Ansätzen zur Analyse und Auswertung von Geschäftsprozessen und zum Verstehen relevanter Wirkungszusammenhänge in Unternehmen verbirgt (Gluchowski et al. 2008). Daher haben die meisten Unternehmen eine langjährige Historie, dispositive Daten aus unterschiedlichen IT-Systemen zusammenzuführen und darauf aufbauend Reporting- und Analyse-Anwendungen zu erstellen. Die begriffliche Einordnung von BI gibt dabei keinerlei Einschränkungen hinsichtlich der Datengewinnung auf Anzahl, Lokalisierung oder Art der Datenquellen noch hinsichtlich der Datenauswertung auf den Einsatzzweck vor.

Aus dem Blickwinkel der Begrifflichkeit ergibt sich unmittelbar, dass Big Data ein spezielles BI-Anwendungsfeld ist, das den Fokus auf die Skalierbarkeit hinsichtlich der 3 V’s

legt, wobei exakte Grenzen nicht eindeutig quantifizierbar und damit unternehmensindividuell sind. Ein Anwendungsfeld wird heute in der Regel zu Big Data gezählt, wenn mind. zwei der drei V's zutreffen. Big Data steht damit nicht in Konkurrenz zu klassischen BI-Methoden und BI-Technologien, sondern stellt eher eine Ergänzung bzw. Erweiterung dar.

Auch wenn begrifflich mit BI ein breiter Anwendungsbereich adressiert wird, zeigt sich in der Praxis häufig ein ernüchterndes Bild. Hier wird häufig de facto unter BI das klassische vergangenheitsorientierte Reporting verstanden, das auf den strukturierten Daten aus dem Data Warehouse aufsetzt, welches wiederum in der Regel von internen Transaktionssystemen gespeist wird (Baars et al. 2014). Aus diesem Verständnis ist erklärbar, dass in der Praxis häufig Big Data nicht unmittelbar als Unterbegriff von BI, sondern vielmehr als komplett neuartige Themenstellung aufgefasst wird und bei initialen Big Data Projekten häufig die bestehende BI-Organisation gar nicht oder nur rudimentär eingebunden wird.

Der Hype um Big Data stellt insofern aus Sicht von BI-Professionals Fluch und Segen zugleich dar. Auf der einen Seite führt die breite Diskussion dazu, dass auch außerhalb der IT über innovative Ideen und neue Geschäftsmodelle nachgedacht werden, die auf dem Rohstoff des 21. Jahrhunderts – den Daten – aufsetzen. Dadurch wird die Bedeutung des Assets Daten neu bewertet. Daten werden zu einem wichtigen Produktionsfaktor und sollen Geschäftsprozesse und Geschäftsmodelle treiben und verändern, die „information-driven company“ wird gefordert. Zudem werden analog zum einschränkenden BI-Verständnis in der Praxis bisher vernachlässigte BI-Themen wieder neu belebt (z. B. Data Mining, Visualisierung, Endbenutzer-Autonomie). Diesen positiven Aspekten steht gegenüber, dass viele Diskussionen um Big Data losgelöst von der bestehenden BI-Erfahrung geführt werden. Schließlich sind die Herausforderungen in der ‚Small Data‘ Welt wie mangelnde Datenqualität, inadäquate Organisationsstrukturen, unzureichende strategische BI-Ausrichtung noch längst nicht gelöst und sind auch für Big Data weiterhin relevant (Steria Mummert Consulting GmbH 2013). Zudem ist Big Data derzeit auf dem Höhepunkt der übertriebenen Erwartungshaltung und wird häufig mit anderen Aspekten (z. B. Cyber Crime, Datenschutz) vermengt. Darüber hinaus klammert die bisher eher technologiegetriebene Debatte um Big Data die Frage nach dem Business Nutzen regelmäßig aus und technikzentrierte Ansätze („Habe Lösung, suche Problem!“) waren schon im BI-Umfeld nicht erfolgreich.

4.2 Erweiterung der analytischen Anwendungsfelder durch Big Data

Im Moment sind die meisten Unternehmen im Themenfeld Big Data noch eher in einer Orientierungsphase. Für die überwiegende Mehrheit der Unternehmen in Europa hat das produktive Big-Data-Zeitalter daher noch nicht begonnen (Gluchowski 2014). Big Data fungiert allerdings als Initiator für das Überdenken von BI in den Dimensionen Fachlichkeit, Organisation und Technik: Big Data treibt in Unternehmen die Abkehr von alten Architekturparadigmen, führt zu einer Business Value Diskussion jenseits von Standardreports und Enterprise Data Warehouse und damit zu einer neuen Arbeitsaufteilung zwischen Fachbereich und IT.

Im Mittelpunkt der Diskussion zwischen Fachbereich und IT zu Big Data steht im Moment natürlich die Frage nach den industriespezifischen Anwendungsfällen (Use Cases). Hier spannt sich ein weites Feld auf, das sich von internen Analysegegenständen aus dem Unternehmen (Prozesse, Planungs-/Prognoseergebnisse etc.) bis hin zu externen Analysegegenständen auf dem jeweiligen Markt (Produkte, Wettbewerber, Kunden, Lieferanten) erstreckt. Durch neue Datenquellen und deren Verknüpfungen durch industrieübergreifende Kooperationen sind zudem komplett neuartige Geschäftsmodelle in neuen Märkten denkbar.

Aus der Kombination der beiden Dimensionen Analysegegenstand und Geschäftsmodellbezug lassen sich in einer Matrix grob vier Klassen von Big Data Anwendungsfällen unterscheiden (vgl. Abb. 4.1). Einzelne Big Data Use Cases lassen sich innerhalb dieser Matrix problemlos einordnen, auch wenn die Übergänge der einzelnen Klassen fließend sind.

In einer ersten Klasse sind Anwendungsfälle einzuordnen, die effizientere Prozesse und Steuerung durch Big Data versprechen. In dieser Klasse wird das grundsätzlich bestehende Geschäftsmodell nicht verändert, allerdings werden unternehmensinterne Abläufe durch Big Data Methoden und Technologien optimiert. Ein häufiger Anwendungsfall ist die Erhöhung der Prognosesicherheit für den zukünftigen Warenabverkauf, Ersatzteilbedarf, Energieverbrauch oder Maschinenverschleiß. In Distribution und Logistik geht es um nachhaltige Prozessoptimierung auf dem Wege einer stärkeren Vernetzung von Fahrzeugen mit der Außenwelt. Z. B. können durch Ortungs- und Verkehrsaufkommensinformationen Disponenten zeitnah Transporte planen, gegebenenfalls Routen und Beladung ändern, Wartungskosten und Stillstandzeiten minimieren. Selbst bei einfachen Prozessen wie der Postbearbeitung können lernende Systeme durch automatisierte Abläufe für mehr Effizienz im Routing sorgen.

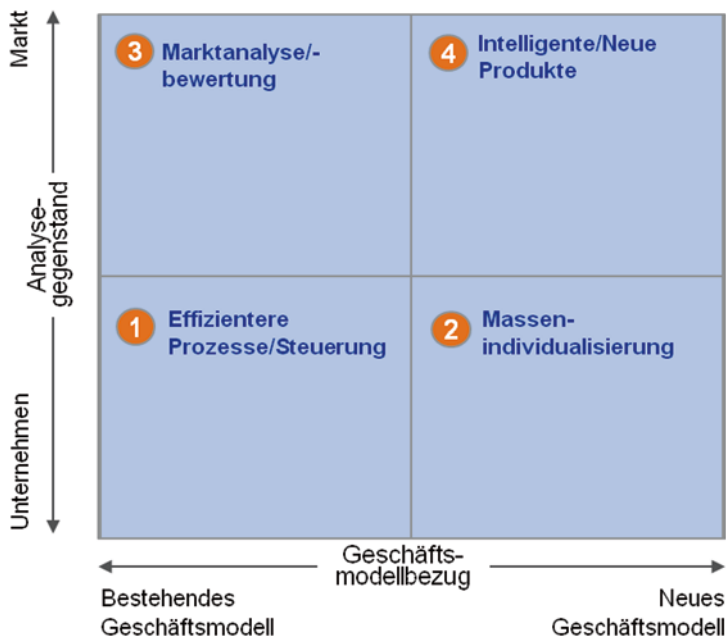


Abb. 4.1 Anwendungcluster für Big Data

In einer zweiten Klasse sind Anwendungsfälle zu beschreiben, die eine Massenindividualisierung durch Big Data erlauben. In Sinne eines fließenden Übergangs zur ersten Klasse wird hier Big Data häufig zur Optimierung kundenbezogener Prozesse genutzt, indem z. B. individuelle Bedürfnisse oder der jeweilige aktuelle Kontext der Kunden noch stärker im Rahmen der Kundenansprache berücksichtigt wird. Die Analyse des Kundenverhaltens durch Einbeziehung einer Vielzahl auch unternehmensexternen Datenquellen und Analyse der Social Influencer Position in den Social Networks, um daraus während des Kundenkontakts individualisierte Next-Best-Offering abzuleiten, wäre ein entsprechender Big Data Anwendungsfall. Zudem fällt in diese Kategorie die Vielzahl von kontextbasierten Diensten basierend auf Mobilfunk- & Telematik-Systemen, um Verhalten und Aufenthaltsort des Kunden bei der Kundenansprache zu nutzen. Auf der anderen Seite des Spektrums der Massenindividualisierung kann sich auch eine komplette Individualisierung eines Produktes oder einer Dienstleistung ergeben, so dass daraus im Extremfall neue Geschäftsmodelle entstehen. Individualisierte Mobilitätsprodukte oder über die persönliche Nutzung und Verhaltensweise prämierte Versicherungsprodukte (pay-as-you-use) sind hier als Beispiele zu nennen.

In einer dritten Klasse sind Anwendungsfälle einzuordnen, die aus der Vielzahl der Einschätzung und Meinungen der unterschiedlichen Marktteilnehmer eine aggregierte Gesamteinschätzung im Sinne einer Marktanalyse und -bewertung ableiten lassen, dabei den Bezug allerdings im etablierten Geschäftsmodell halten. Entsprechende Big Data Anwendungsfälle betreffen z. B. die Erhebung und Analyse der Marken- oder Produktwahrnehmung oder Markttrends und werden daher häufig auch als effizientere Form der Marktforschung klassifiziert. Als potenzielle Datenquellen kommen hier insbesondere Internetinhalte aus den Social Media und Bewertungsportale in Form von unstrukturierten Daten zum Einsatz.

Schließlich sind in einer vierten Klasse Anwendungsfälle zu sehen, die innovative/selbstregulierende Produkte durch Big Data versprechen und somit neue Geschäftsmodelle adressieren, indem bis dato isolierte Datenquellen ggf. industrieübergreifend zusammengeführt werden und mit dem resultierenden Informationsprodukt neue Märkte bzw. Marktsegmente adressiert werden können. Häufig bilden sich neue Allianzen zwischen bisher isolierten Branchen, wie der Telekommunikationsindustrie, der Automobilindustrie und der Dienstleistungsbranche. Viele Produkte und Anlagen verfügen über eine Vielzahl von Sensoren, die kontinuierlich über unzählige interne und externe Zustände (z. B. Wartungszustand, Temperatur, Position, Sonneneinstrahlung, Abstände) Auskunft geben, und es dadurch erlauben, z. B. selbstregulierende Häuser und autonom fahrende Fahrzeuge zu entwickeln. In Zukunft werden immer mehr Produkte und vernetzte Maschinen mit eigener Big Data Intelligenz ausgestattet sein, um selbst auf Einflüsse autonom zu reagieren bzw. entsprechende Reaktionen vorschlagen.

Das Spektrum von Big Data Use Cases ist zusammenfassend formuliert sehr vielfältig und branchenabhängig. Branchenübergreifende „Killer-Anwendungen“ sind nicht in Sicht. Big Data verspricht große Chancen, aber auch große Herausforderungen für Unternehmen. Die öffentliche Diskussion rund um das Thema Big Data fokussiert sich im Mo-

ment sehr auf Fragen des Datenschutzes aufgrund der scheinbaren Transparenz über das Handeln und Verhalten des Einzelnen. Es zeigt sich allerdings, dass Endkunden damit einverstanden sind, dass Unternehmen ihre Daten sammeln und analysieren, sofern sie unmittelbar einen klaren Mehrwert für sich in Form eines Plus an Sicherheit, Information, Convenience oder eines finanziellen Vorteils sehen.

4.3 Überblick über Technologien für Big Data

Eine Vielzahl von technologischen Konzepten und Systemen werden heute mit dem Begriff Big Data verbunden. Nicht immer verbergen sich hinter entsprechenden Auflistungen ausschließlich nur neuartige Technologien, schließlich wird der vieldiskutierte Big Data Begriff vom Marketing gerne aufgegriffen und mit etablierten Technologien verbunden, so dass die Trennschärfe bei den Anwendern abnimmt (TDWI 2011). In einer ersten Kategorisierung können Technologien zur Erzeugung von Big Data (z. B. Sensorik, Geo-Tracking), zur Verarbeitung und –Integration von Big Data (z. B. Streaming, Virtualisierung), zur Speicherung von Big Data (z. B. Analytische Plattformen, Cloud) und zur Auswertung von Big Data (z. B. Visualisierung, Advanced Analytics) unterschieden werden. Für die nachfolgende Betrachtung werden die Technologiecluster selektiert, die maßgeblichen Einfluss auf die Gestaltung von klassischen BI Architekturen haben.

Technologien zum Streaming bzw. zur (Near-)Realtime Datenpopulation erweitern die gängige Batch-orientierte Datenpopulation um die Fähigkeit, Datenquellen mit kontinuierlichem Datenfluss mit niedriger Latenzzeit anbinden zu können. Hier werden Einsatzszenarien unterstützt, in denen die Quelldaten nicht persistent vorgehalten werden (z. B. Daten aus Sensoren), sondern diese nur kontinuierlich hinsichtlich bestimmter auffälliger Datenkonstellation untersucht werden müssen, um definierte Aktionen auszulösen.

Analytische Plattformen sind Datenhaltungssysteme, die sich aufgrund ihrer Architektur und Funktionen besonders gut für Datenanalysen eignen, was vor allem durch Abfrage-Eigenschaften und durch zusätzliche analytische Funktionen bestimmt wird. Analytische Plattformen zeichnen sich u.a. durch zwei wesentliche Charakteristika aus. Zum einen wird die Ausführung von analytischen Anfragen so dicht wie möglich an die Datenhaltung gebracht und zum anderen ist die Hard- und Softwarearchitektur auf die spezifischen Anforderungsprofile analytischer Abfragen optimiert. Die Optimierung wird durch die Anwendung der Konzepte zur Reduktion von Festplattenzugriffen (In-Memory, physische Speicherungsstrategie, Indexierung, Komprimierung), Parallelisierung der Anfragen über eine möglichst große Anzahl an Knoten und die Verwendung von aufeinander abgestimmten Systembausteinen erreicht. Besondere Popularität und Verbreitung unter den analytischen Plattformen hat die Open Source Lösung Hadoop und ihre kommerziellen Distributionen erreicht. Wesentliche Charakteristika von Hadoop sind ein verteiltes Dateisystem und ein Computing Framework, das einen Rahmen bietet, um eine Vielzahl von darauf aufsetzenden Lösungen, beispielsweise Abfragesprachen (z. B. Hive), Suchsysteme (z. B. Solr) oder Machine Learning Umgebungen (z. B. Mahout), in einem gemeinsamen, hoch-

skalierbaren Cluster ablaufen zu lassen, sowie Daten in großer Menge auf kostengünstiger Standardhardware zu speichern. NoSQL-Datenbanken (z. B. Graphen-, dokumentenorientierte- oder Key-Value-Datenbanken), deren Implementierung teilweise auch auf der Hadoop Plattform basieren, werden ebenfalls unter Analytischen Plattformen subsumiert.

Bei der Menge und Heterogenität von Big Data sind zunehmend Technologien im Einsatz, die die bisher verfolgte Maxime der physischen Integration sämtlicher relevanter Quellen in einer singulären dispositiven Umgebung zugunsten einer Integration über eine logische, auf Metadaten basierende Zwischenschicht verzichtet (Logische Integration). Somit werden Abfragen über bisher separate oder gar nicht für Analysen verfügbare Systeme ermöglicht. Entsprechende Virtualisierungslösungen ermöglichen den direkten Zugriff auf unterschiedliche Datenquellen, um eine integrierte Datenanalyse zu erlauben.

Anwender fordern heutzutage ohne Unterstützung des IT-Bereichs selbstständig Analysen auch auf Datenbestände durchführen zu können, die die Big Data Charakteristika aufweisen. Entsprechende Self-Service-BI-Werkzeuge werden heute häufig ebenfalls zu den Big Data Technologien gezählt, da u.a. oftmals dazu In-Memory-Technologien eingesetzt werden, die auf dem Desktop-Rechner des Anwenders laufen. Der Funktionsumfang von Self-Service-BI-Werkzeugen gerade im Bereich Visual Analytics zeigt erkennbare Überschneidungen zum Funktionsumfang mit klassischen BI-Suiten. Insofern hat sich ein interessanter Wettbewerb zwischen den erfolgreichen Herstellern von Self-Service-BI-Werkzeugen und klassischen BI-Suiten entwickelt, die eine zunehmende Ausrichtung aller am Markt befindlicher BI-Werkzeuge auf Analytics als Ergänzung der klassischen Fokussierung auf Reporting bedingt.

4.4 Erweiterung klassischer BI-Architekturen durch Big Data

Aus theoretischer Sicht war die „richtige“ BI-Architektur lange klar definiert. Ausgehend vom logischen Architekturverständnis zergliedern die Standardlehrbücher den Datenintegrations- und -veredlungsprozess in der Regel in vier mandatorische Datenschichten (Data Acquisition Layer, (Core) Data Layer, Analytical Layer, Presentation Layer). Damit wird eine mehrfach redundante Speicherung der operativen Datenbestände aus den relevanten Quellen durch eine physische Datenintegration über verschiedene Ebenen gefordert. Der Standardansatz von W. H. Inmon geht dabei von einer singulären Hub&Spoke Architektur mit einem an der 3NF orientierten Core Warehouse aus. Aus diesem „Single Point of the Truth“ werden gewisse Datenextrakte in der Regel multidimensional in voneinander fachlich abgrenzbaren Data Marts gehalten, die dann wiederum den Presentation Layer mit seinem spezifischen Reporting und Analyse Werkzeugen versorgen. In historisch gewachsenen Landschaften führt idealtypisch die erwartete semantische Integration und der „One Size fits All“-Ansatz ggf. in Verbindung mit einer Technologiekonsolidierung bzw. -standardisierung zu einem kostenoptimalen und vor allem singulärem Enterprise Data Warehouse Approach. Ggf. werden entsprechende Architekturen noch durch einen Operational Data Store (ODS) als temporäre Datensinke auf feingranularer Ebene ergänzt (Inmon 1996; Gluchowski et al. 2008).

Die Einhaltung der skizzierten Paradigmen klassischer BI-Architekturen wird jedoch in der Realität immer schwieriger: So finden sich in der Praxis selten vollständig konsolidierte und zentralisierte Architekturansätze. Vielmehr sind einzelne Informationssilos und heterogene Architekturen in gewachsenen dispositiven Landschaften eher die Regel statt die Ausnahme. Die existierenden Architekturen zeichnen sich in Folge dessen durch eine hohe Komplexität aus, so dass mitunter die fachliche und/oder technische Transparenz über die Datenveredlungs- und –konsolidierungsprozesse zumindest partiell abhandengekommen ist. Zudem ist ein starkes Anwachsen der Schatten-BI in Form von sog. „Spread-Marts“ unter Umgehung von architekturellen Vorgaben zu beobachten (Dittmar 2013).

Der oben skizzierte Lösungsansatz eines Layer bezogenen Aufbaus der dispositiven Landschaft wird traditionell von den Mega-Vendoren bedient, die den gesamten Stack an notwendigen Technologien integriert anbieten. Aber gerade durch die neuen Big-Data-Technologien mit hohem Einfluss auf die klassische BI-Architektur wächst auf Seiten der Anwender erneut die Bereitschaft, in „Best of Breed“ Lösungsportfolien zu investieren. Der Einfluss der neuen Technologien und Konzepte wirkt sich dabei auf alle Schichten einer traditionellen BI-Architektur aus. Einige Big-Data-Technologien konkurrieren mit klassischen Lösungen, aber die Mehrzahl der Anwendungsszenarien verlangt – zumindest in der mittelfristigen Perspektive – eher einen additiven Einsatz zu bestehenden Lösungen.

Das Architekturelement Streaming/(Near)Realtime ergänzt die BI-Architektur um die Fähigkeit, Daten (Near-)Realtime zu empfangen und zu verarbeiten. Eine Schnittstelle für den Empfang von Streams oder Mini-Batches sowie ein auf diese Frequenz der Datenlieferung abgestimmter Datenintegrationsprozess und ein online analytischer Zugang für Nutzer sind die dazu minimal notwendigen Komponenten. Optional wird eine sog. Decision Engine für die automatisierte Entscheidungsableitung eingesetzt. Im Gegensatz zu klassischen Complex Event Processing Lösungen kann hier auch Bezug auf eine längere/vollständige Historie im DWH genommen werden. Bei der Integration von (Near-)Realtime Anwendungen in BI-Architekturen ist die Verwendung einer zunächst separaten Verarbeitung für den schnellen Zugriff und erst eine spätere Zusammenführung mit dem DWH Datenbestand verbreitet (Lambda-Architektur).

Gerade der Einsatz von Analytischen Plattformen in klassischen BI-Architekturen wird derzeit sehr intensiv diskutiert. Diese werden nicht nur als eine Ergänzung, sondern auch als eine alternative Persistenz-Basis für eine BI-Architektur evaluiert. Prominentester Anwendungsfall für die Ergänzung ist die Nutzung von Analytischen Plattformen für die kostengünstige Vorverarbeitung großer Datenmengen und/oder die Möglichkeit der Aufbereitung polystrukturierter Daten als Vorbereitung für die Integration in ein klassisches DWH mit strukturiertem Datenbestand. Anwendungsfall für die Substitution ist der Einsatz von Analytischen Plattformen in den Schichten (Core) Data Layer oder Analytical Layer, um hohen Anforderungen an Abfrageperformance bei gleichzeitig großer Datenmenge gerecht zu werden. Analytische Plattformen werden heute zudem sehr intensiv autark als Analytical Lab für die tiefgehende Analyse großer Datenmengen – auch polystrukturierter Daten – außerhalb des etablierten DWH verwendet.

Um möglichst schnell eine explorative Datenanalyse durch Zugriff auf eine breite Datenbasis unterschiedlichen Typs und unterschiedlicher Aktualität zu ermöglichen, wird

häufig in klassischen BI-Architekturen die Abstraktion zwischen Datenhaltungs- und Verarbeitung- zur Nutzungsschicht durch die Einführung eines zusätzlichen logischen Layers erreicht. Diese Komponente wird mit Virtualisierungs-Lösungen gebaut, die eine Transparenz über verschiedene Datenspeicherquellen und –technologien gegenüber der Nutzungsschicht erreicht. Dies steigert die Flexibilität und Aktionsgeschwindigkeit bei gleichzeitiger Senkung der Kosten für Umsetzung, Betrieb und Infrastruktur, sofern man alternativ von einer physischen Datenintegration ausgeht. Weiterhin wird dieses Konzept verwendet, um einen iterativen, transparenten Umbau der zugrundeliegenden Datenhaltungssysteme zu ermöglichen. Wichtige Voraussetzung bildet ein System-übergeordnetes Stamm- und Metadaten-Management.

Self Service BI erweitert die klassischen Funktionsklassen wie Standard Reporting und Ad-hoc Analyse des Presentation Layers. Die Möglichkeit der Bildung abgeschotteter Analyseräume (Sandboxing) durch den Nutzer kann im Frontend Tool entweder am lokalen Rechner des Nutzers oder im DWH ermöglicht werden. Die wesentliche Herausforderung bei der Öffnung der Reporting-Definitionen und der Möglichkeit der Integration von durch Nutzer eingebrachten Daten bleibt jedoch die Ausgestaltung der Governance, um einen konsistenten Rahmen durch organisatorische Instrumente aufrechtzuerhalten.

Die oben aufgezeigten Erweiterungen der traditionellen BI-Architektur lassen sich in einer zukunftsfähigen Gesamtarchitektur integrieren, die sich als Analytisches Ökosystem beschreiben lassen. In diesem existieren durchaus mehrere analytische Datenpools nebeneinander, da unterschiedliche Granularitäten, Datenquellen oder Analyseziele nicht dogmatisch physisch in einem Data Warehouse integriert werden müssen. Sofern die logische Transparenz besteht, in welchem analytischen Datenpool welche Daten in welcher Form abgelegt sind, können diese bei konkretem Bedarf durchaus auch nur virtuell und temporär zusammengeführt werden. Das Analytische Ökosystem bietet Services zu Datenspeicherung, -veredelung, -distribution, -analyse und –zugriff an. Allerdings ist es nicht für jeden Anwendungsfall (Use Case) sinnvoll, sämtliche Services über alle zur Verfügung gestellten Schichten zu durchlaufen.

4.5 Fazit

Big Data verspricht große Chancen, aber auch große Herausforderungen für Unternehmen. Die Notwendigkeit, schnell und flexibel auf neue Marktsituationen zu reagieren, wird in Zukunft für den Unternehmenserfolg entscheidender sein denn je. Erfolgreich werden dabei nur die Unternehmen sein, welche die Schätze, die sich in den Datenbergen ihres Unternehmen verbergen, zu bergen wissen.

Big Data ist von seinen Konzepten und Ideen weitgehend kongruent zu BI, setzt aber den Fokus auf die Skalierbarkeit hinsichtlich der 3 V's. Auf der einen Seite ist Big Data Treiber für eine Bandbreite neuer innovativer Use Cases, auf der anderen Seite werden alternative und additive Technologien und Konzepte aus Sicht von BI angeboten. In Folge entwickelt sich ein neues Verständnis einer ganzheitlichen Referenzarchitektur, die aus

flexiblen Elementen besteht, die bei definierten Freiheitsgraden miteinander agieren und sich so von starren und zentralistischen Architekturen lösen.

Nichts desto trotz ist der Weg zu einem Analytischen Ökosystem lang und in Abhängigkeit der verfolgten Use Cases unternehmensindividuell auszugestalten. Ein geeigneter erster Schritt zur Bestimmung des für den einzelnen Unternehmenskontext notwendigen Ausschnitts einer Gesamtarchitektur ist die Abbildung der als relevant selektierten Use Cases zur Identifikation der benötigten Komponenten und deren funktionales Zusammenspiel. Eine geeignete Grundlage bildet die Referenzarchitektur eines gesamthaften Analytischen Ökosystems.

Der Identifikation der Untermenge an relevanten Komponenten nachgelagert ist die Besetzung mit konkreten Lösungen. In Anbetracht der gestiegenen Vielzahl und Diversität an Lösungen kommt einem „Proof of Concept“ bei diesem Prozess verstärkt Bedeutung zu. Basierend auf Architekturdefinition und konkretisierter Bebauung in Form einer unternehmensindividuellen Referenzarchitektur ist ein individueller Transformationspfad zum Erreichen des Zielbilds in pragmatischen Schritten abzuleiten.

Literatur

- Business Application Research Center (BARC): Big data analytics. <http://barc.de/docs/survey-big-data-analytics> (2014). Zugegriffen: 1. Nov. 2014
- Baars, H., Funke, K., Müller, P.A., Olbrich, S.: Big Data als Katalysator für die Business Intelligence – Das Beispiel der informa Solutions GmbH. *HMD* **51**(298), 436–446 (2014)
- Dittmar, C.: Abschied von der einen Quelle der Wahrheit – Die aktuelle biMA-Studie zeigt: Viele Ideale sind utopisch. *BI-Spektrum* **8**(4), 22–24 (2013)
- Eaton, C., Deroos, D., Deutsch, T., Lapis, G., Zikopoulos, P.: Understanding big data, analytics for enterprise class hadoop and streaming data. McGraw-Hill Education, New York (2012)
- Gluchowski, P.: Empirische Ergebnisse zu Big Data. *HMD* **51**(298), 401–411 (2014)
- Gluchowski, P., Gabriel, R., Dittmar, C.: Management Support Systeme und Business Intelligence, Computergestützte Informationssysteme für Führungskräfte und Entscheidungsträger, 2. Aufl. Springer, Berlin (2008)
- Inmon, W.H.: Building the data warehouse, 2. Aufl. John Wiley & Sons, New York (1996)
- Merv, A.: It's going mainstream, and it's your next opportunity. In: Teradata Magazine Online, Q1. <http://www.teradatamagazine.com/v11n01/Features/Big-Data/> (2011). Zugegriffen: 1. Nov. 2014
- Plattner, H.: Big data. In: Kurbel, K., Becker, J., Gronau, N., Sinz, E., Suhl, L. (Hrsg.) Enzyklopädie der Wirtschaftsinformatik – Online Lexikon. <http://www.enzyklopaedie-der-wirtschaftsinformatik.de/wi-enzyklopaedie/lexikon/datenwissen/Datenmanagement/Datenmanagement-Konzeptes/Big-Data/index.html> (2013). Zugegriffen: 30. Sept. 2014
- Steria Mummert Consulting GmbH: Business intelligence: status Quo in Europa. Europäische biMA-Studie. <http://www.steria.com/de/bi/bidm-einblicke/europaeische-bima-studie-201213>. (2013). Zugegriffen: 31. März 2014
- TDWI: Big data analytics, best practices report 4 (2011)

Analytische Informationssysteme aus Managementsicht: lokale Entscheidungsunterstützung vs. unternehmensweite Informations- Infrastruktur

Robert Winter

Inhaltsverzeichnis

5.1	Analytische Informationssysteme aus Managementsicht	68
5.1.1	Dimensionen „Umsetzungsunabhängigkeit“ und „Reichweite“	68
5.1.2	Infrastruktur vs. Analytik	71
5.1.3	Unternehmensweite Informations-Infrastruktur	72
5.1.4	Analytik	74
5.1.5	Gestaltungsaufgaben für Informations-Infrastruktur und Analytik	76
5.2	Unternehmensweite Informations-Infrastruktur aus Managementsicht	78
5.2.1	Projekt- vs. Betriebssicht	78
5.2.2	Informationsinfrastruktur-Strategie	79
5.2.3	Informationsinfrastruktur-Organisation	82
5.2.4	Finanzielle Aspekte der Informationsinfrastruktur	84
5.2.5	IT/Business Alignment für die Informations-Infrastruktur	86
5.3	Analytik aus Managementsicht	87
5.3.1	Projekt- vs. Betriebssicht	88
5.3.2	Anwendungspotenziale und Wertbeitrag von Analytik	88
5.3.3	Realisierungsformen von Analytik	90
5.3.4	IT/Business Alignment in der Analytik	91
5.4	Ausblick	92
	Literatur	93

R. Winter (✉)
Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen, Schweiz
E-Mail: robert.winter@unisg.ch

Zusammenfassung

Als soziotechnische Systeme verbinden analytische Informationssysteme genauso wie andere Informationssysteme menschliche und technische Aufgabenträger im Kontext von Organisationen. Analytische Informationssysteme sind fast immer ein komplexes Konglomerat aus Komponenten zwischen den beiden Extrema „Entscheidungsunterstützung einzelner Personen für bestimmte Entscheidungen“ einerseits und „Nutzung von Synergien durch den unternehmensweiten Austausch von Daten“ andererseits. Managementansätze müssen diese Vielgestaltigkeit adressieren, d. h. die Eigenheiten der jeweils unterschiedlichen Leistungen, Ziele, Bedarfs- und Produktionsaspekte sowie Planungs- und Steuerungsansätze differenziert berücksichtigen.

Als exemplarische (Extrem-)Szenarien für analytische Informationssysteme grenzt dieser Beitrag das „Management dezentraler Entscheidungsunterstützung“ vom „Management unternehmensweiter Informations-Infrastrukturen“ ab. Dabei wird jeweils analysiert, welche Planungs- und Steuerungsherausforderungen mit der Entwicklung und dem Betrieb „im Großen“ und „im Kleinen“ verbunden sind. Soweit jeweils anwendbar, werden zentrale Positionierungs- und Gestaltungsfragen wie strategische Ausrichtung, organisatorische Umsetzung und Alignment mit entsprechenden IT-Systemen angesprochen. Eine besondere Rolle spielen aus Managementsicht finanzielle Aspekte analytischer Informationssysteme wie z. B. Wertbeitrag, Leistungsverrechnung und Finanzierung.

5.1 Analytische Informationssysteme aus Managementsicht

Dieser Abschnitt dient zur Charakterisierung der beiden Konzepte „unternehmensweite Informations-Infrastruktur“ (Unterabschn. 5.1.3) und „lokale Entscheidungsunterstützung“ (nachfolgend als Analytik bezeichnet) (Unterabschn. 5.1.4). Diese Konzepte beschreiben zwar beide fachliche, aber einerseits extrem übergreifende und andererseits extrem lokale Komponenten analytischer Informationssysteme (AIS). AIS in der Unternehmensrealität kombinieren meist Elemente beider Extremkonzepte. Um die wichtigsten Erklärungs- und Gestaltungsherausforderungen beschreiben zu können (Unterabschn. 5.1.5), ist es aber zunächst notwendig, einen konzeptionellen Rahmen mit den Dimensionen „Umsetzungsunabhängigkeit“ und „Reichweite“ aufzuspannen (Unterabschn. 5.1.1). Auf dieser Grundlage können die beiden Extremkonzepte ausreichend trennscharf voneinander und von anderen, bestehenden Konzeptionen im AIS-Bereich abgegrenzt werden (Unterabschn. 5.1.2).

5.1.1 Dimensionen „Umsetzungsunabhängigkeit“ und „Reichweite“

Als soziotechnische Systeme verbinden AIS genauso wie andere Informationssysteme menschliche und technische Aufgabenträger im Kontext von Organisationen. Aus Ma-

managementsicht darf sich die AIS-Gestaltung deshalb nicht auf die Formulierung der jeweils relevanten funktionalen und konstruktiven Anforderungen an die IT-Unterstützung beschränken. Auch die Identifikation der adressierten „Kunden“-Gruppen und der abgedeckten Bedürfnisse sowie die Gestaltung der Ziele, Leistungen, Produktionsmechanismen, Verantwortlichkeiten, Abläufe usw. der Informationsversorgung aus fachlicher Perspektive gehört zur Analyse- und Gestaltungsaufgabe des (in diesem Fall analytischen) Informationssystems.

Für die ganzheitliche, d. h. sowohl die fachlichen Aspekte wie auch die IT-Umsetzung umfassende Analyse und Gestaltung von Informationssystemen wurde der Business Engineering-Ansatz (Österle und Winter 2003a) entwickelt. Das Business Engineering-Framework sieht verschiedene Beschreibungs- bzw. Gestaltungsebenen vor, die sich durch unterschiedliche Fach- bzw. IT-Umsetzungsnahe unterscheiden. Die implementierungsunabhängigste Ebene wird als Strategieebene bezeichnet. Auf dieser Ebene werden u. a. die Positionierung, die Ziele und die Leistungen (einschl. Qualitätseigenschaften) des AIS festgelegt. Auf Grundlage der Beantwortung dieser sog. „Was“- (bzw. „Was nicht“-) Fragen erfolgt auf der Organisationsebene die Festlegung von Abläufen (z. B. Leistungs- und Informationsflüsse), Strukturen (z. B. Stellen, Rollen, Informationsobjekte, Bezugsgrößen) und deren Zusammenhängen. Die Antworten auf diese „Wie“- (bzw. „Wie nicht“-) Fragen stellen eine Art Umsetzung der strategischen Ausrichtung dar – aber eben (noch) nicht mit IT-Mitteln, sondern in Form fachlicher Strukturen und Prozesse („Organisation“). Die IT-Umsetzung wird im Business Engineering auf den Ebenen „Software“ und „IT-Infrastruktur“ beschrieben bzw. gestaltet. Hier finden sich mit der fachlichen und technischen Spezifikation von Softwarekomponenten, Datenstrukturen und Hardwarekomponenten die Antworten auf die „Womit“- (bzw. „Womit nicht“-) Fragen. Da fachliche Lösungen andere Granularität, u. U. andere Dezentralität und vor allem andere Innovationszyklen haben als IT-Umsetzungen, wird meist eine explizite Entkopplungs-Architecturebene notwendig sein, die im Business Engineering als Alignment-Ebene bezeichnet wird und auf der u. a. IT-Capabilities, (logische) Applikationen und Domänen beschrieben bzw. gestaltet werden (Aier und Winter 2009).

Abbildung 5.1 illustriert die fünf Ebenen des Business Engineering-Frameworks in der üblichen „Stack“-Darstellung. Diese Darstellung sollte nicht dahingehend interpretiert werden, dass die Informationssystem-Gestaltung immer „top-down“ auf fachlicher Ebene startet und mit der IT-Umsetzung endet. Neben diesem „klassischen“ Projekttyp gibt es auch viele AIS-Vorhaben, die ihren Ausgangspunkt in einer IT-Innovation (z. B. „in-Memory“-Datenhaltung) haben und Änderungen der organisatorischen Einbettung (Abläufe, Verantwortlichkeiten etc.) erfordern bzw. ermöglichen (Winter et al. 2012), neue/veränderte Informationsversorgungsleistungen ermöglichen und/oder die Verfolgung neuer/anderer Ziele ermöglichen. Daneben können als weiterer Projekttyp „Alignment“-Projekte fachliche und IT-Umsetzungs-Strukturen näher „zusammenbringen“, z. B. nach Phasen paralleler Entwicklung mit sinkender Kohärenz zwischen diesen Ebenen. Schließlich gibt es sog. „Housekeeping“- bzw. Reorganisationsprojekte, welche die IT-Umsetzungsland-

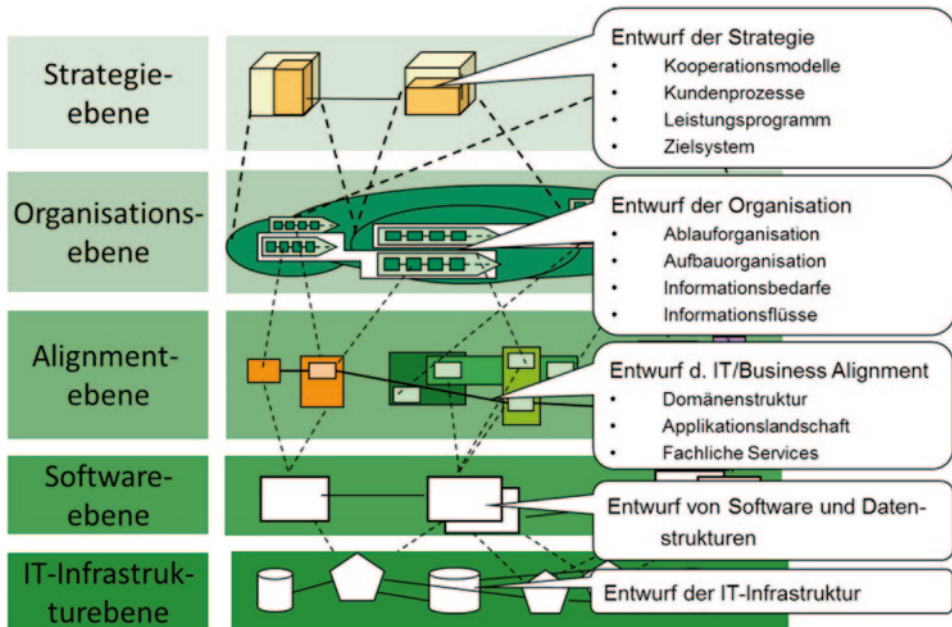


Abb. 5.1 Business Engineering-Framework. (in Anlehnung an Winter 2008)

schaft oder auch die fachlichen Strukturen „aufräumen“, ohne zu unmittelbaren Veränderungen der jeweils anderen Seite zu führen (Winter 2011).

Als erste Dimension des konzeptionellen Rahmens für die Diskussion von AIS aus Managementsicht sollte deshalb die Umsetzungsunabhängigkeit eingeführt werden. Es ist nicht wichtig, wie viele Architekturebenen zwischen „vollständig umsetzungsunabhängig“ (Strategieebene) und „vollständig umsetzungsabhängig“ (IT-Infrastrukturebene) vorgesehen sind. Zusätzliche Architekturebenen erzeugen zwar einerseits zusätzlichen Gestaltungs- und Abgleichungsbedarf, unterstützen aber andererseits die gezielte Verfolgung ebenenspezifischer Ziele.

Sowohl fachliche Lösungen wie auch IT-Umsetzungen können extrem unterschiedliche Reichweite haben. In Anlehnung an typische Unternehmensstrukturen kann die Reichweite von stellen-/aufgaben-spezifisch über abteilungs-/bereichsweit bis unternehmensweit oder sogar unternehmensübergreifend reichen. Die Reichweite ist dabei weitgehend unabhängig von der Umsetzungsunabhängigkeit: Zwar haben strategische Ausrichtungen im Normalfall eher größere Reichweite, während auf allen anderen Beschreibungs- bzw. Gestaltungsebenen die verschiedensten Reichweiten zu finden sind. Aber auch Teile der IT-Infrastruktur können durchaus unternehmensweit standardisiert sein, während die Geschäftseinheiten hinsichtlich Leistungsprogramm und adressierten Kundenprozessen „lokal“ aufgestellt sind. Es ist jedoch häufig festzustellen, dass die Fristigkeit der Festlegungen mit zunehmender Reichweite korreliert: Aus naheliegenden Gründen sind stellen-/aufgaben-spezifische Festlegungen einfacher zu ändern – und werden deshalb auch häufiger

figer und umfassender geändert – als bereichsweite, unternehmensweite oder gar unternehmensübergreifende Festlegungen.

Als zweite Dimension des konzeptionellen Rahmens für die Diskussion von AIS aus Managementsicht wird aufgrund der Orthogonalität mit der Umsetzungsunabhängigkeit die Reichweite (und umgekehrt proportional dazu implizit auch Änderungsdynamik) vorgeschlagen. Auch hier spielt die Zahl und Trennschärfe der Abstufungen zwischen stellen-/aufgabenpezifisch und unternehmensübergreifend zunächst keine zentrale Rolle.

5.1.2 Infrastruktur vs. Analytik

Auf Grundlage des im vorigen Abschnitt aufgespannten konzeptionellen Rahmens mit den Dimensionen Umsetzungsunabhängigkeit und Reichweite lassen sich in erster Annäherung nun unterschiedliche, einerseits primär fachliche oder primär technische sowie andererseits primär übergreifende oder primär „lokale“ AIS-Komponenten unterscheiden:

Data Warehousing ist als Nutzung einer aus den verschiedensten Datenquellen integrierten, themenorientiert strukturierten, persistenten und zeitbezogenen Datenbasis (Inmon 2002, S. 31) als „single version of the truth“ für alle Analysen eher umsetzungsorientiert definiert und hat klar übergreifenden Charakter. Nach der klar umsetzungsorientierten Innovation wurden zwar schon früh Managementansätze für Data Warehousing vorgeschlagen (Jung und Winter 2000b); gleichwohl verweist „Warehousing“ nach wie vor auf eine Umsetzungsarchitektur und nicht auf eine fachliche Anforderung.

Business Intelligence (BI) ist als Nutzung einer großen Bandbreite von Konzepten und insbesondere Werkzeugen zur Entscheidungsunterstützung (Anandarajan et al. 2003, S. 18 f.; Gartner 2004) zwar tendenziell fachlicher orientiert, aber gleichwohl stark durch Umsetzungsaspekte wie z. B. multidimensionale Analysen, Dashboards, Reporting o.ä. geprägt (Wixom und Watson 2010). Unabhängig von der Umsetzungsunabhängigkeit adressiert BI jedoch immer konkrete Nutzungsszenarien, d. h. ist deutlich lokaler ausgerichtet. Die Darstellung von BI auf Grundlage des vorgeschlagenen zweidimensionalen konzeptionellen Rahmens hat deshalb Überschneidungen mit Data Warehousing und erstreckt sich etwas weiter in Richtung Fachlichkeit. Dies kann durchaus dahingehend interpretiert werden, dass die AIS-Komponente Data Warehousing ein „Enabler“ für die AIS-Komponente BI ist (Ramamurthy et al. 2008).

Als Bezeichnung einer für die Betrachtung von AIS aus Managementsicht wichtigen, umsetzungsunabhängigen wie auch übergreifenden Konzeption wurde integrierte Informationslogistik vorgeschlagen (Winter et al. 2008). Auch wenn dieser Begriff aufgrund der Logistik-Metapher durchaus geeignet erscheint, um die gesamthafte Sicht und die fachliche Servicefunktion auszudrücken, wird er gleichwohl nur selten verwendet. Bessere Chancen auf breite Verwendung scheint der Begriff der Infrastruktur zu haben, mit der ebenfalls gesamthafte Sicht und Servicefunktion verbunden werden (Buhr 2003) – wenn gleich „Infrastruktur“ oft unzulässig nur auf die technische Infrastruktur bezogen wird. Da bei AIS Informationen im Mittelpunkt stehen, wird im Folgenden für die gesamthafte

AIS-Betrachtung der Begriff „**unternehmensweite Informations-Infrastruktur**“ verwendet.

Die Findung einer Bezeichnung für ein umsetzungsunabhängiges, lokales Unterstützungskonzept gestaltet sich schwieriger. Auch in diesem Fall ist es wichtig, durch einen geeigneten Begriff keine spezifischen Umsetzungsimplicationen zu erzeugen und gleichzeitig den „lokalen“ Charakter dieses für die AIS-Betrachtung aus Managementsicht wichtigen Konzepts zu betonen. Mögliche, selbsterklärende Begriffe sind hier analytische Prozessunterstützung oder lokale Entscheidungsunterstützung. In den letzten Jahren wird zunehmend der Begriff **Analytik** verwendet, der jedoch nicht selbsterklärend ist. Wir definieren hier Analytik als umsetzungsneutralen (d. h. fachlich orientierten), aber gleichzeitig auf spezifische Entscheidungstypen, Entscheiderrollen, Informationen o. ä. zugeschnittenen AIS-Ansatz. Hier geht es nicht um bereichs- oder gar unternehmenübergreifende Synergien, sondern um die spezifische „lokale“ Entscheidungsunterstützung aus fachlicher Sicht.

Abbildung 5.2 illustriert die Positionierung der vier vorgestellten AIS-Analyse- und Gestaltungskomponenten im durch „Umsetzungsunabhängigkeit“ (horizontale Dimension) und „Reichweite“ (vertikale Dimension) aufgespannten konzeptionellen Rahmen. Auf der Grundlage dieser Positionierung werden die Konzepte unternehmensweite Informations-Infrastruktur und Analytik in den folgenden beiden Abschnitten spezifiziert.

5.1.3 Unternehmensweite Informations-Infrastruktur

Für die Planung, Steuerung, Durchführung und Kontrolle der Gesamtheit der Datenflüsse, die über eine Betrachtungseinheit (Abteilung/Funktion, Bereich, Konzerngesellschaft)

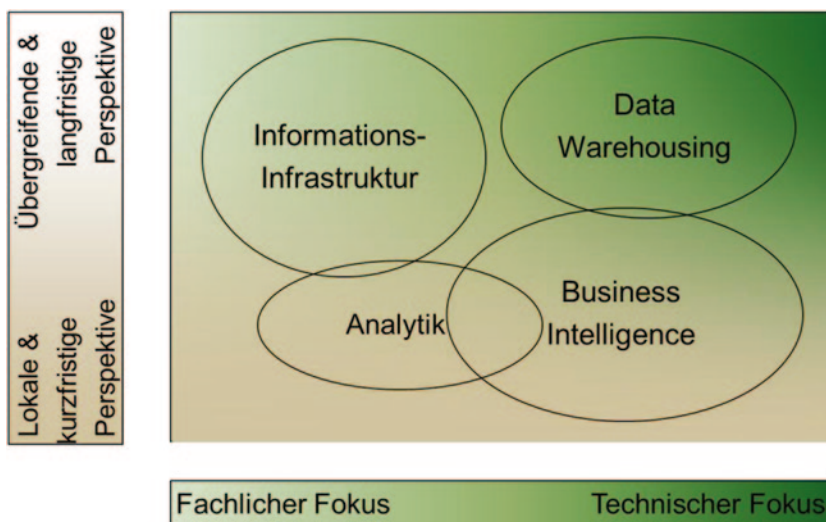


Abb. 5.2 Positionierung zentraler Konzeptionen für analytische Informationssysteme

hinausgehen, sowie die Speicherung und Aufbereitung dieser Daten wurden vor einigen Jahren der Begriff der integrierten Informationslogistik vorgeschlagen (Winter et al. 2008). Im Konzept der integrierten Informationslogistik wird von der IT-Umsetzung vollkommen abstrahiert: Informationslogistik kann im (kaum vorstellbaren) Extremfall auch mit Zettelkästen, scharfem Nachdenken und/oder Gruppensitzungen umgesetzt werden. Schließlich ist zu beachten, dass lokale wie auch übergreifende Datenflüsse, die nicht der Entscheidungsunterstützung dienen (z. B. operative Applikationsintegration), ebenfalls nicht der Informationslogistik zugerechnet werden. Abbildung 5.3 visualisiert dieses Verständnis von Informationslogistik. Welche Datenflüsse jeweils als Komponenten der Informationslogistik angesehen werden, hängt vom jeweiligen Standpunkt (Stelle/Aufgabe, Abteilung, Geschäftsbereich, Unternehmen) ab. Auch hier ist nicht die Zahl der Betrachtungsebenen wichtig, sondern allein deren Vorhandensein und die darauf basierende Abgrenzung des Konzepts.

Schon bei Szyperski und Klein (Szyperski und Klein 1993) und später bei von Maur et al. (2003) wurde Informationslogistik erwähnt, auch wenn seinerzeit die operative Datenintegration nicht ausgeschlossen bzw. sogar explizit einbezogen wurde und der Begriff nicht vollkommen umsetzungsunabhängig bzw. sogar umsetzungsorientiert benutzt wurde. Allerdings muss festgestellt werden, dass sowohl die Forschergemeinschaft wie auch die Praxis den Begriff der integrierten Informationslogistik nur partiell und zögerlich adoptiert.

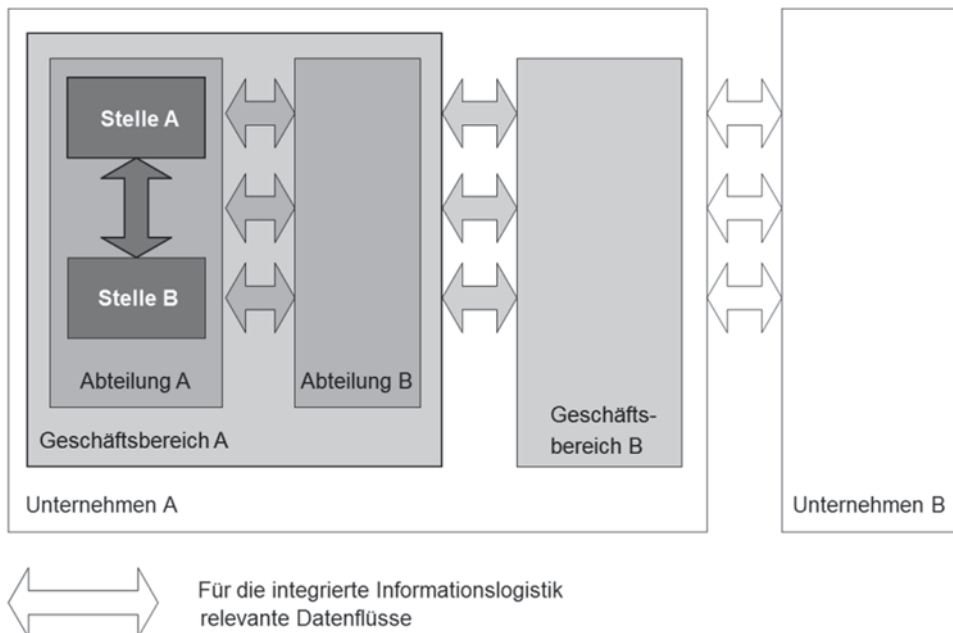


Abb. 5.3 Integrierte Informationslogistik auf verschiedenen Betrachtungsebenen. (in Anlehnung an Winter et al. 2008)

In dieser Hinsicht erfolgversprechender erscheint der Begriff der Informations-Infrastruktur. Dieser kann bei Bedarf hinsichtlich der Reichweite in Form von bereichsweit (d. h. abteilungs-/funktionsübergreifend), konzerngesellschaftsweit (d. h. bereichsübergreifend) oder unternehmensweit (d. h. konzerngesellschaftsübergreifend) spezifiziert werden und ist damit trennschärfer. Weiterhin werden mit „Infrastruktur“ genau die ökonomischen und organisatorischen Herausforderungen verbunden, vor denen übergreifende Informationsbewirtschaftung in komplexen Systemen meist steht: Während die Kosten zu großen Teilen früh und meist zentral anfallen, fallen die Nutzeffekte später und an den verschiedensten Orten an. Während alle lokalen Einheiten gerne eine funktionierende Infrastruktur nutzen möchten, ist die Bereitschaft zur Finanzierung oder auch nur Nutzungsabgeltung sehr beschränkt. Aus Sicht der Gesamtorganisation ist es immer schwer, den Bedarf und Nutzen korrekt einzuschätzen und die Infrastrukturentwicklung effizient und nutzungsgerecht zu organisieren. Für eine Übersicht der mit dem Konzept „Infrastruktur“ verbundenen grundlegenden Herausforderungen sei auf Buhr (2003) verwiesen.

5.1.4 Analytik

Als Analytik wird die Planung, Steuerung, Durchführung und Kontrolle der Datengewinnung und analytischen Informationsnutzung in (Geschäfts-, Unterstützungs- oder Führungs-)Prozessen verstanden. Die genutzte Information muss sich auf die vom jeweiligen Prozess behandelten Objekte beziehen und zur Ausführungszeit in der festgelegten Qualität verfügbar sein (Bucher und Dinter 2008a, b). Dabei ist es unerheblich, ob die unterstützten Prozesse nicht-automatisiert, teilautomatisiert oder vollständig automatisiert sind.

Wichtig an dieser Definition ist im Gegensatz zum Konzept der Informations-Infrastruktur die Fokussierung auf lokale Effektivität und Effizienz. Im Zentrum des Interesses stehen weder Synergien durch übergreifende Datennutzung noch Prozessführung, sondern die Transformation lokal verfügbarer Daten in lokal entscheidungsunterstützende Informationen „zur Laufzeit“. Analog zur Definition von Informations-Infrastruktur richtet sich die „Lokalität“ jedoch nach der Betrachtungseinheit: vom Standpunkt der Stelle/Aufgabe sind nur deren Prozessteile lokal, vom Standpunkt einer Abteilung nur deren (Teil-)Prozesse usw. Auch für die Analytik-Definition gilt, dass die Definition die Planung, Steuerung, Durchführung und Kontrolle der Datengewinnung und analytischen Informationsnutzung umfasst, wobei von der Form der Realisierung vollkommen abstrahiert wird. Analytik wird tatsächlich in einigen Fällen mit Zettelkästen, Gruppensitzungen oder anderen, nicht strukturierten Mitteln umgesetzt. Schließlich ist auch in diesem Fall zu beachten, dass Datennutzungen, die nicht der Entscheidungsunterstützung dienen (z. B. operative Prozessintegration, vollautomatisierte Abwicklungsprozesse), nicht der Analytik im Sinne dieses Beitrags zugerechnet werden.

Auch das BI-Konzept befasst sich nicht nur aus IT-Umsetzungssicht, sondern auch aus fachlicher Sicht mit der Transformation von Daten in entscheidungsrelevante Informationen. Die Gestaltung ist jedoch einerseits primär umsetzungs-getrieben, d.h. folgt über-

spitzt formuliert dem Prinzip „habe Hammer, suche Nagel“. Wenn z. B. für eine bestimmte Steuerungsproblematik mehrere Umsetzungsoptionen (wie multidimensionale Analyse, Entscheidungsunterstützungssystem, interaktiver Report oder Dashboard) denkbar sind, wird der BI-Ansatz möglicherweise zu mehreren Fachkonzepten führen und nicht zu einer einzigen fachlichen Lösung, für die dann zwischen alternativen Umsetzungen ausgewählt werden kann. Andererseits fokussiert BI traditionell auf taktische, mitunter auch strategische oder im Fall von „operational BI“ operative Entscheidungen, wie sie typischerweise in Führungsprozessen getroffen werden (Bucher und Dinter 2008b). Das analytische Potenzial, das gerade für die Effektivität und Effizienz operativer Geschäftsprozesse (und nicht allein für deren Steuerung) aufgrund deren Häufigkeit und Erfolgswirksamkeit genutzt werden könnte, wird dadurch nicht voll ausgenutzt.

Der Begriff Management Support Systems (MSS, deutsch Führungsunterstützungssysteme bzw. Führungsinformationssysteme) ist ein Sammelbegriff für verschiedene Arten von analytischen Systemen (Holten 1999, S. 29). Unter den Begriffen Management Information Systems („MIS“, Gallagher 1961), Decision Support Systems („DSS“, Gorry und Scott Morton 1971) und Executive Information Systems („EIS“, Rockart und Treacy 1980) werden verschiedene Klassen analytischer Informationssysteme verstanden, deren Entwicklung bis in die 1960er Jahre zurückreicht. Im Lichte des hier skizzierten konzeptionellen Rahmens beschreiben diese Begriffe Spezialkonzepte, die sich kaum noch in ihrer ursprünglichen Form finden (Gluchowski und Kemper 2006, S. 12).

Unter MIS werden Systeme zur Versorgung des mittleren Managements mit Berichten zur Unterstützung alltäglicher, strukturierter Entscheidungen verstanden, die auf Unternehmensdaten aus den operativen Systemen basieren, welche teilweise historisiert sein können und die in der Regel engen Subjektbezug aufweisen (Davis und Olson 1985, S. 6). Daneben wird der Begriff MIS auch teils synonym für die MSS-Gesamtheit verwendet.

DSS sind Systeme, die basierend auf Datenanalysemodellen und Datenbanken das mittlere Management interaktiv beim Treffen von semistrukturierten, nicht standardisierten Entscheidungen unterstützen sollen (Davis und Olson 1985, S. 11 f.). Hierzu werden teils auch OLAP- und Data Mining Tools gezählt (Laudon und Laudon 2006, S. 481). Verwandt sind die Expertensysteme, die basierend auf einer Wissensbasis mit Methoden der künstlichen Intelligenz Entscheidungen in eng abgegrenzten Bereichen automatisiert treffen können (Holten 1999, S. 38).

EIS schließlich sind Systeme zur Unterstützung unstrukturierter, strategischer Entscheidungen, die in erster Linie für das Top-Management gedacht sind. Da sie sich zwar auf „lokale“ Entscheidungen fokussieren, aber immer eine bestimmte Umsetzungsarchitektur – oder sogar ein konkretes Umsetzungskonzept wie bei klassischen MIS – implizieren und auf (meist taktische, aber im Fall von EIS auch strategische) Führungsprozesse zugeschnitten sind, lassen sich MSS bzw. die verschiedenen MSS-Applikationstypen unter dem hier verwendeten BI-Begriff subsumieren. Aus fachlicher Sicht muss Organisation und Strategie ausreichend einbezogen werden, und die Unterstützung sollte sich nicht auf Führungsprozesse beschränken.

In letzter Zeit wird viel über „Big Data“ diskutiert. Allerdings wird bei näherer Betrachtung schnell klar, dass „Big Data“ nicht auf die fachlichen Aspekte der (lokalen) Analytik abzielt, sondern aus Sicht der Datenidentifikation, -sammlung und -bereitstellung ein Analytik-übergreifendes Potenzial zu schaffen versucht, das von den hier skizzierten vier AIS-Komponenten dem Data Warehousing am nächsten liegt – aber im Gegensatz zum traditionellen Data Warehousing-Verständnis nicht auf Transaktionssystem-generierte, mehrheitlich strukturierte Daten fokussiert, sondern auch Menschen- und Sensor-generierte, oft auch unstrukturierte und unternehmensexterne Daten einbezieht, die in einer Menge, Heterogenität und Geschwindigkeit anfallen, welche die Möglichkeiten traditioneller Technologien übersteigt.

5.1.5 Gestaltungsaufgaben für Informations-Infrastruktur und Analytik

Auf Grundlage des weiter oben eingeführten Business Engineering-Frameworks lassen sich die fachlichen Gestaltungsaufgaben für Informations-Infrastruktur und Analytik wie folgt strukturieren: Auf **Strategieebene** geht es darum, die Bereitstellung

- der richtigen Informationen
- in der richtigen Qualität (z.B. rechtzeitig, korrekt) und
- in wirtschaftlicher Weise

sicherzustellen, um alle relevanten Informationsbedarfe zu befriedigen. Aus Infrastruktur-Sicht steht dabei als Hauptziel die Realisierung von Synergien im Mittelpunkt („Das Ganze ist mehr als die Summe seiner Teile“), so dass als Leistungen insbesondere der Datenaustausch zwischen Einheiten festzulegen ist. Aus Analytik-Sicht steht dagegen als Hauptziel die richtige Unterstützung der richtigen Entscheidungen im Mittelpunkt, so dass als Leistungen die prozessbezogenen Datenbereitstellungen festzulegen sind.

Neben dem Leistungsprogramm sind die dabei zu verfolgenden Ziele und deren Gewichtung (z. B. Datenqualitätsziele, aber auch Zufriedenheits- oder Wertbeitragsziele) sowie die zu bedienenden (internen oder externen) Kunden festzulegen. Analog zu dieser Kundenperspektive sind aus Lieferantenperspektive die beziehbaren Vorleistungen (z. B. operative Rohdaten), die dabei zu verfolgenden Ziele und deren Gewichtung sowie die verfügbaren (internen und externen) Datenlieferanten festzulegen. Der Kernbeitrag des AIS-Managements besteht jedoch nicht in Kunden- und Lieferantenmanagement, sondern im Management der Transformation der bezogenen Daten in entscheidungsrelevante Informationen. Für diese Kernleistung sind die dabei zu verfolgenden Ziele, die dafür notwendigen Kompetenzen und das alle Teilaspekte umfassende Planungs- und Steuerungssystem festzulegen (Dinter und Winter 2009).

Auf **Organisationsebene** geht es darum, auf Grundlage der strategischen Ausrichtung (Positionierung, Lieferanten- und Kundenleistungen, Ziele, Transformationskompetenzen, Governance) eine geeignete Ablauf- und Aufbauorganisation festzulegen. Da

es sich bei Vorleistungen um Daten (in bestimmter Qualität) und bei Kundenleistungen um entscheidungsrelevante Informationen (in bestimmter Qualität) handelt, entspricht die AIS-Ablauforganisation in weiten Teilen den Daten- bzw. Informationsflüssen und entspricht die AIS-Aufbauorganisation in weiten Teilen Daten- bzw. Informationslandschaften. Dadurch wird die Organisationsgestaltung einfacher als im Business Engineering-Normalfall, in dem Daten-/Informationsflüsse wie auch -strukturen häufig stark von der Erstellung der Kernleistungen und den entsprechenden Verantwortungsstrukturen divergieren. Mit zunehmender Erfahrung in Informationslogistik und Entscheidungsunterstützung müssen auf Organisationsebene nicht immer wieder völlig neue Lösungen gesucht werden, sondern es kann auf existierende Referenzmodelle zurückgegriffen werden. So beschreiben z. B. (Klesse und Winter 2007) vier verschiedene Grundmodelle der Organisationsgestaltung für Datenintegrationsdienstleister.

Als Ergebnis der Organisationsgestaltung liegen AIS-Prozessmodelle und -Aufgaben/-Verantwortungsstrukturen auf verschiedenen Aggregationsebenen vor, die mit entsprechenden AIS-Daten-/Informationsstrukturen verknüpft sind. Neben den Kernprozessen und -strukturen sind natürlich auch Unterstützungs- und Führungsaufgaben wie Planung, Steuerung, Leistungsverrechnung sowie Kunden- und Lieferantenmanagement festzulegen.

Auf **Software- und IT-Infrastrukturebene** geht es darum, für die fachlichen „Was“- (bzw. „Was nicht“-) sowie „Wie“- (bzw. „Wie nicht“-) Festlegungen geeignete Softwaresysteme und IT-Infrastrukturen zu spezifizieren. Im Hinblick auf die Infrastruktur kommen dabei verschiedene Kombinationen und Variationen von Data Warehousing (allenfalls mit „Big Data“-Erweiterungen), operativer Datenintegration und/oder unternehmensübergreifender Datenintegration in Frage (von Maur et al. 2003). Im Hinblick auf die Analytik stehen verschiedene Lösungsvarianten prozessorientierter BI zur Verfügung (Bucher und Dinter 2008a; Bucher 2009). Hinsichtlich der Gestaltung des Fachkonzepts für AIS-Softwaresysteme sowie entsprechender IT-Infrastrukturen sowie der Umsetzung in entsprechende IT-Lösungen finden sich spezifische Ansätze in anderen Beiträgen dieses Buches.

Auf **Alignmentebene** geht es darum, die Passung von IT-Lösungen und fachlichen Festlegungen sicherzustellen. Während eines erfolgreichen Entwicklungsprojekts ist dies definitionsgemäß der Fall. Einerseits können jedoch unterschiedliche Entwicklungsprojekte Interferenzen auf fachlicher oder IT-Ebene erzeugen; andererseits entstehen im Zeitverlauf durch fachlich getriebene Innovationsprojekte, technisch getriebene Innovationsprojekte sowie „Housekeeping“- bzw. Reorganisationsprojekte (z. B. Master Data Management, Data Cleansing o. ä.) mehr oder weniger große Diskrepanzen. Natürlich kann es auch vorkommen, dass dem Alignment aufgrund z. B. eines dominanten IT- oder Fach-„Mindsets“ zu geringe Aufmerksamkeit geschenkt wird.

Ein zentrales Instrument zum Management von Alignment ist die aktive Bewirtschaftung der Applikationsarchitektur. Weitere Konzepte, mit denen auf Alignmentebene die Zuordnung von Softwaresystemen zu fachlichen Bedarfen dokumentiert und weiterentwickelt werden können, sind fachliche Services (zur Anwendung in der Informationslogistik (Dinter 2008)) und Domänen (Aier und Winter 2009).

5.2 Unternehmensweite Informations-Infrastruktur aus Managementsicht

Dieser Abschnitt betrachtet das Management von Informations-Infrastrukturen und bezieht sich damit auf die in Abschn. 5.1.5 skizzierten Gestaltungsaufgaben. Zunächst wird in Abschn. 5.2.1 diskutiert, inwieweit sich Aufbau/Veränderung und Betrieb von Informations-Infrastrukturen unterscheiden und was daraus für entsprechende Unterstützungsansätze folgt. Die Abschn. 5.2.2 und 5.2.3 adressieren die aus Managementsicht zentralen Ebenen Strategie und Organisation. Da aus Managementsicht gerade die finanziellen Aspekte von Informations-Infrastrukturen wichtig sind, ist Abschn. 5.2.4 ganz dem Wertbeitrag, der Leistungsverrechnung und der Finanzierung gewidmet. Abschließend werden in Abschn. 5.2.5 Fragen des IT/Business Alignment behandelt.

5.2.1 Projekt- vs. Betriebssicht

Das St. Galler Managementmodell kennt verschiedene Blickwinkel, unter denen Unternehmen als soziale Systeme analysiert und gestaltet werden können (Rüegg-Sturm 2003). Einer dieser Blickwinkel ist der sog. Entwicklungsmodus, der entweder „Optimierung“ oder „Erneuerung“ sein kann. Für den Betrieb und die höchstens inkrementelle Veränderung in Unternehmen sind andere Konzepte und Ansätze notwendig als für die grundsätzliche Neu- bzw. Umgestaltung. Die Unterscheidung zwischen Optimierung und Erneuerung findet sich in der Wirtschaftspraxis im Begriffspaar „Run the Business“ vs. „Change the Business“, das in einer zunehmenden Zahl von Fällen auch als Leitschnur für die Strategie- und Organisationsgestaltung Anwendung findet.

Während die meisten Methoden der betriebswirtschaftlichen Funktionslehren (wie z. B. Marketing, Finanzmanagement, Personalwesen etc.) nicht dediziert auf Veränderungsvorhaben zugeschnitten sind und damit dem „Run the Business“ zugeordnet werden können, versteht sich Business Engineering als „betriebswirtschaftliche Konstruktionslehre“, d. h. stellt Modelle und Methoden für „Change the Business“ bereit (Österle und Winter 2003b).

In Bezug auf AIS sollte deshalb sorgfältig zwischen „Change“ und „Run“ unterschieden werden. Während die Neu- und Umgestaltung im Normalfall projektweise erfolgt und deshalb aus Managementsicht primär als Programm- bzw. Projektmanagement geplant und durchgeführt werden sollte, erfolgt der Dauerbetrieb im Normalfall als Portfolio miteinander vernetzter Lösungen, so dass klassische Informationsmanagement-Ansätze (z. B. Referenzprozesse für Portfolio- und Betriebsmanagementprozesse nach (Österle et al. 1992)) zur Anwendung kommen sollten. Abbildung 5.4 illustriert diese Unterscheidung: Aus separaten Projekten bei der Neu- bzw. Umgestaltung wird ein Lösungsbündel im Betrieb.

AIS sind im Hinblick auf Neu- bzw. Umgestaltungsmanagement nicht fundamental anders als andere unternehmensweite Infrastrukturen (z. B. Enterprise Architecture

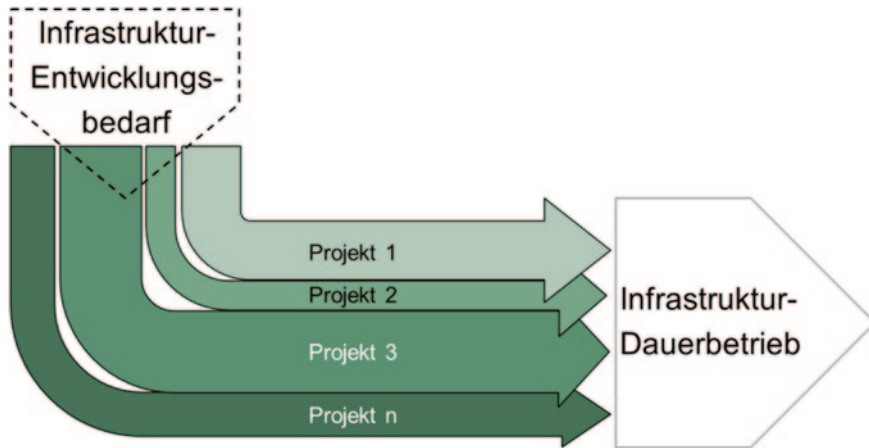


Abb. 5.4 Entwicklungssicht=Projektsicht vs. Betriebssicht=Gesamtsicht. (in Anlehnung an Winter et al. 2009)

Management (EAM)). Gerade im Betrieb kommt jedoch die primäre Infrastruktur-Zielsetzung, die Realisierung von Synergien, zum Tragen. Deshalb ist hier eine spezifische, auf AIS ausgerichtete Ausgestaltung wichtig. Während von durchschnittlichen IS, die meist lokale Zielsetzungen haben, hier wenig gelernt werden kann, bieten Synergie-orientierte, unternehmensweite IS wie z. B. ERP-Systeme oder EAM-Systeme durchaus übertragbare Erkenntnisse an.

5.2.2 Informationsinfrastruktur-Strategie

Die Gliederung der Festlegungen der Informationsinfrastruktur-Strategie in Abschn. 5.1.5 macht bereits deutlich, dass solche AIS als Teil einer Wertschöpfungskette gesehen werden sollten, die von der „Urproduktion“ von Daten, meist in Transaktionssystemen, bis zur finalen Nutzung daraus abgeleiteter Informationen in Entscheidungsprozessen reicht. Die Komponenten einer AIS-Strategie können deshalb durchaus in Anlehnung z. B. an das SCOR-Modell (Supply Chain Council 2008) als Lieferantenstrategie („Source“), Kundenstrategie („Deliver“) und Leistungsstrategie („Make“) formuliert werden, die durch einen entsprechenden Prozess koordiniert werden. Die empirische Untersuchung von Strategien in der Informationslogistik zeigt, dass eine derartige Strukturierung zielführend ist, auch wenn die Untergliederung der Entwicklungs- und Wartungsstrategie nicht Supply Chain-orientiert, sondern eher klassisch funktional erfolgt (Dinter und Winter 2009). Abbildung 5.5 zeigt die sich ergebende Strategie-Strukturierung.

Die Teilstrategien legen in ihrer jeweiligen Ausprägung (Lieferantenfokus, Transformationsfokus, Kundenfokus) das Leistungsprogramm, die Prozesspartner (falls anwendbar), die zu verfolgenden Ziele und deren Gewichtung (z. B. Qualitätsziele oder Wert-

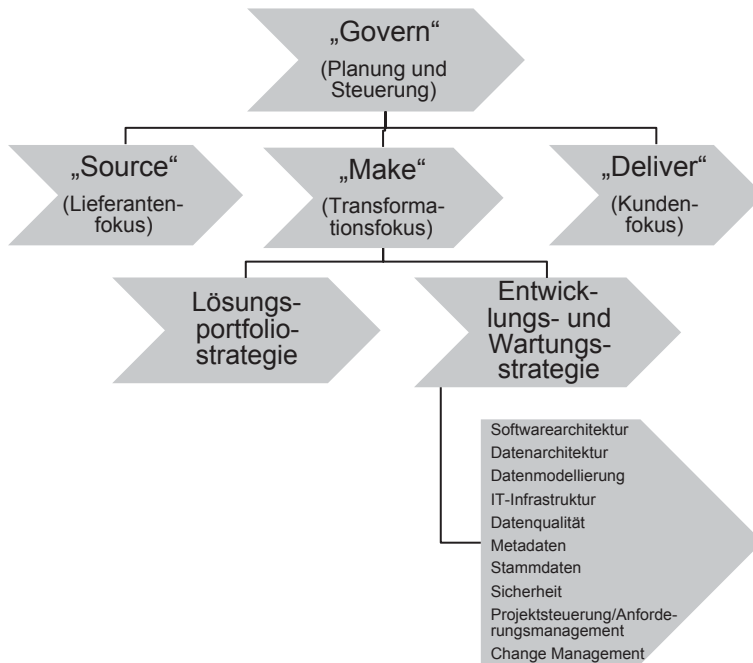


Abb. 5.5 Komponenten der AIS-Strategie. (in Anlehnung an Dinter und Winter 2009)

schöpfungsziele) sowie die dafür notwendigen Kompetenzen fest. Darüber hinaus wird die Koordination der Gesamt-Leistungserbringung festgelegt („AIS-Governance“).

Für die Kundenstrategie ist dabei zu beachten, dass aus Kundensicht nicht der AIS-Ressourceneinsatz relevant ist, sondern die bereitgestellte Entscheidungsunterstützung. Leistungen sollten also nicht in Form von Ressourcennutzung definiert und abgerechnet werden, sondern in Form von (Kunden-)Prozessunterstützung. Abbildung 5.6 zeigt den Entwicklungspfad für Leistungsbeschreibungen, der auch in anderen Bereichen des Informationsmanagements zu beobachten ist. Das Zielsystem der Kundenstrategie wird im Normalfall tendenziell effektivitätsorientiert sein (die richtigen Leistungen erbringen), während für die Lieferanten- und insbesondere die Transformationsstrategie ressourcenorientierte Leistungsdefinitionen und Effizienzziele definiert werden.

Während das zentrale Ziel von Analytik die Bereitstellung verlässlicher Entscheidungsgrundlagen für die Entscheidungen ist, steht für die Informations-Infrastruktur die Realisierung von Synergien an oberster Stelle. Durch die Planung, Steuerung, Durchführung und Kontrolle übergreifender Datenflüsse (und durch die Speicherung und Aufbereitung solcher Daten) muss größerer Nutzen generiert werden, als wenn solche Daten ausschließlich lokal genutzt würden – ansonsten schaffen übergreifende AIS keinen Wertbeitrag und sollten besser lokal und ohne Koordination betrieben werden.

Beispiele für „Das Ganze ist mehr als die Summe seiner Teile“ sind:

- die Nutzung von Schadensdaten für die Kalkulation von Prämien neuer Versicherungsverträge (Synergien zwischen Abteilungen eines Geschäftsbereichs),

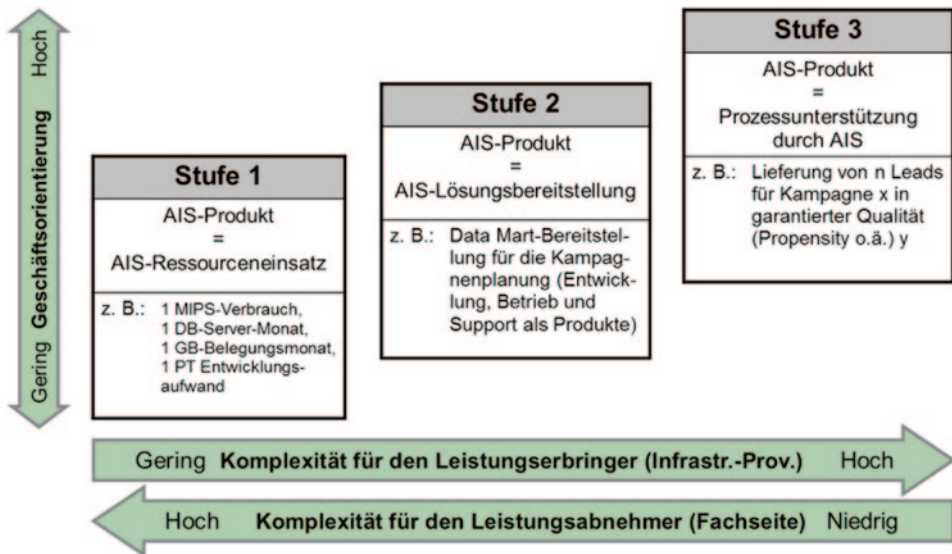


Abb. 5.6 Leistungsbegriff – Vom Ressourceneinsatz zur Prozessunterstützung (Herrmann et al. 2004) in Anlehnung an (Zarnekow und Brenner 2003)

- die Nutzung der Verkaufsdaten der Banksparte für Verbundverkäufe der Versicherungsparte (Synergien zwischen Geschäftsbereichen eines Unternehmens) oder
- die Nutzung von Flugbuchungs-Daten für die Generierung individualisierter Hotel- oder Mietwagen-Angebote (Synergien zwischen Unternehmen eines Wertschöpfungsnetzwerks).

Insbesondere die Effektivitätsziele der Kundenstrategie sollten Synergien adressieren und damit den ersten Teilaspekt des Informations-Infrastruktur-Globalziels „Bereitstellung der **richtigen Informationen** in der **richtigen Qualität** (z. B. rechtzeitig, korrekt) und **in wirtschaftlicher Weise**, um alle relevanten Informationsbedarfe zu befriedigen, für deren Befriedigung Daten aus anderen Einheiten benötigt werden“ umsetzen. Die Operationalisierung der Effektivitätsziele erfolgt auf der Grundlage von Informationsbedarfsanalysen (Winter und Strauch 2003).

Die Effizienzziele der Lieferanten- und Transformationsstrategie adressieren hingegen primär den zweiten Teilaspekt des Informationsinfrastruktur-Globalziels, die Sicherung der Leistungsqualität. Zur Operationalisierung existiert umfangreiches Schriftgut zum Qualitätsmanagement in der Informationslogistik (z. B. Otto et al. 2008).

Nicht zu vernachlässigen ist die Zielbildung und -operationalisierung der Koordinationsfunktion („Informationsinfrastruktur-Governance“). Steuerungsziele können Wertschöpfungsziele (Bereitstellungsvolumen, Bereitstellungskosten), Kundenziele (Transparenz, Zufriedenheit, Abdeckung), Produktionsziele (Standardisierung, Flexibilität), Innovationsziele und nicht zuletzt Mitarbeiterziele umfassen (siehe z. B. Sommer et al. 2008).

5.2.3 Informationsinfrastruktur-Organisation

Auf Grundlage des auf Strategieebene festgelegten Leistungsprogramms und Zielsystems werden auf Organisationsebene Ablauf- und Aufbauorganisation abgeleitet. Dabei wird ein Unternehmen nur in den seltensten Fällen „bei Null“ anfangen und entsprechende Strukturierungen völlig neu erarbeiten. Einerseits gibt es im Normalfall bereits Prozesse und Organisationseinheiten, die übergreifende Datenflüsse bewirtschaften; andererseits sind in vielen Unternehmen konvergierende Strukturierungen zu beobachten, die im Sinne von „Referenzlösungen“ situationsspezifisch adaptiert werden können. Klesse und Winter (2007) beschreiben vier Grundmodelle der Organisationsgestaltung für (im Normalfall interne) Datenintegrationsdienstleister:

- 1. **Full Service Provider** sind sowohl bei Informationsnutzung wie auch bei Plattformbetrieb und Datenaufbereitung/Informationsbereitstellung massgeblich involviert. Sie sind jedoch als „AIS-Serviceprovider“ normalerweise nicht in IT-Betrieb und fachliche Analysen involviert.
- 2. **Business Service Provider** sind bei der Informationsnutzung massgeblich involviert, während sie bei Datenaufbereitung/Informationsbereitstellung und insbesondere beim Plattformbetrieb nur am Rande involviert sind. Der Schwerpunkt sind also fachnahe AIS-Services.
- 3. **IT Platform Provider** sind insbesondere beim Plattformbetrieb, aber auch bei der Datenaufbereitung/Informationsbereitstellung massgeblich involviert, während sie bei der Informationsnutzung nur am Rande eine Rolle spielen.
- 4. **Competence Center (CC)** sind bei keiner der AIS-Hauptaktivitäten „im Lead“, sondern sollen die jeweils verantwortlichen Einheiten inhaltlich und konzeptionell unterstützen. In Unternehmen mit AIS-CCs wird der Hauptanteil des AIS-Betriebs von Fachbereichen und IT wahrgenommen, während das AIS-CC lediglich koordinierende Funktion hat und sich „Enabler“ versteht.

Die unterschiedliche Ausrichtung dieser Grundmodelle wird in Abb. 5.7 illustriert.

Plattformbetrieb	Datenaufbereitung/ Informations- bereitstellung	Informations- nutzung
Full Service Provider		
		Business Service Provider
IT Platform Provider		
Competence Center (rein koordinierende Funktion)		

Abb. 5.7 Ausrichtung der vier Grundmodelle für die Organisation von Informationsinfrastruktur-Betrieb

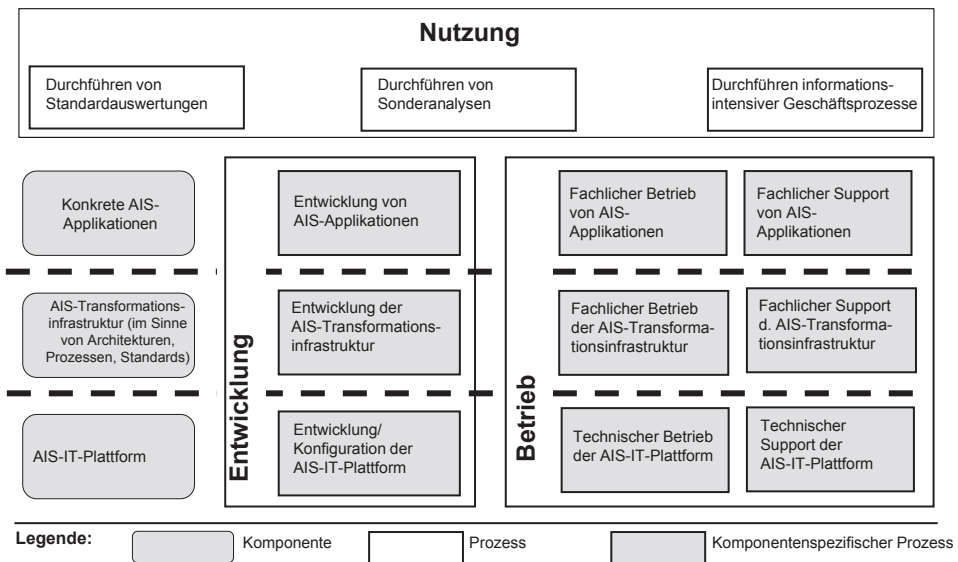


Abb. 5.8 : IIL-Prozesslandkarte. (in Anlehnung an Klesse und Winter 2007)

Welche Organisationsform im Einzelfall von Unternehmen tatsächlich gewählt wird, hängt übrigens weder von der Branche noch von der Unternehmensgröße ab. Es scheint vielmehr so zu sein, dass Organisationen das Grundmodell gewählt haben, das zur Zeit der unternehmensweiten AIS-Einführung dominant war und dass es eine Tendenz weg von den drei anderen Grundmodellen hin zum Competence Center gibt (Klesse und Winter 2007). Diese Tendenz steht in Einklang mit der Industrialisierung (d. h. zunehmender Arbeitsteilung/Spezialisierung und Kooperation) in anderen Bereichen des Informationsmanagements, wie sie z. B. bei (Zarnekow und Brenner 2003) beschrieben wird.

Standardisierung und Arbeitsteilung sind nur möglich, wenn einheitliche oder zumindest kompatible Vorstellungen über die AIS-Ablauf- und -Aufbauorganisation bestehen. Die Standardisierung anderer Bereiche des Informationsmanagements, z. B. auf ITIL-Grundlage, wie auch die mittlerweile mindestens 15-jährige Erfahrung der meisten Unternehmen sind hier hilfreich. Die Unterscheidung der Managementaufgaben Betrieb („Run Information Infrastructure“), Entwicklung („Change Information Infrastructure“) und Nutzungsunterstützung („Use Information Infrastructure“) einerseits sowie der AIS-Komponenten „Applikationen“ (z. B. Analyse, Reporting-Anwendungen), „Transformations-Infrastruktur“ (z. B. ETL-Prozesse, Datenqualitätsmanagement) und „IT-Plattform“ (z. B. Data Warehouses) andererseits führt zu der in Abb. 5.8 illustrierten Strukturierung, die als Grundlage sowohl für eine Prozesslandkarte (Ablauforganisation) wie auch eine entsprechende Organisationsstruktur dienen kann. Eine solche Strukturierung ist nicht nur zwischen verschiedenen Unternehmensbereichen, sondern auch unternehmensübergreifend als gemeinsame Grundlage nutzbar.

Von Unternehmensberatungen und Verbänden, aber auch aus der Forschung liegen für verschiedene IIL-Detailprozesse und -strukturen Referenzmodelle vor (z. B. Auth 2003; Herrmann 2006). AIS-Organisationsgestaltung kann auf dieser Grundlage weniger als „Gestaltung“, sondern mehr als Selektion und Adaption verfügbarer Betriebs-, Prozess- und Struktur-Referenzmodelle angesehen werden.

5.2.4 Finanzielle Aspekte der Informationsinfrastruktur

Die Informations-Infrastruktur produziert keine signifikanten Anteile von Daten und nutzt keine signifikanten Anteile von Informationen. Im Sinne des SCOR-Modells haben wir sie deshalb ja als Lieferketten-„Hub“ positioniert. Dies hat auch Implikationen für die finanzielle Planung und Steuerung von AIS. Jede Komponente einer Lieferkette muss in der Lage sein, Mehrwert zu erzeugen, der ihre Existenz rechtfertigt. Über den generierten Mehrwert sollen die Aktivitäten wertorientiert gesteuert werden können.

Die Bestimmung des Informationsinfrastruktur-Mehrwerts ist schwieriger als für andere Komponenten des Informationsmanagements, da Nutzenaspekte nur durch die (Informations-)Kunden in Form ihres Mehrwerts für die Entscheidungsfindung bewertet werden können, während Kostenaspekte durch (Daten-)Lieferanten und den Transformationsprozess der Integrationsinstanz determiniert werden (Jung und Winter 2000a) – was zu unterschiedlichen Zuordnungsobjekten und -verfahren führt. Andere „Veredler“ in Wertschöpfungsketten sehen sich jedoch einer durchaus ähnlichen Situation gegenüber, so dass hier neben mehr als fünfzehn Jahren AIS-Erfahrung auch auf Erfahrungen aus anderen IS-Infrastrukturbereichen zurückgegriffen werden kann.

Eine dieser Erfahrungen ist, dass der relativ höhere Infrastruktur-Investitionsanteil früherer Projekte zu einer gewissen Pionier-Unwilligkeit führt und nur durch die Einführung eines Investitionsmodells behoben werden kann, das frühe Investitionen/Projekte auf Kosten späterer subventioniert, welche die dann ausgebaute Infrastruktur nutzen. Diese Erkenntnis muss in ein AIS-Infrastruktur-Investitionsmodell umgesetzt werden, für das im Normalfall die Unternehmensleitung zentrale Mittel bereitstellen muss. Für ähnlich gelagerte Infrastruktur-Investitionsprobleme in Netzwerken empfiehlt (Schober 1999) faire gemeinsame Investitionsstrategien der Partner (in diesem Fall Datenlieferanten, Datenintegrationsdienstleister und Informationsnutzer) auf Grundlage eines „gegenseitige[n] Verständnis[es] für die Lage des anderen Partners“.

Ross und Beath (2002) empfehlen, bei Infrastruktur-Investitionen zwischen Maßnahmen mit transformierendem und solchen mit erhaltendem Charakter zu unterscheiden. Für transformierende Projekte empfehlen sie die sog. „Executive Level Allocation“, d. h. die Kostenübernahme durch Projekt-Sponsoren mit übergreifender Verantwortung (z. B. Unternehmens- oder Konzernleitung), weil die Infrastrukturkosten die lokalen (Projekt-) Budgets übersteigen und verschiedene lokale Einheiten betreffen. Die Wirtschaftlichkeit transformierender Projekte ist dabei oft nicht durch „Business Cases“ allein zu rechtfertigen, da eben nicht nur substitutive, sondern auch innovative und/oder strategische Nutz-

effekte erreicht werden (zur Unterscheidung siehe Nagel 1990). Im Gegensatz dazu haben erhaltende bzw. erneuernde Projekte substitutiven Charakter, so dass klassische Business Cases gerechnet werden können und im Normalfall lokale (Projekt-)Budgets ausreichen sollten; Die Nutzeneffekte sind in solchen Fällen lokal bewertbar, was sie von transformierenden Projekten unterscheidet, deren Wertbeitrag aufgrund der Projektkomplexität nicht erfasst werden kann.

Zu berücksichtigen ist bei der Informations-Infrastruktur natürlich immer wieder, dass das Ganze mehr wert sein muss als die Summe seiner Teile. Über die Mehrwerte der einzelnen AIS-Entwicklungsprojekte hinaus müssen deshalb auch Synergien bewertet werden. Für Informations-Infrastrukturen lassen sich auf Grundlage der allgemeinen Überlegungen von (Ansoff 1965) Synergieeffekte z. B. in Form des vereinfachten Transfers von Wissen zwischen Organisationseinheiten oder in Form der Vermeidung redundanter Mehrfach-Informationsaufbereitungen realisieren – diese Effekte sind der Kern der ökonomischen Rechtfertigung von Informations-Infrastrukturen. Diese ist allerdings aufgrund des übergreifenden Charakters meist nur mindestens mittelfristig und immer nur auf Programmebene möglich, da die Synergien eben gerade nicht einzelnen „Change“-Projekt zugerechnet werden können.

Nicht nur die Bewertung des Aufbaus bzw. der Weiterentwicklung von Informations-Infrastrukturen und die Zurechnung dieser Werte zu Vorhaben ist schwierig – auch beim Betrieb dieser Infrastrukturen werden Mehrwerte geschaffen, die sich nur schwer quantifizieren und den nutzenden bzw. bereitstellenden Einheiten zurechnen lassen. Die Veredelung von (aus dem operativem Geschäft verfügbaren und externen) Daten in Informationen berührt auch aus betriebswirtschaftlicher Sicht viele offene ökonomische Fragen – z. B. die Tatsache, dass sich bei Nutzung einer Infrastruktur, die überwiegend fixe Gemeinkosten erzeugt, Nutzungskosten nicht vernünftig kalkulieren bzw. begründen lassen, oder die Tatsache, dass entscheidungsorientiert die erste Informationsnutzung ganz anders bewertet werden müsste wie die zweite. Klesse (2007) diskutiert diese Probleme ausführlich für Data Warehousing und schlägt Lösungsansätze vor, deren Übertragbarkeit auf Informations-Infrastruktur zu prüfen wäre. Unter anderem differenziert er zwischen (a) für die Informationsnutzer sichtbaren Leistungskomponenten und (b) dem internen Leistungs- und Kostenmodell des Leistungserbringers. Innerhalb (a) wird unterschieden nach den Informationsprodukten und -dienstleistungen sowie den nicht einem bestimmten Informationsprodukt zuordenbaren Kundendienstleistungen wie bspw. allgemeinem Support. Für (b) unterscheidet (Klesse 2007) zwischen applikationsbezogenen Prozessleistungen sowie Plattformleistungen (zu denen Infrastruktur-Produktkomponenten gezählt werden). Er zeigt, wie das interne Kosten- und Leistungsmodell (b) in ein Preismodell überführt wird und zusammen mit dem aus den sichtbaren Leistungskomponenten resultierenden Abrechnungsmodell (a) in einer Kosten- und Leistungsrechnung abgebildet werden kann.

Darüber hinaus finden sich in der Controllingliteratur und entsprechenden Lehrbüchern verschiedene generische Verrechnungsmethoden, z. B. Umlageverfahren anhand von Schlüsseln, Leistungsverrechnung mit internen Preisen oder Prozesskostenrechnung, die allerdings die Besonderheiten von AIS nicht berücksichtigen. Bei der Gestaltung eines

Verrechnungsmodells für Informations-Infrastrukturen sollte u. a. bedacht werden, dass die Nutzung im Gegensatz zu transaktionalen Systemen meist auf freiwilliger Basis erfolgt, der Mehrwert der Informations-Infrastruktur nicht deren Nutzung, sondern die Nutzung der daraus erhaltenen Information ist, und dass eine Preisbildung für das Produkt „Information“ unter objektiven Gesichtspunkten schwierig ist. Im Allgemeinen sollte ein Verrechnungsansatz so gewählt werden, dass er die Nutzung und die Wertgenerierung der Informations-Infrastruktur nicht (z. B. durch hohe Verrechnungspreise) negativ beeinflusst, sondern idealerweise sogar fördert. Ein weiteres Prinzip für die Umsetzung eines Verrechnungsmodells für Informations-Infrastrukturen ist, dass die verrechneten Kosten für den Leistungsempfänger nachvollziehbar und steuerbar sein müssen, denn nur dadurch kann sowohl Akzeptanz wie auch die beabsichtigte Steuerungswirkung erreicht werden. Darüber hinaus muss ein Verrechnungsansatz für AIS in das Controllingsystem des Gesamtunternehmens eingebettet werden, weswegen bei der optimalen Gestaltung des Verrechnungsansatzes neben den Besonderheiten von AIS auch die Eignung als Teil des „großen Ganzen“ beachtet werden muss.

5.2.5 IT/Business Alignment für die Informations-Infrastruktur

Mit zunehmender AIS-Maturität steigt der Anteil von adaptierten und integrierten Standardlösungen und sinkt entsprechend der Anteil individuell entwickelter Softwarelösungen. Die Umsetzung von Fachkonzepten durch Adaption und Integration von Standardlösungen bringt andere Herausforderungen mit sich als die individuelle Umsetzung von Fachkonzepten; außerdem erzeugen Innovations- und Releasezyklen sowie Marktentwicklungen im Bereich der Standardlösungen eine Dynamik, die es in der Individualentwicklung so nicht gibt. Da sich AIS in dieser Hinsicht nicht von operativen Softwarelösungen unterscheidet, werden für IT/Business Alignment für die Informations-Infrastruktur auch die gleichen Ansätze vorgeschlagen.

Um die Passung von IT-Lösungen und fachlichen Festlegungen angesichts fachlich getriebener Innovationsprojekte, technisch getriebener Innovationsprojekte sowie „Housekeeping“- bzw. Reorganisationsprojekten immer wieder herzustellen bzw. zu verbessern, ist im Business Engineering mit der Alignmentebene eine spezielle Architekturebene vorgesehen. Auf der Alignmentebene werden Strukturen abgebildet, die „zwischen“ fachlichen Festlegungen und Softwarelösungen stehen und diese bei entsprechender Gestaltung und Bewirtschaftung entkoppeln (Aier und Winter 2009). Ein traditioneller Artefakttyp, der auf die Alignmentebene gehört, sind Applikationen. Applikationen sind Cluster von Zuordnungen zwischen IT-Funktionalitäten und Datenstrukturen auf der einen Seite sowie fachlichen Funktionen und Informationsstrukturen auf der anderen Seite. Dieser Applikationsbegriff ist nicht zu verwechseln mit dem umgangssprachlichen Begriff, der sich in vielen Unternehmen findet und Applikationen (aus fachlicher Sicht) mit IT-Softwaresystemen mehr oder weniger gleichsetzt.

Applikationslandschaften liefern auf aggregierter Ebene schon von Beginn an wertvolle Unterstützung für die Zuordnung von Softwareartefakten (z. B. IT-Funktionalitäten und Datenstrukturen) zu organisatorischen Artefakten (z. B. Aktivitäten, Rollen, Informationsobjekten). Bei entsprechender Gestaltung können Applikationen in beiden Richtungen als Puffer wirken: Nicht jede Änderung fachlicher Festlegungen muss unbedingt zu einer Änderung auf Softwareebene führen (z. B. können Standardlösungen anders konfiguriert oder anders integriert werden) und nicht jede Änderung von Softwarelösungen muss unbedingt zu einer Änderung fachlicher Strukturen führen (z. B. kann die Änderung aufgrund der geringeren Granularität und der Umsetzungsunabhängigkeit der Alignmentebene aufgefangen werden).

Auf (Referenz-)Architekturebene ist eine Zuordnung bestimmter Typen von Integrations-Softwaresystemen zu bestimmten Typen fachlicher Integrationsanforderungen möglich (Jung 2006). Die Applikationsarchitektur (auf der Alignmentebene) kann somit nicht nur zur Entkopplung dienen, sondern bei fundamentalen Entwicklungsprojekten auch Anhaltspunkte für geeignete oder weniger geeignete Umsetzungsformen aufzeigen.

Bei detaillierterer Betrachtung (als für die Applikationslandschaft) kann die Zuordnung von Softwarefunktionalitäten zu fachlichen Anforderungen durch fachliche Services dokumentiert und weiterentwickelt werden. Auch wenn aufgrund der für AIS inhärenten Kopplung durch Daten die Potenziale serviceorientierter Gestaltung begrenzt sind, finden sich Überlegungen zu fachlichen Services in der Informationslogistik z. B. bei (Dinter 2008).

Bei größerer Betrachtung (als für die Applikationslandschaft) kann die Zuordnung von Softwaresystemen zu fachlichen Strukturen durch Domänen dokumentiert werden. Domänenmodelle bilden eine wichtige Grundlage, um die organisatorische Verankerung von Datenstrukturen, -flüssen und -verantwortlichkeiten auf unternehmens- oder bereichsweiter Ebene sicherzustellen und entsprechende Analysen für die Weiterentwicklung des IT/Business Alignment zu ermöglichen.

Nicht vergessen werden sollte insgesamt der Charakter von „Alignment“: Hier geht es darum, sich tendenziell immer wieder auseinander entwickelnde Strukturen (z. B. aufgrund unzureichender Governance oder durch Zukauf nicht ausreichend anpassbarer Lösungen) auf fachlicher und IT-Ebene immer wieder miteinander kohärenter zu machen. Solches „Kohärenzmanagement“ wird oft durch die Funktion „Unternehmensarchitekturmanagement“ verantwortet. Auch wenn die wesentlichen Aktivitäten auf Alignmentebene im Kontext „Change“ notwendig und ausgelöst werden, hat Architekturmanagement auch einen kleineren „Housekeeping“-Anteil, der auch im Tagesgeschäft kleinere Anpassungen vornimmt bzw. ein Auseinander-Entwickeln zu verhindern versucht.

5.3 Analytik aus Managementsicht

Dieser Abschnitt betrachtet das Management der Analytik und bezieht sich damit ebenfalls auf die in Abschn. 5.1.5 skizzierten Gestaltungsaufgaben. Die Diskussion in Abschn. 5.3.1, inwieweit sich der (Erst-)Aufbau, die Veränderung und der Betrieb von Analytik konzeptionell

tionell unterscheiden, führt zu anderen Ergebnissen als für die Informations-Infrastruktur. Da es sich beim Prozess- und Nutzungsfokus von Analytik um ein für traditionell daten- bzw. informationsfokussierte AIS vergleichsweise neues Konzept handelt, werden in Abschn. 5.3.2 die Anwendungspotenziale und der Wertbeitrag skizziert. In Abschn. 5.3.3 wird gezeigt, in welchen typischen Formen und in welchen Entwicklungsschritten Analytik realisiert wird. Abschließend werden in Abschn. 5.3.4 Fragen des IT/Business Alignment aus Analytik-Sicht angerissen.

5.3.1 Projekt- vs. Betriebssicht

Aus der prozess- und nutzungsorientierten Sichtweise von Analytik resultiert ein fundamentaler Unterschied zu Informations-Infrastruktur hinsichtlich der Unterscheidung zwischen „Change“ und „Run“: Analytik erfolgt nicht nur bei Neu- oder Umgestaltung projektweise und kann damit separiert voneinander betrachtet werden; auch im Betrieb steht nicht „das Ganze“ im Mittelpunkt, sondern die jeweils Prozess-, Aufgaben- und vielleicht sogar Nutzer-spezifisch erreichten Nutzeneffekte.

Für das Analytik-Management folgt aus dieser Betrachtung, dass sowohl die Neu- bzw. Umgestaltung wie auch der Betrieb als Portfolio von Einzelprojekten bewirtschaftet werden können. Zwar entstehen durch die gemeinsame Nutzung bestimmter Daten mittelbare Kopplungen. Doch ist im Prinzip – im Gegensatz zu Informations-Infrastruktur – die gesamthafte Analytik eben genau nicht mehr als die Summe ihrer Einzelkomponenten; gesamthafte Zusammenhänge spielen nur eine untergeordnete Rolle, wodurch Gestaltung und Management stark vereinfacht werden.

5.3.2 Anwendungspotenziale und Wertbeitrag von Analytik

Für eine enge Verzahnung von Analytik mit der Prozessausführung spricht, dass (Bucher und Dinter 2008b)

- viele Geschäfts- und Unterstützungsprozesse Daten erzeugen, die analytisch eingesetzt werden können,
- viele Geschäfts- und Unterstützungsprozesse Auswertungen bzw. Analyseergebnisse als direkten Input für die Prozessausführung benötigen und
- im Idealfall die Geschäfts- bzw. Unterstützungsprozesse direkt den geeigneten Kontext für die Datenauswertung bereitstellen.

Bucher et al. (2009) unterscheiden zwischen direkten und indirekten Nutzeffekten. In einer Untersuchung von Analytik-Anwendungsfällen identifizieren sie beschleunigte Prozessausführung, qualitative Verbesserung der Prozessleistung, Risikominimierung, Steigerung der Effizienz der Ressourcennutzung, Steigerung der Kosteneffizienz der Prozessausfüh-

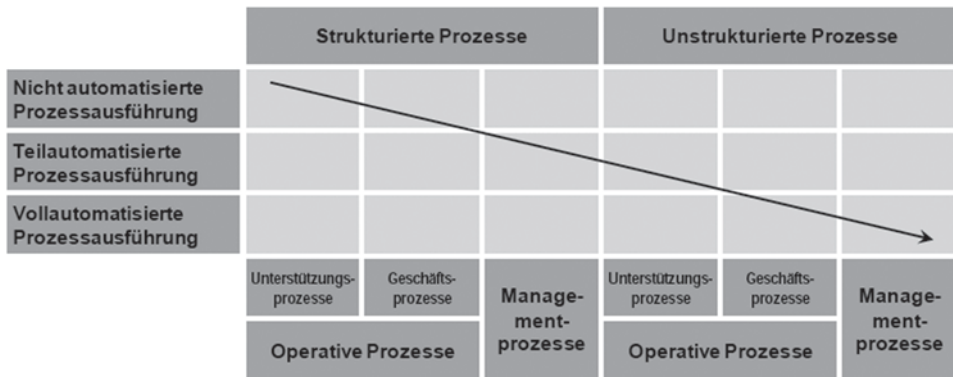


Abb. 5.9 Eignungsgrad verschiedener Prozessstypen für analytische Prozessunterstützung. (Bucher und Dinter 2008b)

rung sowie ganz allgemein eine Stärkung der Prozessorientierung. Als generelle, indirekte Nutzeffekte werden durch die betreffenden Unternehmen Leistungsverbesserungen für unternehmensexterne und -interne Stakeholder, Steigerung von Kundenzufriedenheit, Kundenprofitabilität und Kundenloyalität sowie Steigerung der Mitarbeiterzufriedenheit mit den Prozessen angegeben. Gesamthaft kann empirisch eine Korrelation zwischen dem Ausmaß analytischer Prozessunterstützung und – zumindest den indirekten – Nutzeffekten gezeigt werden (Bucher und Dinter 2008a). Natürlich bringt Analytik auch Kosten mit sich. Diese fallen jedoch „näher“ an den Nutzungsprozessen an und können deshalb den Anwendungsszenarien viel direkter und zeitnäher zugeordnet werden als alles, was in direktem Zusammenhang mit der Informations-Infrastruktur steht.

Allerdings sind nicht alle Geschäfts- und Unterstützungsprozesse gleich gut für Analytik geeignet. Abbildung 5.9 zeigt nach (Bucher und Dinter 2008b) einen von links oben nach rechts unten verlaufenden Pfeil, in dessen Richtung die Eignung für Analytik abnimmt. Das höchste Potenzial haben strukturierte, nicht automatisierte Unterstützungs- und Geschäftsprozesse, während unstrukturierte Führungs-/Managementprozesse das geringste Potenzial aufweisen. Das liegt daran, dass mit zunehmender Strukturiertheit Analytik gezielter spezifiziert und wirksamer integriert werden kann. Außerdem entfaltet Analytik „den größten Nutzen, wenn sie zur Unterstützung, Begründung und Rechtfertigung von Prozess-inhärenten Entscheidungen eingesetzt wird, die von Menschen im Kontext der Prozessausführung auf Grundlage von aufbereiteten Informationen getroffen werden“ (Bucher und Dinter 2008b).

Als Analytik-Beispiele beschreiben Bucher und Dinter (2008b) im Detail die Generierung von Last-Minute-Angeboten bei einem Reiseveranstalter sowie die Überprüfung von Zahlungsverkehrstransaktionen hinsichtlich Geldwäscheverdacht bei einer Bank. Beide Prozesse sind stark strukturiert (Entscheidungsregeln) und datenorientiert, können jedoch aus verschiedenen Gründen nicht voll automatisiert werden und haben große Bedeutung für das jeweilige Unternehmen (Compliance-Vorschrift bzw. Wertbeitrag). Gemäß Abb. 5.9 liegt also hohe Eignung für analytische Prozessunterstützung vor.

	Unterstützungs- prozesse	Geschäfts- prozesse	Führungs- prozesse
Nicht automatisierte Prozesse	<ul style="list-style-type: none"> ▪ Zentralisiertes Beschaffungsmanagement 	<ul style="list-style-type: none"> ▪ Lead- und Pipeline-Management ▪ Kampagnenmanagement ▪ Kundenbindungsmanagement ▪ Kundenberatung und Kundenbetreuung ▪ Auswirkungenanalyse bei Produktionsstörungen ▪ Gewährung von Hypotheken 	<ul style="list-style-type: none"> ▪ Budgetierung
Teilautomatisierte Prozesse	<ul style="list-style-type: none"> ▪ Ganzheitliches Zahlungsmanagement ▪ Geldwäscheerkennung im Zahlungsverkehr 	<ul style="list-style-type: none"> ▪ Erstellung von Last Minute-Reiseangeboten ▪ Steuerung des Kraftwerkeinsatzes ▪ Vertriebsportfoliomanagement im Energiehandel 	
Vollständig automatisierte Prozesse		<ul style="list-style-type: none"> ▪ Prüfung der Kreditwürdigkeit im Internethandel 	

Abb. 5.10 Analytik-Anwendungsbeispiele. (Bucher und Dinter 2008b)

Abbildung 5.10 ist eine Liste von Analytik-Anwendungsbeispielen, die für die Analyse der Anwendungspotenziale und Wertbeiträge der hier zitierten Arbeiten untersucht wurden. Ohne den Anspruch auf auch nur annähernde Vollständigkeit erheben zu können, zeigt sich in der Verteilung der Anwendungsbeispiele eine Tendenz zu nicht oder nur teilweise automatisierten Geschäfts- und allenfalls Unterstützungsprozessen, die mit den in Abb. 5.9 illustrierten Aussagen konsistent ist.

5.3.3 Realisierungsformen von Analytik

Als Bestimmungsfaktoren für die Realisierungsform „analytischer Prozessunterstützung“ werden in (Bucher und Dinter 2008b) der Maturitätsgrad der Datenbereitstellung sowie der Maturitätsgrad der Integration der Analytik in die Prozessausführung ermittelt. Diese Arbeiten wurden später in komplexeren Reifegradmodellen fortgeführt, die sich jedoch nicht auf Analytik fokussieren (Raber 2013).

Die seinerzeit untersuchten Analytik-Lösungen unterscheiden sich neben den beiden genannten Maturitätsgraden durch das Ausmaß, in dem nur Basis-Zugangs- und Basis-Analyseinstrumente eingesetzt werden oder in dem darüber hinaus auch fortgeschrittene Zugangs- und Analyseinstrumente eingesetzt werden. Als Konsequenz schlagen Bucher und Dinter (2008a) vor, die folgenden Realisierungsformen von Analytik zu unterscheiden:

1. **Informationsexperten:** die Maturität der Datenbereitstellung ist hoch, hinsichtlich der Integration von Analytik sind die Lösungen jedoch erst am Anfang.

2. **Integrationsexperten:** die Maturität der Integration von Analytik in die Prozessausführung ist hoch, hinsichtlich der Datenbereitstellung (z. B. Schnelligkeit, Konsistenz, Verlässlichkeit) bestehen jedoch noch Defizite.
3. **Prozessanalytikexperten:** die Maturität sowohl der Datenbereitstellung wie auch der Integration der Analytik in die Prozessausführung ist hoch.

Nur vergleichsweise wenige Unternehmen gehörten zum Zeitpunkt der Untersuchung zur Gruppe der Prozessanalytikexperten. Die meisten untersuchten Unternehmen sind entweder Informationsexperten oder noch häufiger Integrationsexperten. Der Weg zum Prozessanalytikexperten läuft also offenbar über eines (aber nicht beide) dieser Zwischenstadien.

5.3.4 IT/Business Alignment in der Analytik

Im Gegensatz zur Informations-Infrastruktur, bei der übergreifende Synergien im Mittelpunkt stehen und deshalb Applikationsarchitekturen eine zentrale Rolle für das IT/Business Alignment spielen, ist für Analytik die gegenseitige Ausrichtung von fachlichen Anforderungen und entsprechenden IT-Fähigkeiten lokaler und sollte deshalb auch feingranularer sein.

Um Unterstützungs- und Geschäftsprozesse analytisch unterstützen zu können, müssen Daten aus heterogenen Quellen in Echtzeit oder Nahe-Echtzeit integriert werden. Datenintegration, Datenanalyse und Informationsbereitstellung müssen mit minimaler Latenz und hoher Zuverlässigkeit erfolgen. Im Gegensatz zu klassischen BI-Ansätzen kommt der Stabilität und Verfügbarkeit analytischer Systeme, Prozesse und Informationen vor diesem Hintergrund eine deutlich höhere Bedeutung zu (Bucher und Dinter 2008b).

Im Mittelpunkt des IT/Business Alignment für Analytik stehen als Konsequenz nicht Applikationen oder gar Domänen, sondern fachliche Services. Fachliche Services fassen Verknüpfungen zwischen funktionalen/informativ-fachlichen Festlegungen einerseits sowie IT-Funktionalitäten (einschl. Datenzugriffen) andererseits zu lose gekoppelten Einheiten zusammen, die gröber geschnitten sein sollten als die verknüpften Strukturen Schelp und Winter 2008; Aier und Winter 2009).

Nach (Besemer 2007) und (Wu et al. 2007) können für AIS Front-end-Services und Back-end-Services unterschieden werden. Als typische Front-end-Services werden

- Abfrage- und Reporting-Services, Dashboards und Scorecard-Services,
- analytische und Data-Mining-Services und
- Alarmierungs- und Benachrichtigungs-Services

genannt; als typische Back-end-Services dagegen

- Services, die Daten aus Quellsystemen abziehen (u. U. auch Real-time-Synchronisation),

- Transformations- und Integrations-Services sowie
- Services für das Datenqualitäts- und Metadatenmanagement.

Hersteller- und Anwendungsszenario-unabhängige Referenz-Servicekataloge für Analytik haben sich noch nicht weit verbreitet, so dass die Strukturierung der Alignment-Ebene auch bei Nutzung von Standard-Werkzeugen noch Individualentwicklung darstellt. Für derartige Entwicklungsprojekte liefern umsetzungsorientierte Servicestrukturierungen der Werkzeug-Hersteller ebenso wie fachlich orientierte, funktionale Strukturierungen aus der operativen Informationsverarbeitung Anhaltspunkte. Bucher (2009) skizzieren auf dieser Grundlage die Orchestrierung sog. „composite processes“ (Ferguson 2005) für analytische Prozessunterstützung. Durch Kombination konventioneller, nicht-analytischer fachlicher Services mit analytischen Prozessunterstützungs-Services wird auf Alignment-Ebene eine Integration geschaffen, welche die beiderseitigen Anforderungen (aus fachlicher Sicht wie auch aus IT-Umsetzungssicht) transparent und vor allem änderungsfreundlich zu kombinieren verspricht.

5.4 Ausblick

AIS sind nicht mehr wie in der zweiten Hälfte der 1990er Jahre ein Innovationsthema, sondern stellen inzwischen einen etablierten Bestandteil des Informationsmanagements von Unternehmen fast jeder Größenordnung dar. Wie bei allen anderen „großen“ Innovationsthemen des Informationsmanagements (z. B. Datenintegration, Prozessintegration, Applikationsintegration, unternehmensübergreifende Integration) standen aus Managementsicht zunächst projektbezogene Gestaltungs- und Steuerungsfragen wie z. B. Projekt-Wirtschaftlichkeit oder die Gestaltung innovativer Einzellösungen im Mittelpunkt. Mit zunehmender Reife von Informations-Infrastruktur und Analytik verschiebt sich das Interesse nicht nur aus Umsetzungssicht, sondern auch aus Managementsicht in Richtung Dauerbetrieb (statt Entwicklungsprojekte), Industrialisierung (statt kreative Einzelfertigung), gesamthafte Betrachtung (statt isolierte Betrachtung) und IT/Business Alignment.

Gleichwohl ist es, im Gegensatz zu den meisten anderen Typen von Informationssystemen, speziell bei AIS so, dass Kostenverursachung und Wertbeitrag nicht nur zeitlich auseinander fallen, sondern typischerweise auch in unterschiedlichen Organisationseinheiten anfallen: Während sowohl die Infrastruktur als auch lokale Lösungen mehr oder weniger in der IT entwickelt werden, tritt der Nutzen in den verschiedenen Geschäftsbereichen auf, wenn Entscheidungsträger mit Unterstützung von AIS bessere Entscheidungen treffen als ohne.

Auch wenn sich die ersten „Betriebsmodelle“ für AIS abzeichnen, ist im Hinblick auf die strategische Ausrichtung und die Organisation des AIS-Dauerbetriebs noch viel zu tun. Breit akzeptierte Leistungs(referenz)kataloge fehlen ebenso wie situativ anpassbare Betriebs(referenz)modelle oder fachseitig akzeptierte Investitions- und Verrechnungsmodelle. Mehr noch als lokale Entscheidungsunterstützung, für die Nutzeneffekte und ein

gewisser Teil der Kosten aufgrund ihrer „Lokalität“ zeitlich und organisatorisch eher zusammenfallen, muss sich die unternehmensweite Informations-Infrastruktur deutlicher als „Veredler“ und Teil einer Informations-Lieferkette positionieren. Zu dieser Positionierung als „Business within the Business“ gehören neben den „Basics“ des Servicemanagements auch ein etablierter Strategieprozess sowie entsprechende Marketing- und Kommunikationsfähigkeiten.

Die Orientierung an Referenzlösungen und der Einsatz einer bewährten Methode sind zwar notwendige Voraussetzungen für das effektive und effiziente Management von AIS; diese Instrumente versagen aber angesichts zunehmender Professionalisierung und Differenzierung, wenn Lösungen und Gestaltungsansätze nicht situativ angepasst werden können. Um situativ anpassbare Referenzlösungen und anpassbare Gestaltungsmethode entwickeln zu können, bedarf es noch massiver Anstrengungen zur Identifikation der „richtigen“ Ausrichtungsvarianten und Kontexte von AIS. Die Dissertation von (Bucher 2009) stellt einen systematisch konstruierten Ansatz für Analytik dar und die Dissertation von (Raber 2013) enthält eine Komponente zur (beschränkten) situativen Anpassung entsprechender Reifegradmodelle. Viele weitere Arbeiten in dieser Richtung sind wünschenswert.

Danksagung Der Autor möchte Stefan Bischoff und Johannes Epple für ihre Unterstützung bei der Überarbeitung und Weiterentwicklung dieses Beitrags für die 5. Auflage danken.

Literatur

- Aier, S., Winter, R.: Virtuelle Entkopplung von fachlichen und IT-Strukturen für das IT/Business Alignment – Grundlagen, Architekturgestaltung und Umsetzung am Beispiel der Domänenbildung. *Wirtschaftsinformatik* **51**(2) 175–191 (2009)
- Anandarajan, M., Anandarajan, A., Srinivasan, C.: *Business intelligence techniques – a perspective from accounting and finance*. Springer, Berlin (2003)
- Ansoff, H.I.: *Corporate strategy: an analytic approach to business policy for growth and expansion*. McGraw-Hill, New York (1965)
- Auth, G.: *Prozessorientierte Organisation des Metadatenmanagements für Data-Warehouse-Systeme*. Universität St. Gallen, Bamberg (2003)
- Besemer, D.: Getting started now on SOA for BI. *DM Rev.* **37**, 26–27 (2007)
- Bucher, T.: *Ausrichtung der Informationslogistik auf operative Prozesse – Entwicklung und Evaluation einer situativen Methode*, Dissertation Universität St. Gallen, St. Gallen (2009)
- Bucher, T., Dinter, B.: Anwendungsfälle der Nutzung analytischer Informationen im operativen Kontext. In: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F., Picot, A., Speitkamp, B., Wolf, P. (Hrsg.) *Multikonferenz Wirtschaftsinformatik 2008 (MKWI 2008)*, S. 41–42. GITO-Verlag, München (2008a)
- Bucher, T., Dinter, B.: Process orientation of information logistics – an empirical analysis to assess benefits, design factors, and realization approaches. In: *Proceedings of HICSS-41, IEEE* (2008b).
- Bucher, T., Gericke, A., Sigg, S.: Process-centric business intelligence. *Bus. Process Manage. J.* **15**(3), 408–429 (2009)
- Buhr, W.: *What is infrastructure?* Universität Siegen (2003)

- Davis, G., Olson, M.: Management information systems – conceptual foundations, structure and development. McGraw Hill, New York (1985)
- Dinter, B.: Einsatzmöglichkeiten serviceorientierter Architekturen in der Informationslogistik. In: Töpfer, J., Winter, R. (Hrsg.) *Active Enterprise Intelligence*, S. 221–242. dpunkt.verlag, Berlin (2008)
- Dinter, B., Winter, R.: Information logistics strategy – analysis of current practices and proposal of a framework. In: *Proceedings HICSS-42*, IEEE (2009)
- Ferguson, M.: Techniques for integrating business intelligence into the enterprise – Part III. <http://www.b-eye-network.co.uk/view/3161> (2005). Zugegriffen: 1. Aug. 2015
- Gallagher, J.D.: Management information systems and the computer. American Management Association, New York (1961)
- Gartner: The Gartner glossary of information technology acronyms and terms. <http://www.gartner.com/it-glossary/> (2004). Zugegriffen: 1. Aug. 2015
- Gluchowski, P., Kemper, H.-G.: Quo Vadis Business Intelligence? *BI-Spektrum* **1**(1), 12–19 (2006)
- Gorry, G.A., Scott Morton, M.S.: A framework for management information systems. *Sloan Manage. Rev.* **13**(1), 55–70 (1971)
- Herrmann, C.: Referenzprozesse für die Wartung von Data-Warehouse-Systemen. Dissertation Universität St. Gallen, St. Gallen (2006)
- Herrmann, C., Klesse, M., Melchert, F.: Ergebnisdokumentation 3. CC BPM Workshop, Institut für Wirtschaftsinformatik, Universität St. Gallen, St. Gallen (2004)
- Holten, R.: Entwicklung von Führungsinformationssystemen. Deutscher Universitäts-Verlag, Wiesbaden (1999)
- Inmon, W.: Building the Data Warehouse, 3. Aufl. Wiley, New York (2002)
- Jung, R.: Architekturen zur Datenintegration: Gestaltungsempfehlungen auf der Basis fachkonzeptueller Anforderungen. Deutscher Universitätsverlag, Wiesbaden (2006)
- Jung, R., Winter, R.: Data Warehousing Strategie. Springer, Berlin (2000a)
- Jung, R., Winter, R.: Ökonomische Beurteilung von Entwicklungsvorhaben im Umfeld des Data Warehousing. In: Schütte, R., Rothowe, T., Holten, R. (Hrsg.) *Data Warehouse Management-handbuch – Konzepte, Software, Erfahrungen*, S. 25–37. Springer, Berlin (2000b)
- Klesse, M.: Leistungsverrechnung im Data Warehousing – Entwicklung einer Methode. Dissertation Universität St. Gallen, St. Gallen (2007)
- Klesse, M., Winter, R.: Organizational forms of data warehousing: an explorative analysis. In: *Proceedings Of HICSS-40*, IEEE (2007)
- Laudon, J., Laudon, K.: Management information systems: managing the digital firm. Prentice Hall International, New Jersey (2006)
- von Maur, E., Schelp, J., Winter, R.: Integrierte Informationslogistik – Stand und Entwicklungstendenzen. In: von Maur, E., Winter, R. (Hrsg.) *Data Warehouse Management*, S. 3–23. Springer, Berlin (2003)
- Nagel, K.: Nutzen der Informationsverarbeitung. Oldenbourg, München (1990)
- Österle, H., Winter, R.: Business Engineering. In: Österle, H., Winter, R. (Hrsg.) *Business Engineering – Auf dem Weg zum Unternehmen des Informationszeitalters*, S. 3–19. Springer, Berlin (2003a)
- Österle, H., Winter, R.: Business Engineering – Auf dem Weg zum Unternehmen des Informationszeitalters. Springer, Berlin (2003b)
- Österle, H., Brenner, W., Hilbers, K.: Unternehmensführung und Informationssystem – Der Ansatz des St. Galler Informationssystem-Managements 2 Aufl. B. G. Teubner, Stuttgart (1992)
- Otto, B., Wende, K., Schmidt, A., Hüner, K., Vogel, T.: Unternehmensweites Datenqualitätsmanagement: Ordnungsrahmen und Anwendungsbeispiele. In: Dinter, B., Winter, R. (Hrsg.) *Integrierte Informationslogistik*, S. 211–230. Springer, Berlin (2008)

- Raber, D.: Reifegradmodellbasierte Weiterentwicklung von Business Intelligence im Unternehmen. Dissertation, Universität St. Gallen (2013)
- Ramamurthy, K., Sen, A., Sinha, A.P.: An empirical investigation of the key determinants of data warehouse adoption. *Decis. Support Syst.* **44**(4), 817–841 (2008)
- Ross, J.W., Beath, C.M.: New approaches to it investment. *MIT Sloan Manage. Rev.* **43**(2), 51–59 (2002)
- Rockart, J.F., Treacy, M.E.: Executive information support systems. CISR Working Paper No. 65, Cambridge (1980)
- Rüegg-Stürm, J.: Das neue St. Galler Management-Modell – Grundkategorien einer integrierten Managementlehre – Der HSG-Ansatz, 2. Aufl. Haupt Verlag, Bern (2003)
- Schelp, J., Winter, R.: Entwurf von Anwendungssystemen und Entwurf von Enterprise Services – Ähnlichkeiten und Unterschiede. *Wirtschaftsinformatik* **50**(1), 6–15 (2008)
- Schober, F.: Kostenallokation für interorganisationale Informationssysteme. (Hrsg.) 4. Internationale Tagung Wirtschaftsinformatik 1999, Saarbrücken (1999)
- Sommer, T., Wendlandt, B., Trost, U., Zirkel, M.: Business Intelligence-Strategie bei der Volkswagen AG. In: Dinter, B., Winter, R. (Hrsg.) *Integrierte Informationslogistik*, S. 261–284. Springer, Berlin (2008)
- Supply Chain Council: Supply-chain operations reference-model. <http://www.supply-chain.org>. Zugriffen: 9. Feb. 2008 (2008)
- Szyperski, N., Klein, S.: Informationslogistik und virtuelle Organisationen – Die Wechselwirkung von Informationslogistik und Netzwerkmodellen der Unternehmung. *Die Betriebswirtschaft* **53**(2), 187–208 (1993)
- Winter, R.: Business Engineering – Betriebswirtschaftliche Konstruktionslehre und ihre Anwendung in der Informationslogistik. In: Dinter, B., Winter, R. (Hrsg.) *Integrierte Informationslogistik*, S. 17–38. Springer, Berlin (2008)
- Winter, R.: Business Engineering Navigator – Gestaltung und Analyse von Geschäftslösungen „Business-to-IT“. Springer, Berlin (2011)
- Winter, R., Strauch, B.: Demand-driven information requirements analysis in data warehousing. *J. Data Warehous.* **8**(1), 38–47 (2003)
- Winter, R., Schmaltz, M., Dinter, B., Bucher, T.: Das St. Galler Konzept der Informationslogistik. In: Dinter, B., Winter, R. (Hrsg.) *Integrierte Informationslogistik*, S. 1–16. Springer, Berlin (2008)
- Winter, R., Aier, S., Fischer, C., Gleichauf, B., Riege, C., Saat, J., Schelp, J.: Management von Integrationsprojekten. Springer, Berlin (2009)
- Winter, R., Bischoff, S., Wortmann, F.: Revolution or evolution? Reflections on in-memory appliances from an enterprise information logistics perspective. Tagung Innovative Unternehmensanwendungen mit In-Memory Data Management (IMDM 2011), Mainz, Gesellschaft für Informatik e. V. (GI) (2012)
- Wixom, B.H., Watson, H.J.: The BI-based organization. *Int. J. Bus. Intell. Res.* **1**(1), 13–28 (2010)
- Wu, L., Barash, G., Bartolini, C.: A service-oriented architecture for business intelligence. Service-Oriented Computing and Applications, SOCA '07, IEEE International Conference, Newport Beach, CA, S. 279–285 (2007)
- Zarnekow, R., Brenner, W.: Auf dem Weg zu einem produkt- und dienstleistungsorientierten IT-Management. *HMD – Praxis Wirtschaftsinformatik* **40**(232), 7–16 (2003)

Carsten Bange

Inhaltsverzeichnis

6.1 Einsatzbereiche und Werkzeugkategorien 98

6.2 Marktentwicklung 100

6.2.1 Trends im Markt für Datenmanagement 104

6.2.2 Trends im Markt für BI-Anwenderwerkzeuge 105

6.3 Datenintegration und -aufbereitung 105

6.4 Datenspeicherung und -bereitstellung 108

6.5 Anwendungen 111

6.5.1 Cockpits und Scorecards 112

6.5.2 Standard Reporting 114

6.5.3 Ad-hoc Reporting 117

6.5.4 Analyse 118

6.5.5 Planung und Simulation 123

6.5.6 Legale Konsolidierung 124

6.5.7 Data Mining 124

6.6 Fazit 125

Literatur 126

Zusammenfassung

Seit den 60er Jahren versuchen Unternehmen das Management mit Hilfe analytischer Informationssysteme bei der Entscheidungsfindung zu unterstützen. Häufig wechseln die Schlagworte wie Management Information Systems, Decision Support Systems oder Executive Information Systems stehen jedoch für den mäßigen Erfolg, den die Anbieter

C. Bange (✉)
Business Application Research Center, Würzburg, Deutschland
E-Mail: bange@barc.de

bis Mitte der 90er Jahre in diesem Markt erringen konnten. Erst das Data Warehouse im Kern einer mehrschichtigen Architektur zum Aufbau analytischer Informationssysteme ermöglichte einen seit vielen Jahren stabilen Architekturrahmen, der grob in Back-End Komponenten zur Datenintegration und Datenspeicherung (Datenmanagement in BI-Systemen) sowie Front-End Komponenten für verschiedene Aufgaben der Aufbereitung, Visualisierung und Verteilung von Daten strukturiert werden kann. Die Erweiterung des Einsatzfeldes analytischer Informationssysteme im Zuge der Digitalisierung von Geschäftsprozessen und -modellen führt zu einer Ergänzung des Data Warehouse Konzeptes durch neue Softwarekomponenten und einer zunehmenden Heterogenität in den Werkzeuglandschaften und Architekturen. Die grundsätzliche Strukturierung des Marktes in Datenmanagement- und Anwenderwerkzeuge bleibt aber bestehen.

Der vorliegende Beitrag gibt einen Überblick der allgemeinen Marktentwicklung und der verfügbaren Produkte und Anbieter für die verschiedenen Systemebenen analytischer Informationssysteme. Trotz voranschreitender Marktkonsolidierung existieren weltweit schätzungsweise mehr als 300 Softwareanbieter für einzelne oder mehrere Aufgaben in solch einer Architektur. Die Nennung von Angeboten kann daher schon aus Platzgründen nur exemplarisch erfolgen, umfassendere Listen und detailliertere Informationen zu Softwarewerkzeugen und Softwareauswahl sind beim Business Application Research Center (www.barc.de) verfügbar.

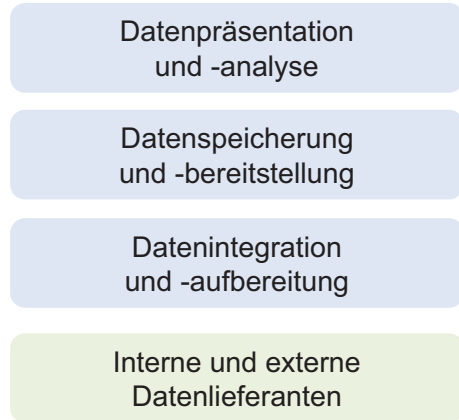
6.1 Einsatzbereiche und Werkzeugkategorien

Analytische Informationssysteme werden mit verschiedenen Softwareanwendungen für die Sammlung, Aufbereitung und Verteilung entscheidungsrelevanter Daten zur Planung, Steuerung und Kontrolle der Leistung von Organisationen aufgebaut. Als übergreifender Begriff für diese Werkzeuge hat sich „Business-Intelligence-Lösungen“ etabliert, die in einem breiten Verständnis sowohl Datenmanagement- als auch Datenpräsentations- und analysewerkzeug umfassen (Gluchowski 2001).

Die Sammlung entscheidungsrelevanter Daten stellt den Unterbau zur Datenversorgung (im „Back-End“) dar, der anspruchsvolle Anforderungen erfüllen muss:

- Daten müssen aus heterogenen, unternehmensinternen und -externen Quellen integriert werden, um eine Verknüpfung von Information zu erlauben. Operative Informationssysteme und andere interne und externe Quellen liefern hier Daten über Geschäftsprozesse, Kunden, Lieferanten, Marktgeschehen und andere entscheidungsrelevante Sachverhalte.
- Daten müssen über einen langen Zeithorizont gespeichert werden, um Trends erkennen zu können und Vorhersagen zu berechnen. Hieraus resultieren Fragen der Historisierung von Daten und Strukturen.
- Daten sollen sowohl detailliert als auch aggregiert in einem Informationsmodell zur Verfügung stehen, das entscheidungsrelevante Sachverhalte in ihrem Kontext darstellt. Dies hat Implikationen in der Wahl des Datenmodells aber auch der physischen Umsetzung des Datenmodells in Datenbankarten (z. B. relational oder multidimensional).

Abb. 6.1 Wesentliche
Aufgabenbereiche in
Business-Intelligence-Systemen



Die wichtigsten Aufgaben für Werkzeuge im Rahmen der Datensammlung sind daher **Datenintegration** durch Extraktion und Überführung relevanter Daten zwischen den Systemkomponenten, **Transformation** von Daten in benötigte Formate und Inhalte und **Speicherung** von Daten.

Die Anwenderwerkzeuge bzw. Anwendungen übernehmen (im „Front-End“) auf dieser Grundlage Aufgaben der Bereitstellung, Präsentation und Verteilung von Daten in verschiedener Form (z. B. als Bericht oder Analysetabelle) und stellen bestimmte Funktionen zur Weiterverarbeitung bereit (z. B. Analyse, Planung oder Mustererkennung durch Data Mining).

Die angesprochenen Aufgaben lassen sich anschaulich und vereinfachend in die in Abb. 6.1 gezeigten Segmente gruppieren.

Wichtige **Querschnittsaufgaben** in analytischen Informationssystemen sind das Metadatenmanagement (Datendokumentation) sowie die Datenqualitätssicherung.

Metadaten liefern sowohl eine betriebswirtschaftlich-semantische als auch eine technisch-strukturelle Beschreibung der Daten. Sie dienen der Dokumentation der Informationsobjekte, insbesondere hinsichtlich ihrer Speicherparameter, Herkunft, Struktur, Zusammensetzung und inhaltlichen Beschreibung. Neben betriebswirtschaftlich relevanten Fragestellungen wie die Definition oder Kalkulationsvorschrift von Kennzahlen können so auch technische Informationen wie Aktualisierungszeitpunkt oder Datenherkunft eines Reports bereitgestellt werden. Metadaten finden sich in den mehr oder weniger offenen Metadatenspeichern (Repositories) der beteiligten Werkzeuge. Zur Zusammenführung sind Repository-Werkzeuge interessant, die als zentrale Speichersysteme für Metadaten angeboten werden, sowie Datenintegrationswerkzeuge, die Metadaten aus Vorsystemen extrahieren und Transformations- und Überführungsregeln dokumentieren.

Mangelnde **Datenqualität** ist ein wesentliches Problem in analytischen Informationssystemen, da die Qualität von Daten in operativen Systemen für die besonderen Auswertungszwecke häufig nicht ausreicht. Diese Qualitätsmängel umfassen fehlende, mehrfach

vorkommende, falsch verknüpfte, falsch definierte und natürlich auch einfach inhaltlich falsche Daten (Wang und Strong 1996). Mängel in der Datenqualität treten als kritisches Element wegen der höheren Anforderungen häufig erst in analytischen Informationssystemen zu Tage. Das Data Warehouse ist daher regelmäßig der Ort mit der höchsten Datenqualität im Unternehmen. Eine Adressierung des Problems ist besonders im Prozessschritt der Datenintegration, also an der Schnittstelle zwischen operativen und dispositiven Systemen, von hoher Relevanz. Mangelnde Datenqualität betrifft sowohl Stamm- als auch Bewegungsdaten. Der Großteil der Stammdaten wird lediglich aus den Vorsystemen übernommen und im analytischen System angereichert, z. B. um Gruppierungs- und Hierarchieinformationen zur Aggregation. Fehler in den Vorsystemen oder nicht integrierte Stammdaten haben demnach einen stark negativen Effekt in BI-Systemen und sollten eigentlich in den Vorsystemen bereinigt oder über Stammdatenmanagement-Systeme (Master Data Management – MDM) integriert und harmonisiert werden. Defizite im Bereich der Stammdaten sind tatsächlich häufig anzutreffen und müssen im Rahmen der analytischen Informationsverarbeitung korrigiert werden. Dies ist üblicherweise dann ein wesentlicher Teil des Datenqualitätsmanagements.

Auf Basis der vorgestellten Systemarchitektur und Aufgabenbeschreibung können Werkzeugkategorien abgeleitet werden. Das Business Application Research Center (BARC) nutzt beispielsweise ein Marktstrukturmodell zur Erhebung der Business-Intelligence-Software-Umsätze aus Lizenzverkauf und Wartungseinnahmen (). Diese setzt sich zum einen aus dem Bereich „BI-Anwenderwerkzeuge“ zusammen (mit dem Teilsegment „BI“, bestehend aus Reporting&Dashboarding, Analyse und Data-Mining-/Predictive Analytics Lösungen und dem Teilsegment „Performance Management“, bestehend aus Planungs-, Konsolidierungs- und Strategiemanagement-Lösungen). Zum anderen umfasst das Marktstrukturmodell den Bereich „BI-Datenmanagement“ (Lösungen für Datenintegration, Datenqualitätsmanagement und Stammdatenmanagement, sowie relationale oder spezielle analytische Datenbanken in Business-Intelligence-Systemen). Entsprechend beinhaltet der „BI-Gesamtmarkt“ die Summe der Umsätze in den Segmenten „BI-Anwenderwerkzeuge“ und „BI-Datenmanagement“ (Abb. 6.2).

6.2 Marktentwicklung

Business-Intelligence-Lösungen bleiben auf absehbare Zeit hoch priorisiert auf der Agenda von CIOs und Fachbereichen. Schon heute zählt BARC rund 260 Anbieter, die über 600 BI-Produkte in Deutschland anbieten, mit weiter wachsender Tendenz (BARC 2014a). Für die kommenden Jahre erwartet BARC eine weiter anhaltende Marktdynamik und eine wachsende Zahl an Anbietern im Markt. Neben Start-ups im Big-Data-Umfeld sind es insbesondere amerikanische Anbieter, die mit hohen Venture-Capital-Summen eine internationale Expansion anstreben und verstärkt in den Markt eintreten.

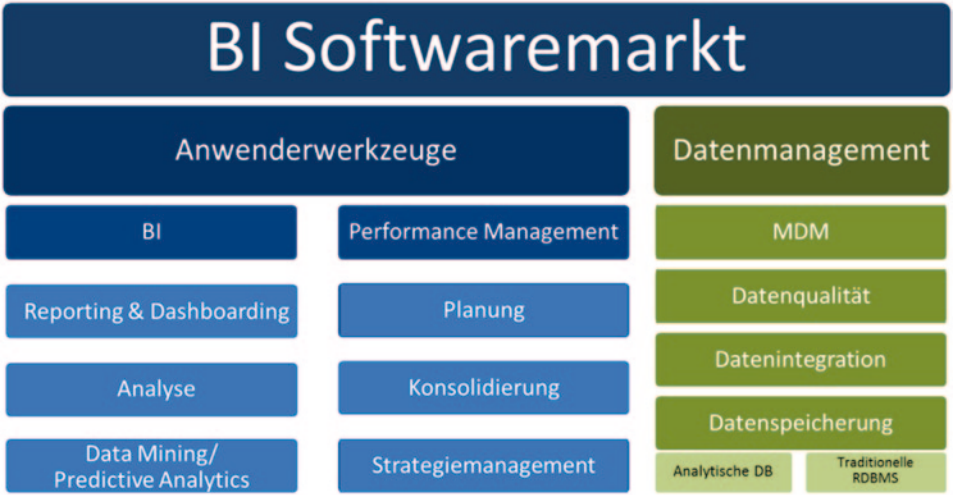


Abb. 6.2 BARC-Marktstrukturmodell des BI-Software-Marktes. (BARC 2014a)

Betrachtet man die erzielten Umsätze, zeigt sich eine seit einigen Jahren anhaltende positive Entwicklung des Gesamtmarkts für Business Intelligence in Deutschland. BARC definiert den BI-Software-Umsatz als Summe aus verkauften Lizenzen sowie Wartungs- oder Mietgebühren für BI-Werkzeuge, BI-Anwendungen und Datenmanagementkomponenten für Business-Intelligence-Lösungen in Deutschland. Entsprechend werden auch Mietmodelle oder die Bereitstellung durch einen Dienstleister (Application Service Providing oder Software as a Service) erfasst.

Die Lizenz- und Wartungsumsätze in seinen beiden Marktsegmenten „BI-Anwenderwerkzeuge“ und „BI-Datenmanagement-Werkzeuge“ und ihrer Subsegmente erhöhten sich gegenüber 2012 um insgesamt 10,8% auf 1,47 Mrd. €. Das Wachstum lag damit im Bereich der durchschnittlichen jährlichen Marktwachstumsrate (Compound Annual Growth Rate - CAGR) der letzten sieben Jahre von 11,0% (Abb. 6.3).

Bei näherer Betrachtung des BI-Gesamtmarkts zeigt sich eine erhebliche Marktverschiebung zugunsten des Datenmanagement-Segments, die sich in früheren BARC-Marktstudien bereits ankündigt hatte. Der Umsatz mit solchen Produkten stieg 2013 um ganze 20% (Vorjahr 16%), während das Wachstum im Segment der BI-Anwenderwerkzeuge nur noch drei Prozent betrug (Vorjahr 11,3%) (Abb. 6.4). Damit sind erstmals beide Marktsegmente in etwa gleich groß (49 zu 51%).

Die Datenmanagement-Anbieter profitieren derzeit offenbar vor allem von der steigenden Nachfrage nach Big-Data-Technologien, vom wachsenden Bedarf an Datenintegrationstechnik sowie speziell im SAP-Umfeld von der Nachfrage nach der Analytischen Datenbank „Hana“. Ohne die enormen Zuwächse im Backendbereich bei SAP um 47% gegenüber dem Vorjahr wäre der Gesamtmarkt lediglich um 6,6 statt 10,8% gewachsen.

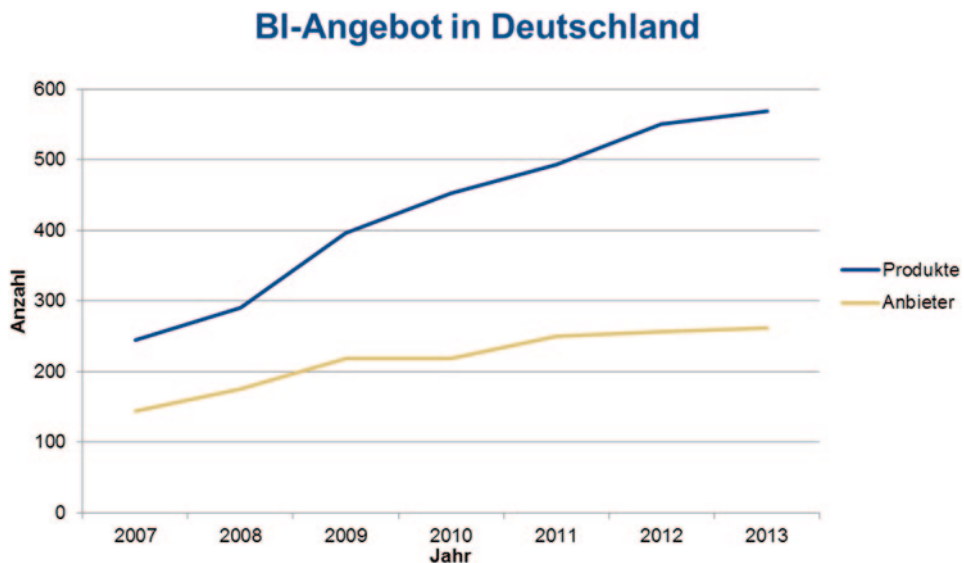


Abb. 6.3 Die Zahl der Anbieter und angebotenen BI-Produkte in Deutschland nimmt seit Jahren zu

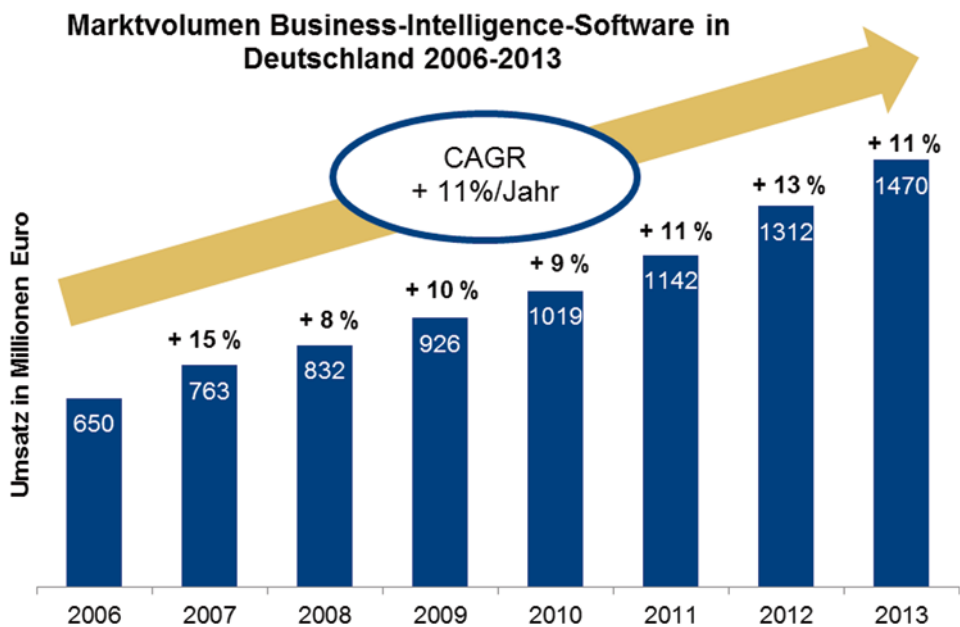


Abb. 6.4 Das Marktwachstum 2013 entspricht dem langfristigen durchschnittlichen für BI-Software seit 2006. (Bange et al. 2014)

Rang	Firma	BI & DM Software- Umsatz 2013 (in Mio. €)	Veränderung zu 2012	Marktanteil in Prozent
1	SAP	371,0	22%	25,5%
2	Oracle	180,0	6%	12,4%
3	Microsoft	154,2	16%	10,6%
4	IBM	142,2	5%	9,8%
5	SAS	109,5	-10%	7,5%
6	Informatica	51,0	13%	3,5%
7	Qlik	35,0	19%	2,4%
8	MicroStrategy	34,0	3%	2,3%
9	Teradata	30,1	-5%	2,1%
10	Software AG/ IDS Scheer	19,2	16%	1,3%

Abb. 6.5 Top 10 Anbieter im deutschen BI-Gesamtmarkt. (Summe aus BI-Anwenderwerkzeugen und Datenmanagement-Werkzeugen)

Vergleichbare Impulse scheinen dem Segment für BI-Anwenderwerkzeuge 2013 gefehlt zu haben. Immerhin konnte SAP auch hier seine Stellung als Marktführer ausbauen, reichte jedoch trotz Umsatzzuwächsen nicht an die zum Teil hohen Wachstumsraten der Konkurrenten heran. So konnten einige Anbieter ihre Einnahmen mehr als verdreifachen, allerdings ausgehend von einem wesentlich niedrigeren Umsatzniveau.

Die „Top 5“-Anbieter im BI-Gesamtmarkt bleiben weiterhin die vier Softwaregeneralisten SAP, Oracle, Microsoft und IBM sowie SAS Institute als weltweit größter BI-Spezialist. Diese Unternehmen bieten sowohl Datenmanagement- als auch BI-Anwenderwerkzeuge an („full stack“-Anbieter). Klar abgesetzt hat sich in dieser Gruppe die SAP auf der Spitzenposition mit 371 Mio. € Umsatz.

Ihre vier Verfolger erzielen zwischen 180 und 110 Mio. € Umsatz. Zusammen entfällt auf diese Fünfergruppe rund 65 % (Vorjahr: 61 %) des Marktvolumens.

Klarer Marktführer des BI-Gesamtmarktes in Deutschland ist die SAP AG (Abb. 6.2). Die Walldorfer konnten 2013 mit einem Plus von 22 % auf 371 Mio. € kräftig zulegen. Oracle konnte hingegen nur um sechs Prozent (2012: fünf Prozent) zulegen, und Microsoft schaffte einen Umsatzzuwachs in Höhe von 16 % (2012: acht Prozent).

Nach Marktanteilen konnte SAP diesen von 20 auf 25 %, Microsoft von acht auf zehn Prozent ausbauen. Oracle blieb hingegen bei zwölf Prozent, während IBM und SAS Institut Marktanteile verloren. Dank SAP vereinnahmten diese „Top-5“-Anbieter 65 % des Marktes – vier Prozent mehr als 2012 (61 %) (Abb. 6.5).

Insgesamt unterliegt die Umsatzentwicklung vieler Hersteller im BI-Gesamtmarkt starken Schwankungen von Jahr zu Jahr, wobei manche schon seit längerem unterdurchschnittlich wachsen. So konnten 2013 im BI-Gesamtmarkt 60 Anbieter einen Softwareumsatz von über einer Million Euro verbuchen, 26 dieser 60 Anbieter weisen allerdings sinkende Umsätze, kein oder nur ein leichtes (plus null bis fünf Prozent) Wachstum auf. Andererseits können Anbieter wie Splunk, Prevero, Blue Yonder, Datawatch und Jaspersoft hohe Wachstumsraten bis zu 100 % aufweisen.

6.2.1 Trends im Markt für Datenmanagement

Das Marktsegment der Datenmanagementwerkzeuge ist seit längerem durch die Konzentration auf wenige Anbieter geprägt. Diese Tendenz scheint sich weiter zu verstärken. So entfielen 2013 allein auf die „Top 5“ der Anbieter 77 % des Segmentumsatzes. Blickt man auf die „Top 10“ sind es sogar 92,5 % (2012: 90 %). Auch finden sich nur 19 Anbieter im BI-Gesamtmarkt, die Werkzeuge für das Datenmanagement zur Entscheidungsunterstützung anbieten und Umsätze von mehr als einer Million Euro erzielen.

Zugleich laufen die Geschäfte im Datenmanagement (Backend) gut und sorgten für ein Umsatzplus in 2013 um ganze 20 % gegenüber dem Vorjahr auf 722 Mio. €. Dies entspricht 49 % des BI-Gesamtmarktes in Deutschland 2013. Getragen wurde das starke Wachstum von verschiedenen Effekten:

- Anwenderunternehmen investieren nach langer Zeit der Zurückhaltung vor dieser neuen Technologie nun auch in analytische Datenbanken. Die Zuwächse in diesem Bereich bei SAP, Oracle, IBM, Microsoft, Exasol oder EMC/Pivotal sind hierfür ein guter Beleg.
- Datenmanagementprojekte sind tendenziell langfristiger angelegt und nicht so konjunkturanfällig.
- Bei wachsender Verbreitung von Self-Service-BI und der Etablierung neuer, flexiblerer Datenmanagementarchitekturen, die ggf. auch Big Data Datenquellen einbeziehen sollen, entsteht Investitionsbedarf.
- Die SAP als größtes deutsches Softwareunternehmen erzielt mit hohem Marketingaufwand und ihrer Evolution der technischen Plattform in Richtung HANA inzwischen auch signifikante Umsätze, da Kunden HANA in ersten Szenarien ausprobieren und/oder es als strategisches Werkzeug verstehen und auch in ihre Lizenzbündel mit aufnehmen. Außerdem kam mit der Übernahme von Sybase Datenmanagement-Produkte ins Portfolio. Der Marktanteil der SAP im Segment Datenmanagement wächst gegenüber dem Vorjahr von 20 % auf jetzt 24 %.

6.2.2 Trends im Markt für BI-Anwenderwerkzeuge

Im Jahr 2013 entfielen auf das Marktsegment der BI-Anwenderwerkzeuge 51 % (2012: 59 %) aller Lizenz- und Wartungsumsätze im BI-Markt in Deutschland. Das Gesamtvolumen erhöhte sich dabei von 725 Mio. auf nun 748 Mio. €. Prozentual verlangsamte sich das Wachstum erheblich: Gelang 2012 noch ein Umsatzplus bei BI-Anwenderwerkzeugen von 11,3 % waren es 2013 nur noch 3,2 %. Im gleichen Zeitraum wuchs das Segment der Datenmanagementwerkzeuge um fast 20 %.

Hier sind klare Abhängigkeiten zu sehen von der schlechten gesamtwirtschaftlichen Entwicklung Anfang 2013, die sich auf das Geschäft vieler Anbieter ausgewirkt hat und im deutlich besseren zweiten Halbjahr 2013 nicht wieder aufgeholt werden konnte. Während der konjunkturell schwachen Phasen in den letzten Jahren waren die Auswirkungen auf Investitionen in BI-Software noch geringer. Inzwischen scheinen Unternehmen Investitionen in BI-Anwenderwerkzeuge auch konjunkturabhängig zurückzuhalten. Wesentliche Gründe hierfür sind wahrscheinlich der inzwischen erreichte Grad an Reife und Verbreitung von BI-Lösungen in Unternehmen.

Im Einzelnen konnte SAP ihre Marktführerschaft für BI-Anwenderwerkzeuge weiter ausbauen und steigerte ihre Umsätze gegenüber dem Vorjahr um sechs Prozent auf 198 Mio. €. Der Marktanteil im BI-Frontend-Segment erhöhte sich so von 25 % auf jetzt 26,5 %. Anders als im Datenmanagement-Segment ist insgesamt jedoch bislang keine so starke Marktkonzentration auf einige, wenige Hersteller zu beobachten. Vielmehr finden sich 50 Anbieter mit über einer Million Euro Umsatz.

Das höchste Wachstum entfiel 2013 auf die Anbieter für Predictive Analytics und Prozess-Mining Blue Yonder und Celonis sowie den Spezialisten für die Zusammenführung, Indizierung und Auswertung semi-strukturierter Daten (insb. Log-Daten) Splunk. Sie alle konnten weit überdurchschnittliche Steigerungen melden. Starkes Wachstum zeigten auch die (Commercial) Open Source Anbieter JasperSoft (Reporting und Analyse) und Rapid-Miner (Data Mining/Predictive Analytics).

6.3 Datenintegration und -aufbereitung

50–80 % des Aufwandes eines Business-Intelligence-Projektes stecken regelmäßig in der Definition und Umsetzung der Datenintegration (DI) und des Datenqualitätsmanagements. Gleichzeitig wachsen sowohl das Datenvolumen als auch die Anzahl der zu integrierenden Datenquellen (BARC 2014b). Daher kommt diesem Prozessschritt und der eingesetzten Werkzeuge zur Effizienzsteigerung eine besondere Bedeutung zu.

Im Rahmen der Datenintegration wird die Überführung von Daten zwischen den verschiedenen Vorsystemen und den Datenhaltungskomponenten der analytischen Informationssysteme organisiert. Üblicherweise werden stündlich, täglich oder monatlich die relevanten Daten aus diversen Vorsystemen extrahiert, verändert (z. B. bereinigt oder um Information ergänzt) und in ein Data Warehouse oder andere Datenspeicher geladen.

Neben konventionellen Data-Warehouse-Systemen auf Basis von klassischen relationalen oder Analytischen Datenbanken sind hier in letzter Zeit auch Hadoop-Systeme für polystrukturierte Daten, oder anwendungs- oder werkzeugspezifische Datenspeicher (z. B. In-Memory Datenbanken von BI-Werkzeugen oder NoSQL Datenbanken für bestimmte Anwendungen) häufiger zu finden. Für die direkte Geschäftsprozess-Steuerung (operative BI) kann auch eine unmittelbare Übertragung von Daten oder Benachrichtigungen direkt nach ihrer Erzeugung im operativen System angestrebt werden („Real-Time-Integration“). Die Datenintegrationsschicht kann auch zur Rückführung von Daten aus einem Data Warehouse in operative Systeme oder zwischen Data Warehouses genutzt werden, wenn beispielsweise in einem Warehouse Plandaten gesammelt werden, die auch in anderen Systemen genutzt werden sollen.

Softwareseitig können drei Gruppen von Werkzeugen unterschieden werden, die eine Datenintegration durch Möglichkeiten der komfortablen Entwicklung, Dokumentation und Steuerung der notwendigen Prozesse unterstützen und damit deutliche Vorteile gegenüber Eigenentwicklungen aufweisen:

- **Spezialwerkzeuge** zur Datenintegration, die ausschließlich Komponenten zur Extraktion, Transformation, Qualitätssicherung und Überführung von Daten enthalten. Nach ihren Hauptaufgaben werden diese Werkzeuge auch mit den Akronymen ETL (Extraktion, Transformation, Laden) für die zyklische Übertragung und EAI (Enterprise Application Integration) für die direkte (Real-Time/Near-Time) Übertragung von Daten benannt,
- **Module in Business-Intelligence-Suiten**, die teilweise wie Spezialwerkzeuge funktionieren, teilweise aber auch dafür optimiert sind, die spezifische Datenstruktur des BI-Werkzeuges der Suite zu befüllen, und
- **Datenbank-Komponenten**, die Teil des Produktangebotes von Standard-Datenbanken sind.

Tabelle 6.1 zeigt Beispiele für Anbieter und Werkzeuge in diesen drei Kategorien.

Eine ausführlichere und jährlich aktualisierte Liste findet sich im BARC Guide Business Intelligence (BARC 2014a).

Viele der Spezialwerkzeuge für Datenintegration sind inzwischen zu Datenintegrationsplattformen herangewachsen. Sie unterstützen verschiedenste Formen der Überführung von Daten (Batch, satzweise, Streaming, Föderation), verschiedene Orte der Transformation (im Vorsystem, in der DI-Engine („ETL“) oder im Zielsystem („ELT“)), eine breite Palette von Datenspeichertechnologien (von relationalen Datenbanken bis Hadoop und NoSQL-Datenbanken) und bieten ergänzende Komponenten für Datenqualitäts- oder Stammdatenmanagement.

Zur **Verbesserung der Datenqualität** in analytischen Informationssystemen wird ein Datenqualitätsmanagement als übergreifende Aufgabe mit Einflüssen in die Organisation, IT-Architektur und Werkzeuglandschaft eingesetzt. Nachdem Anfang der 2000er Jahre viele Spezialwerkzeuge für die Überprüfung von Daten in den Vorsystemen (Data Pro-

Tab. 6.1 Datenintegrationswerkzeuge (Auswahl)

Segment	Hersteller	Produktbezeichnung
Spezialwerkzeug	Ab initio	Co>Operating system
	Aruba informatik	Enterprise query
	Attivio	Active intelligence engine
	Attunity	CloudBeam
	Comlab	ARES
	Embarcadero	DT/Studio
	Group 1 software	Sagent data flow solution
	IBM	InfoSphere information server
	Informatica	PowerCenter
	Information builders	iWay data migrator
	Javlin	CloverETL
	Kofax	Kapow enterprise big data integration platform
	Open text	Integration center
	Pentaho	Data integration (Kettle)
	SAS	Data integration server
	saxxes	SX integrator
	SnapLogic	Elastic integration platform
	Syncsort	DMExpress
	Talend	OpenStudio
	Theobald	Xtract suite
Teil einer business intelligence Suite	Actuate	BIRT designer pro
	Cubeware	Importer
	Qlik	Expressor
	Jaspersoft	JasperETL
	Jedox	Palo ETL server
	Tonbeller	Siron
Datenbank-Ergänzung	Action (Pervasive)	Data integrator
	IBM	SQL warehousing tool
	Microsoft	SQL server integration services
	Oracle	Warehouse builder, data integrator
	SAP	Data services

filing) sowie der automatisierten Bereinigung von Daten (Data Cleansing) entstanden, werden diese Aufgaben inzwischen von so gut wie allen Datenintegrationswerkzeugen mit angeboten. Beispiele für verbleibende Spezialanbieter von Datenqualitätswerkzeugen sind Ataccama, emagixx, Harte Hanks Trillium, Human Inference, Kammachi, ISO, Omikron, Tolerant, Uniserv, oder Verdantis.

Ein wachsender Bereich in der Datenintegration ist die **Föderation von Daten**, auch „Datenvirtualisierung“ genannt. Neben Möglichkeiten zur Föderation von Datenquellen in Datenintegrationsplattformen, z. B. von IBM, Informatica, Oracle oder SAP gibt es auch einige Spezialanbieter im Markt. Bekannte Anbieter sind Composite Software (inzwischen Teil von Cisco), Denodo und Data Virtuality.

Neuester Trend in der Datenintegration ist die „**Self-Service Datenintegration**“, teilweise auch „Data Wrangling“ genannt. Neben der Möglichkeit zur Datenintegration durch einen Fachanwender in vielen „Self-Service“ BI-Werkzeugen, erscheinen neue Anbieter am Markt, die Datenintegration für weniger geschulte und Entwicklungsorientierte Anwender verfügbar machen wollen. Neue und mit viel Venture Capital ausgestattete Anbieter sind beispielsweise Trifacta, Paxata und Tamr.

Für die Datenintegration gibt es einige (Commercial) **Open Source Angebote** bzw. Komponenten. Prominentestes Beispiel ist Talend, die gut über Venture Capital finanziert sind und im Rahmen ihres Commercial Open Source Models sowohl Open Source als auch lizenzkostenpflichtige Software anbieten.

6.4 Datenspeicherung und -bereitstellung

Die Speicherung entscheidungsrelevanter Daten erfolgt klassischerweise in Datenbanken, die nur zu diesem Zweck aufgebaut werden. Kernidee ist die Schaffung einer themenorientierten, integrierten, zeitbezogenen und dauerhaften Sammlung von Daten zur Entscheidungsunterstützung des Managements – dem **Data Warehouse** (Inmon 1996; Bauer und Günzel 2013). Das Konzept ist insbesondere auf die Speicherung von strukturierten Daten und ihrer Zentralisierung möglichst in einer Datenbank ausgelegt.

Data Warehouses als umfassende, unternehmensweite Datensammlung werden häufig in klassischen relationalen Datenbanken aufgebaut. Die Nutzung bestimmter technischer Möglichkeiten wie materialisierte Sichten, spaltenorientierte Indizierung oder Speicherung von Tabellen im Hauptspeicher umgehen die inhärenten Probleme des relationalen Modells mit dem Umgang mit hochdimensionalen Datenmodellen und ihrer performanten Abfrage. Teilweise ist die Nutzung der Möglichkeiten und ihre ständige Anpassung und Einstellung auch mit einem hohen Aufwand verbunden.

Daher hat sich inzwischen neben dem Einsatz **klassischer relationaler Datenbanken** die Nutzung Analytischer Datenbanken etabliert. **Analytische Datenbanken** sind für den Einsatzzweck in analytischen Informationssystemen optimiert und haben keinen Anspruch einer universellen Nutzbarkeit, z. B. auch für Transaktionssysteme. Dieser Grad der Vorbestimmtheit ermöglicht die Nutzung von Datenbankarchitekturen und Speicherformen, die für analytische Informationssysteme besonders gut geeignet sind. Vorteile im Vergleich zu klassischen relationalen Datenbanken sind insbesondere im Bereich der Abfrageperformance zu erwarten, aber auch bei Skalierbarkeit und Wartbarkeit, den weiteren wesentlichen Leistungsparameter von Datenbanken im analytischen Einsatz.

Für Data Marts als funktions- oder abteilungsspezifisch aufgebaute Datensammlungen kommen auch **multidimensionale Datenbanken** in Betracht, die vor allem Vorteile in der Geschwindigkeit der Datenbereitstellung durch Vorkalkulation und Datenhaltung im Hauptspeicher aufweisen. Multidimensionale Datenbanken haben gerade für Data Marts mit Finanzdaten und bei Planungsanwendungen eine große Bedeutung.

In den letzten Jahren ist die Zahl der BI-Werkzeuge mit eigenen, **proprietären Datenspeichern** sprunghaft angestiegen. BI-Werkzeug halten ihre eigenen Daten häufig in In-Memory Strukturen, aber bieten keine offenen Schnittstellen für eine Abfrage dieser Daten aus anderen BI-Werkzeugen an. Beispiele hierfür sind Board, CommaSoft, HumanIT, Oracle Endeca, Qlik, MicroStrategy, Tableau, Tibco Spotfire, Thinking Networks.

Die aufgebaute Datenmanagement-Infrastruktur für strukturierte Daten zielt auf ein **On-Line Analytical Processing (OLAP)** ab, also der schnellen Beantwortung analytischer Fragen. Besondere Anforderungen ergeben sich hier vor allem hinsichtlich Geschwindigkeit der Informationslieferung, Komplexität der Analysemöglichkeiten, Sicherheit und zu verarbeitende Datenmengen. Aus einer modellierungsorientierten Betrachtungsweise steht für den Anwender häufig der Aufbau von mehrdimensionalen und hierarchischen Datenmodellen (Würfel / Cubes) im Vordergrund. Dies entspricht der Sichtweise von Entscheidungsträgern auf Entscheidungssituationen, in denen vernetzte Informationen zu Objekten wie Kunden, Produkten, Verkaufsstätten oder Absatzregionen benötigt werden.

Die physische Umsetzung dieser logischen Konstrukte kann dann direkt in einer multidimensionalen Datenbank, anderen speziellen Speicherformen, durch besondere Modellierungstechniken in einer relationalen Datenbank oder in einer semantischen Metadaten-Schicht erfolgen.

Mit der wachsenden Bedeutung analytischer Informationssysteme im Zuge der Digitalisierung von Geschäftsprozessen und –modellen sollen heute aber ergänzend auch große Mengen **polystrukturierter Daten**, z. B. aus IT- und Web-Logs, Social Media oder Sensoren in Maschinen und anderen Gegenständen („Big Data“) gespeichert und für die Entscheidungsfindung herangezogen werden. Da diese Daten häufig nicht wirtschaftlich in relationalen Datenbanken gespeichert werden können, kommen immer öfter das **Hadoop** File System (HDFS) und seine ergänzenden Komponenten oder für bestimmte Datentypen oder Anwendungen auch sogenannte **NoSQL-Datenbanken** zum Einsatz. NoSQL ist ein Sammelbegriff für nicht relationale Datenbanken und umfasst ganz verschiedene Datenbankarten, z. B. Dokumenten bzw. XML-orientierte Datenbanken wie CouchDB, MarkLogik, MongoDB; Key-Value Stores wie BerkleyDB, Google Big Table, Membase, Redis; oder Graphen-Datenbanken wie Neo4j.

Die Kombination verschiedener Speichertechnologien erscheint momentan für den Aufbau eines umfassenden analytischen Informationssystems, das strukturierte und polystrukturierte Daten umfasst, notwendig. Der Aufbau eines gesamthafte, über verschiedene physische Speicher hinweg gültigen und abfragbaren logischen Modells der entscheidungsunterstützenden Daten ist hierbei eine wesentliche Herausforderung, die mit dem Konzept des „**logischen Data Warehouse**“ beschrieben wird (Tab. 6.2).

Tab. 6.2 Datenbanken für analytische Informationssysteme (Auswahl)

Speicher-konzept	Hersteller	Produkt
Klassische, relationale datenbanken	Actian	Ingres
	Firebird foundation	Firebird
	IBM	DB2
	Microsoft	SQL server
	Oracle	Database
	Oracle	MySQL
	PostgreSQL	PostgreSQL
	SAP	Sybase adaptive server enterprise (SAP Sybase ASE)
Multidimensionale datenbanken	IBM	Cognos TM1
	Infor	Infor BI OLAP server
	Jedox	Jedox OLAP server
	Microsoft	SQL server analysis services
	Oracle	Hyperion essbase
	SAS	OLAP server
Analytische Datenbanken	Actian	The actian analytics platform
	Calpont	InfiniDB
	Exasol	EXASolution
	IBM	PureData for analytics (Netezza)
	IBM	PureData system for operational analytics
	Infobright	Enterprise edition
	Kognitio	The kognitio analytical platform
	Microsoft	Analytics platform system
	Oracle	Exadata database machine
	Oracle	Times ten
	Parstream	Parstream
	Pivotal	Greenplum database
	SAND	Analytic platform
	SAP	HANA
	SAP	Sybase IQ
	SAS	SAS scaleable performance data server
	Teradata	Aster database
	Teradata	Database
	Vertica (a HP company)	Analytics platform
	Yaacomo	Yaacomo

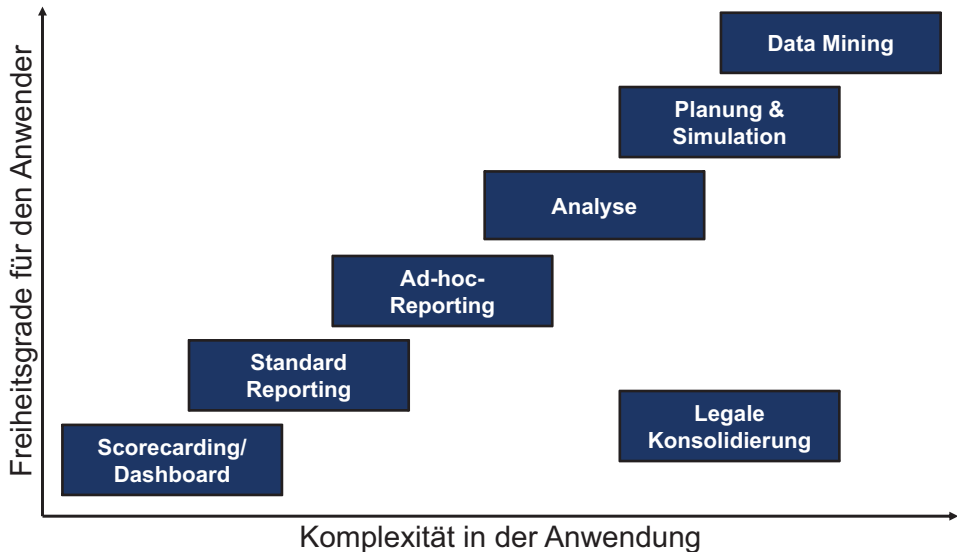


Abb. 6.6 Klassen von Business-Intelligence-Anwendungen

6.5 Anwendungen

Der Aufbau eines Data Warehouse sowie die Modellierung und Aufbereitung der Daten in den Datenmanagement-Komponenten eines Business Intelligence-Systems dient hauptsächlich dem Zweck, entscheidungsrelevante Informationen für Entscheidungsträger darzustellen und weiterzuverarbeiten. Sind Entscheidungsregeln oder Modelle definiert worden, können Entscheidungen gerade im operativen Kontext auch automatisiert werden.

Relevante Auswertungsobjekte wie Kunde, Produkt, Vertriebsregionen, Mitarbeiter, Kosten etc. werden in Business-Intelligence-Systemen für Reporting, Analyse und Planung aufbereitet, dargestellt und kommuniziert. Werkzeuge und umzusetzende Anforderungen in Business-Intelligence-Systemen sind allerdings so unterschiedlich wie der Umgang mit Kennzahlen im Unternehmen. Häufig werden die BI-Hauptanwendungen Cockpits/Scorecards, (Ad-hoc) Reporting, Analyse, Planung, Konsolidierung und Data Mining (Abb. 6.6) in getrennten Werkzeugen angeboten, um den unterschiedlichen Anwenderprofilen und der Anwendungscomplexität gerecht zu werden. Softwareseitig existieren hierfür sowohl Spezialanbieter für jeden Aufgabentyp als auch Hersteller von Business-Intelligence-Suiten, die verschiedene Anwenderwerkzeuge und teilweise auch Datenintegrations- und Datenspeicherungsmöglichkeiten mit anbieten. Nach einer Übernahmewelle im BI-Markt haben sich im Jahr 2007 auch die breit aufgestellten Softwareanbieter Oracle, IBM und SAP durch Zukäufe deutlich verstärkt.

Anwender reichen vom „Power User“, der ständig und intensiv mit BI-Werkzeugen arbeitet, bis hin zum Gelegenheitsnutzer. Eine Kategorisierung der Einsatzbereiche in sieben Klassen mit unterschiedlichen Freiheitsgraden in der Analyse und Komplexität in der Anwendung zeigt die wesentlichen Aufgaben und unterschiedlichen Nutzungsarten von Business-Intelligence-Systemen (Abb. 6.6). Sie werden im Folgenden genauer vorgestellt. Für jede Klasse wird auch eine Auswahl relevanter Anbieter genannt, die im Rahmen dieses Beitrags aber nur exemplarisch bleiben können.

Studien des Business Application Research Center (BARC) in Würzburg liefern umfangreiche Softwareanbieterlisten („BARC Guide“ (BARC 2014a)) sowie Softwarevergleiche für verschiedene Teilgebiete der Business Intelligence („BARC-Studien“) (www.barc.de). Dort wird eine sehr umfangreiche Liste aller bekannten Werkzeuge in den jeweiligen Segmenten vorgestellt.

6.5.1 Cockpits und Scorecards

Cockpits (auch Dashboards genannt) dienen der übersichtlichen und einfachen Darstellung aggregierter Information, z. B. in Unternehmensportalen oder als eigene Führungs- bzw. Managementinformationsanwendung (Abb. 6.7). Viele Anwendungen in diesem Bereich werden auf Basis von Internettechnologie zur Anzeige im Browser entwickelt

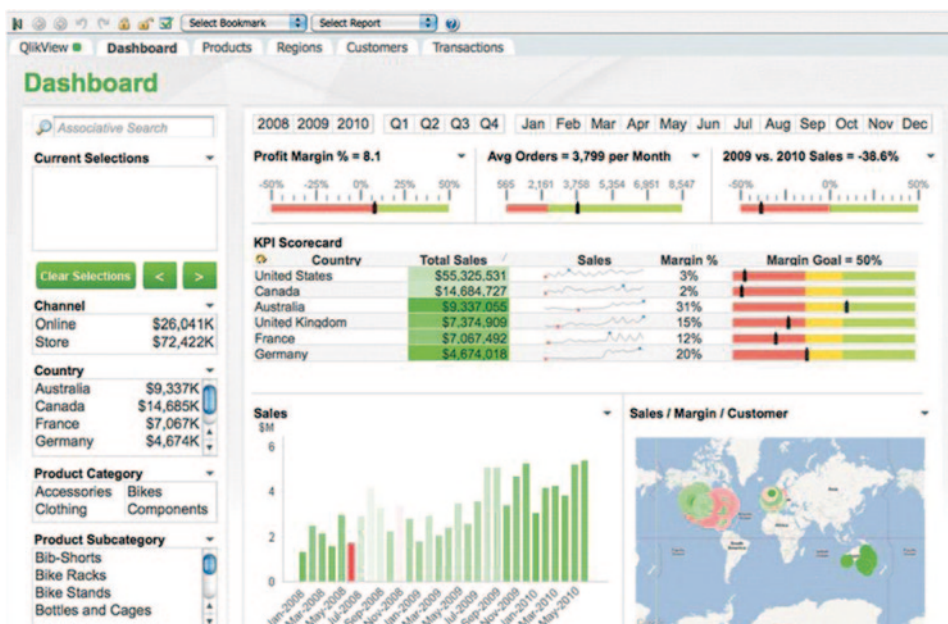


Abb. 6.7 Cockpit zur kondensierten Information des Managements mit KPIs

oder werden in speziellen Applikationen auf mobilen Endgeräten bereitgestellt. Die Darstellung von Dashboards und Berichten ist auch die häufigste Nutzung von „mobile BI“.

Wegen der individuellen Anforderungen werden in Softwarewerkzeugen häufig lediglich Bausteine oder Entwicklungsumgebungen bereitgestellt, mit denen eigene Anwendungen durch grafische Entwicklung und Parametrisierung weit gehend programmierfrei definiert werden können. Viele Reporting- und Analysewerkzeuge können individuelle Sichten im Sinne von Cockpits bereitstellen. Weiterhin existieren Spezialwerkzeuge für den individuellen Aufbau von Business-Intelligence-Cockpits, z. B. von Arcplan, Board, Cubeware, IBM Cognos, Information Builders, Oracle, Qlik, SAP Design Studio oder SAS.

Im Rahmen des Strategiemanagements von Organisationen kommen häufig Scorecards zum Einsatz. Sie erfordern zusätzlich zu der Cockpit-Ansicht von Kennzahlen weitere Funktionen. Die Balanced Scorecard (BSC) als prominente Managementmethode der strategischen Unternehmensführung beispielsweise propagiert sowohl einen ganzheitlichen Blick auf die Key Performance Indicators (KPIs) eines Unternehmens als auch eine Umsetzung von Visionen und Strategien in konkrete Kennzahlen und Maßnahmen. Für die vielfältigen Aufgaben der Dokumentation, Maßnahmenplanung, Kommunikation und Überwachung von Kennzahlen spielen Softwarewerkzeuge eine wesentliche Rolle. Sie erhöhen die Produktivität bei Aufbau und Betrieb einer Balanced Scorecard und erlauben damit eine unternehmensweite Verbreitung eines Performance-Management-Konzeptes, insbesondere durch

- Integration oft verteilter Performance-Daten,
- Kommunikation der strategischen Kennzahlen im Unternehmen (z. B. als „Strategy Map“ – vgl. Abb. 6.8),
- Unterstützung der Zusammenarbeit und Förderung des direkten Informationsaustausches zwischen Mitarbeitern,
- Analyse der Performance-Daten,
- Ableitung konkreter Maßnahmen und Nachverfolgung von Projekten.

Dieses gibt allen Mitarbeitern Zugang zu der Strategie einer Organisation und unterstützt sowohl Kommunikation als auch Feedback. Auf diese Weise kann die Unternehmensstrategie kontinuierlich weiterentwickelt werden.

Gerade die Möglichkeiten der Kommunikation strategischer Ziele, die Abbildung von Ursache-Wirkungs-Ketten oder die Maßnahmenplanung und -verfolgung sind in Spezialwerkzeugen zum Aufbau von Balanced-Scorecard-Systemen enthalten. Beispiele für Anbieter sind Actuate PerformanceSoft, Corporate Planning, Corvu, Hyperspace, IBM Cognos, Infor, Oracle, Procos, Prodacapo, QPR, SAP, SAS, Tonbeller.

Die Funktionen der Datenintegration können über die übliche Infrastruktur von Business-Intelligence-Systemen abgebildet werden. Für die Datenanalyse und das Berichtswesen sind in Balanced-Scorecard-Werkzeugen häufig eigene Möglichkeiten enthalten oder es werden spezialisierte Reporting- und Analysewerkzeuge eingesetzt.

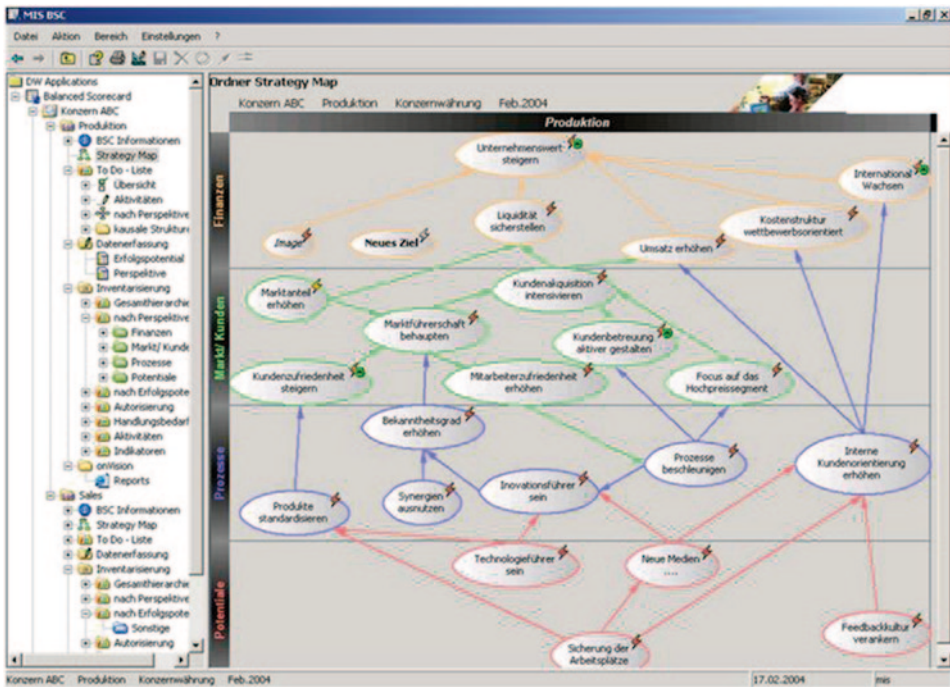


Abb. 6.8 Strategiekarte zur Abbildung von Ursache-Wirkungszusammenhängen einer Balanced Scorecard

6.5.2 Standard Reporting

Die üblicherweise am häufigsten genutzte Kommunikation von Kennzahlen erfolgt im Rahmen des betrieblichen Berichtswesens. Daten werden in der Regel tabellarisch mit ergänzenden Grafiken dargestellt – entweder in papier- oder bildschirmorientierten Ausgabeformaten. Dem Layout der Berichte kommt in der Regel hohe Bedeutung zu, so dass neben der tabellarischen auch die grafische Darstellung von Daten umfangreich unterstützt wird. Die wesentlichen zu unterstützenden Aufgaben sind:

- *Berichtsdefinition*, bei der die Formatierung aber auch die Flexibilität bei der Abfrage definiert wird. Abbildung 6.9 zeigt eine typische Definition eines Umsatzberichtes.
- *Informationsdistribution* zur Verteilung der Berichte an verschiedene Empfänger (häufig als E-Mail) in verschiedenen Formaten (Excel-Datei, PDF-Dokument). Alternativ auch die Steuerung des Zugriffs auf Berichte über Portale oder andere Mechanismen.
- *Informationsdarstellung*, durch Anzeige statischer Berichte auf verschiedenen Medien (z. B. Papier, Bildschirm, Mobile Geräte) und auch Ermöglichung gewisser Interaktionsmöglichkeiten, z. B. die Detaillierung eines Datenbereiches durch Anklicken.

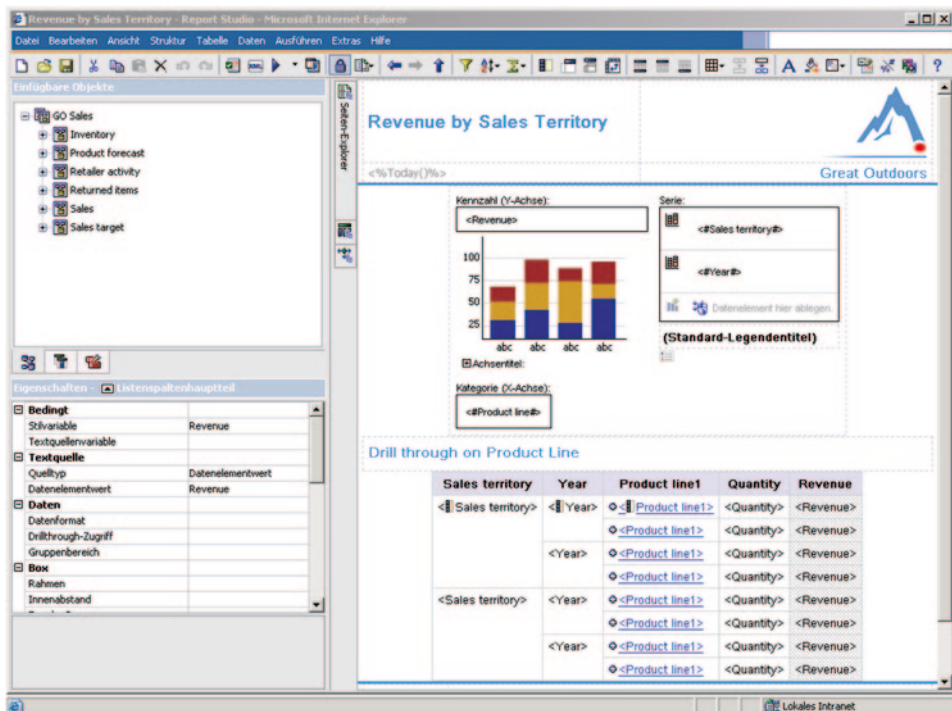


Abb. 6.9 Berichtsdefinition mit Platzhaltern für Grafiken, Tabellen und weiteren grafischen Elementen (rechter Bildschirmbereich) und Datenbankzugriff zur Verknüpfung des Berichts mit seinen Datenquellen (linker Bildschirmbereich)

Während die Definition von Berichten und ihrer Distribution die Erlernung der Anwendung von entsprechenden Werkzeugen für diese Aufgaben benötigt, sollte der Aufruf von Berichten keine Schulung erfordern. Moderne Berichtswerkzeuge mit individualisierten Web-Portalen und der Ausgabe der Berichte auf mobilen Endgeräten erfüllen diesen Anspruch.

Die Trennung von inhaltlicher und Layout-Definition eines Berichtes (Abb. 6.9) von seiner Ausführung durch Befüllung mit Daten und Anzeige (Abb. 6.10) ist das wesentliche Prinzip von Standard-Berichtswerkzeugen. Ohne eine neue Entwicklung können ständig aktuelle Berichte angeboten werden, da die Berichtsdefinition zur Laufzeit mit denjenigen Daten befüllt wird, die zu diesem Zeitpunkt in der Datenbank vorgefunden werden. Zusätzlich erfolgt die Datenabfrage zum Laufzeitpunkt des Berichtes auch rechtegesteuert. Insofern kann die gleiche Berichtsdefinition mit den Rechten verschiedener Anwender erfolgen, so dass in jedem erzeugten Bericht nur diejenigen Daten erscheinen, für die ein Anwender autorisiert ist.

Neben klassischen Papierberichten erlauben Portale im Intra- und Internet den einfachen webbasierten Zugriff auf Berichte und weitere Informationen. Eine wesentliche Entwicklung in den letzten Jahren ist die verstärkte Unterstützung von mobilen Endgeräten

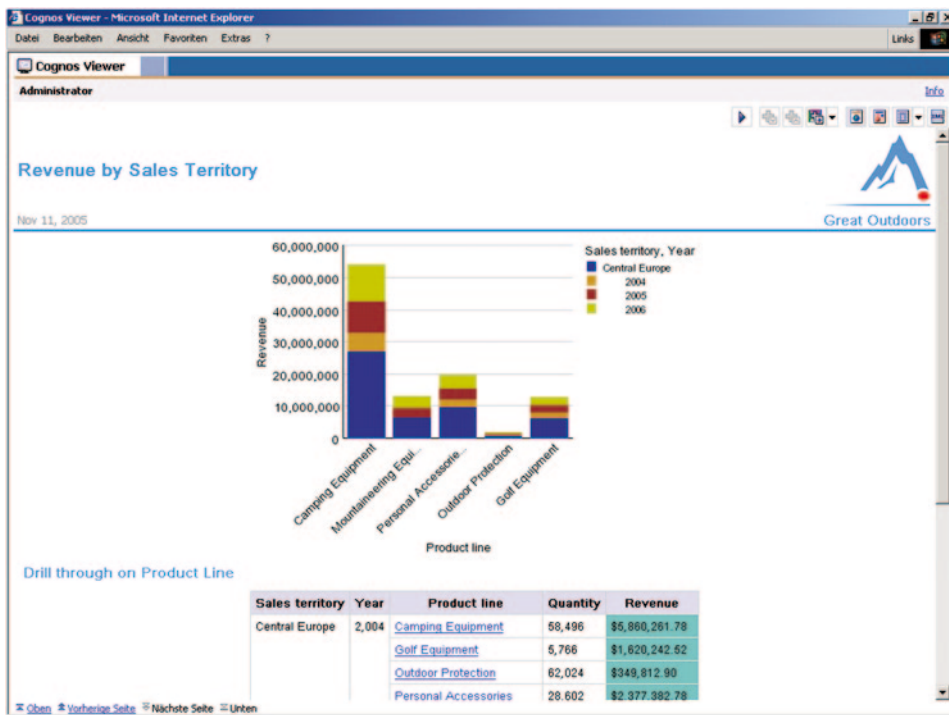


Abb. 6.10 Anzeige eines einfachen Umsatzberichtes in einer Web-basierten Berichtsumgebung (Abb. 6.10 zeigt die Ausgabe des in Abb. 6.9 definierten Berichts an.)

zur Berichtsdarstellung, um eine ortsunabhängige Informationslieferung zu gewährleisten („mobile BI“). Die Nutzung mobiler Endgeräte stellt Unternehmen allerdings vor einige, vorab auch häufig unterschätzte, Herausforderungen. Dies betrifft insbesondere die aufwändige Entwicklung und Wartung der Anwendungen wegen der Vielfalt der Gerätetypen, niedrige Abfrage-Performance wegen niedriger Bandbreiten, Sicherheit in Informationsübertragung, Speicherung und Zugriff auf mobilen Endgeräten sowie die Darstellung von Information mit eingeschränktem Platz.

Die vereinfachte und kostengünstige Erzeugung und Verteilung von Berichten ermöglicht erst die weitere Verbreitung von Information in Organisationen. Unterstützt wird diese Entwicklung von neuen, auf modernen Softwarearchitekturen basierenden IT-Komponenten, die im Hintergrund auch sehr hohe Datenvolumina verarbeiten können und Berichte an eine Vielzahl an Empfängern automatisiert und in verschiedenen Formaten versenden können. Neben der reinen technischen Möglichkeit zur Berichtserzeugung ist auch die Schaffung inhaltlicher Standards zum Berichtsdesign wichtig, um Empfängern ein schnelles Erfassen der Information zu erlauben. Unternehmensweite Standards für die Nutzung von Grafiktypen, Farben oder Beschriftungen sind hier sehr hilfreich und können selbst entwickelt werden oder Vorschlägen von Autoren wie Tufte, Few oder Hichert folgen (Abb. 6.11).

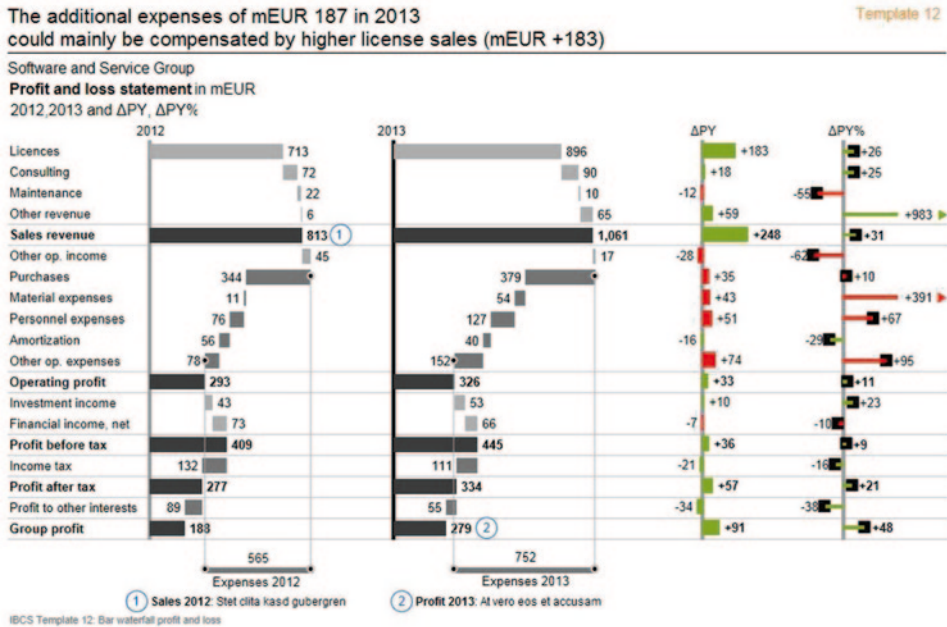


Abb. 6.11 Managementcockpit auf Basis von Hichert@IBCS Information Design Standards (umgesetzt in Arcplan)

Typische Standard-Berichtswerkzeuge sind Entwicklungswerkzeuge für IT-Personal. Die Anzeige der Berichte erfolgt dann in endanwenderfreundlichen Berichtsportalen oder kann in andere Anwendungen integriert werden. Die großen Anbieter liefern mit SAP Crystal Reports, IBM Cognos 8 BI, Oracle BI Publisher, Microsoft SQL Server Reporting Services oder SAS Enterprise BI Server entsprechende Werkzeuge. Spezialisten mit Marktbedeutung in diesem Bereich sind Actuate (Open Text), Arcplan, Information Builders, MicroStrategy. Gerade im Bereich der Entwicklungswerkzeuge für Standardberichtswesen finden sich relevante Open Source Lösungen mit BIRT, JasperReports und Pentaho Reporting.

6.5.3 Ad-hoc Reporting

Ergänzend zum Standardberichtswesen ist eine immer stärkere Anforderung die interaktive Navigation in Berichten und eigene Zusammenstellung von Berichten durch den Fachanwender selbst zu unterstützen – ohne langwierige Anforderungsdefinition und Umsetzung durch die IT. Dieser Trend fand einen Namen mit „Self Service BI“. Zusätzlich zur Rolle des Berichtsempfängers wird hier vor allem auch ein „Power User“ im Fachbereich unterstützt, der Berichte in gewissen Umfang erstellen und anderen Fachanwendern bereitstellen kann.

Ad-hoc Reporting bedarf zunächst einmal eines umfassenden Datenfundamentes, damit auch außerhalb des Standardberichtswesens benötigte Daten einbezogen werden können. Dieses Fundament wird in der Regel mit einem Data Warehouse oder abteilungsbezogenen Data Mart aufgebaut.

Während die Umsetzung des Standardreportings in der Regel über Mitarbeiter mit guten technischen Kenntnissen erfolgt, muss das Ad-hoc Reporting auch einem trainierten Fachanwender möglich sein. Werkzeuge müssen daher grundsätzlich eine höhere Unterstützung für Anwender bieten, die nicht mit Programmiersprachen und klassischen Entwicklungsumgebungen vertraut sind. Anbieter orientieren sich daher in ihren Lösungen eher an Microsoft Office Standards und bieten ergänzende Assistenten an. Ein wesentliches Hilfsmittel ist die Vereinfachung des Zugriffs auf Datenquellen durch eine semantische Schicht. Dieses stellt dem Fachanwender die verfügbaren Daten in ihren Zusammenhängen und in verständlichen Begriffen dar. Technisch gesprochen handelt es sich um eine Metadatenebene, die durch die Administratoren aufgebaut wird und eine Abstraktion zwischen Datenbanken und Fachanwender schafft. Mit Hilfe der semantischen Schicht kann ein Fachanwender Daten aus verschiedenen Datenquellen sehen und sie zu einem Bericht ad hoc kombinieren.

Alternativ zur logischen Integration der Daten in einem semantischen Modell wird auch zunehmende eine lokale, vom Fachanwender gesteuerte physische Integration von Daten in (häufig proprietären und lokalen) Datenspeichern von BI-Werkzeugen umgesetzt. Vorteil dieser Vorgehensweise ist die hohe Kontrolle für den Fachanwender und die Möglichkeit der Erreichung einer hohen Performance durch die Nutzung von Hauptspeicher-Datenbanken. Nachteil ist die Verteilung von Daten in zahlreichen dezentralen Datenbanken und die damit verbundene Problematik der Data Governance.

Wesentliche Werkzeuge in diesem Segment sind arcplan Enterprise, IBM Cognos 8 BI, Infor PM, Information Builders WebFocus, MicroStrategy, Oracle BI Foundation Suite, Qlik QlikView, SAP BusinessObjects WebIntelligence, SAS Enterprise BI Server. Viele Werkzeuge, die im Segment Analyse eingeordnet sind, erlauben auch ein Ad-hoc Reporting. Dies erfolgt häufig Excel-basiert auf Basis einer multidimensionalen Datenbank. Auch Standardberichtswerkzeuge erhöhen die Möglichkeiten für den Berichtsempfänger. Insofern sind die Grenzen zwischen Ad-hoc Reporting und Analyse sowie zwischen Ad-hoc Reporting und Standardreporting fließend (Abb. 6.12).

6.5.4 Analyse

Analyse geht über das Ad-hoc Reporting hinaus, indem es dem Anwender noch weitergehenden Zugriff auf Daten erlaubt sowie bestimmte Funktionen bereitstellt, mit denen Daten analysiert und so neue Informationen generiert werden können. Während Ad-Hoc Reporting stärker die Erstellung von Berichten fokussiert, wird bei der Analyse die explorative Untersuchung und Visualisierung von Daten sowie die Anwendung von mathematisch/statistischen Methoden auf Daten stärker unterstützt. Besondere Relevanz und

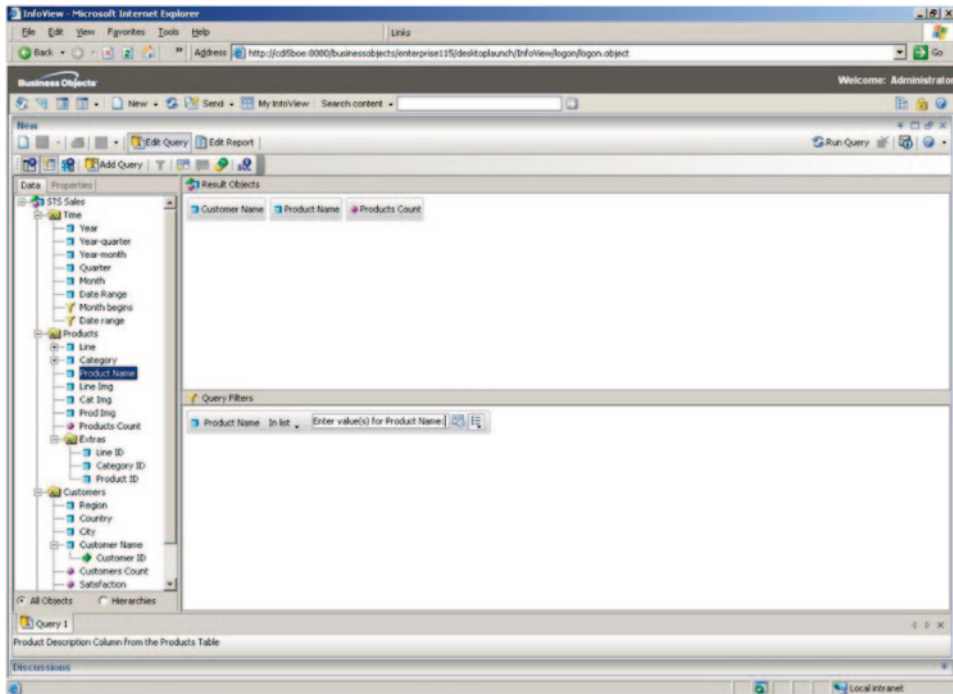


Abb. 6.12 Zur Berichtsdefinition mit einem Ad-hoc Berichtswerkzeug werden dem Berichtsdesigner die verfügbaren Datenobjekte geordnet und verständlich aus einer semantischen Schicht angeboten (linkes Fenster). Berichtsojekte, -layout und Datenfilter werden dann ohne Programmierung über grafische Oberflächen spezifiziert

Herausforderung ist dabei auch die Änderung von Datenstrukturen, um beispielsweise neue Produkt- oder Unternehmenssegmente zu simulieren. Dies wird häufig durch lokale Datenbestände, z. B. als abhängige Data Marts in multidimensionalen oder proprietären Datenbanken, realisiert (Abb. 6.13).

Neben der seit Jahren vorherrschenden dimensionalen Analyse finden sich in den letzten Jahren auch immer mehr Werkzeugen für eine mengenorientierte bzw. assoziative Analyse (z. B. QlikView, CommaSoft). Ein weiterer Markttrend ist die stärkere Unterstützung der Datenvisualisierung zur visuellen Analyse (z. B. Tableau, Tibco Spotfire) und fortgeschrittenen Analyse durch Einsatz mustererkennender Verfahren (s. Kapitel Data Mining). Alle Vorgehensweisen ergänzen sich und bieten überlappende Funktionalität zur Analyse, jeweils mit einem anderen Schwerpunkt und Vorteilen in unterschiedlichen Einsatzszenarien.

Die dimensionale Analyse erlaubt dem Endanwender eine flexible Navigation in einem modellierten, dimensional Datenraum. Neben Navigationsoperationen entlang von Hierarchien und der Möglichkeit zum freien Aufbau von Kreuztabellen mittels Pivotierung werden analytische Verfahren und statistische Auswertungsfunktionen zur Verfügung gestellt. Im Mittelpunkt der Analyse stehen Kennzahlen, die mehrstufig aggregiert werden

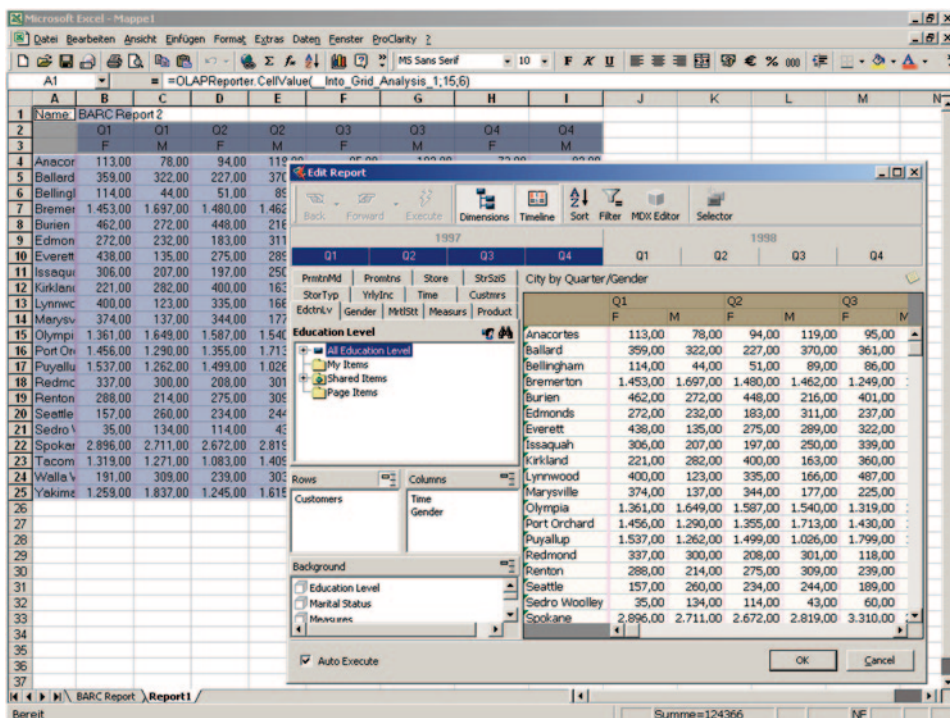


Abb. 6.13 Datenanalyse als Add-In in Microsoft Excel: Excel wird funktional ergänzt und ist die Anzeigumgebung für Daten aus einer dahinterliegenden Datenbank

können. Diese Eigenschaften werden von multidimensionalen Datenbanken (s. Abb. 6.8) sehr gut unterstützt.

Als Anwenderwerkzeug und „Fenster“ zur multidimensionalen Datenbank wird häufig Excel genutzt, das neben dem Datenzugriff um weitere Funktionen für die Analyse durch ein „Add-In“ ergänzt wird. IBM TM1, Infor PM, Jedox Palo, Microsoft SQL Server Analysis Services, MIK OLAP, Oracle Hyperion Essbase, PowerOLAP sind typische Vertreter. Alternativ bieten aber zahlreiche Anbieter auch eigenständige Lösungen an, die nicht Excel als primäres Anwenderwerkzeug nutzen. Beispiele hier sind Bissantz, Board, Cubeware, CUBUS EV oder Targit.

Für einige Anwender ist auch die Bereitstellung umfangreicherer Analyseverfahren sinnvoll. Dies können betriebswirtschaftliche Verfahren wie beispielsweise eine ABC-Analyse sein (Abb. 6.14)

Zusätzlich kann auch die Visualisierung von Daten nicht nur zur Präsentation sondern auch zur Analyse eingesetzt werden. Neben Spezialprogrammen zur visuellen Analyse (z. B. Tableau, Tibco Spotfire) bieten inzwischen auch viele Business-Intelligence-Suiten Möglichkeiten in dieser Richtung an (Abb. 6.15).

Ergänzend zu den seit vielen Jahren verfügbaren Möglichkeiten zur dimensional Analysis werden inzwischen immer mehr Werkzeuge für eine mengenorientierte bzw.

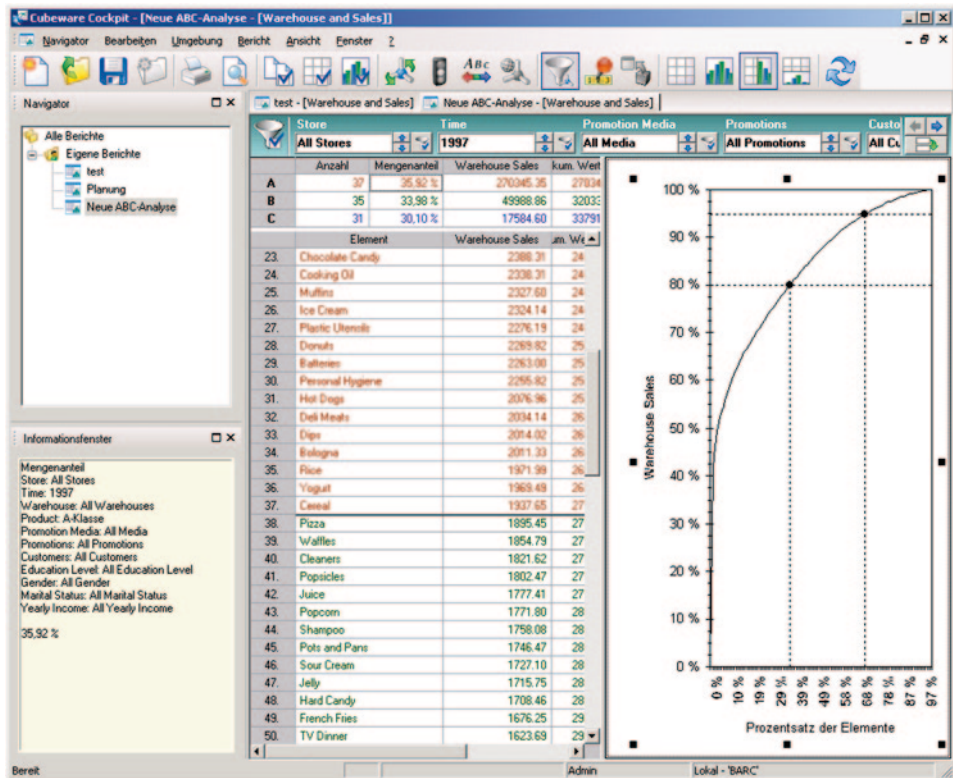


Abb. 6.14 Automatisierte ABC-Analyse

assoziative Analyse angeboten. In der mengenorientierten Analyse werden Informationsobjekte (z. B. Kunden) über ihre Attributausprägungen zu Gruppen zusammengefasst oder die Gruppenzugehörigkeit einzelner Informationsobjekte identifiziert. Im Datenmodell fehlen in der Regel vordefinierte dimensionale Strukturen. Stattdessen werden Verknüpfungen zwischen Datenbanken definiert und nicht oder nur eingeschränkt weiter modelliert. Es ist so möglich, zwischen abhängigen Attributen frei zu navigieren und die Eigenschaften der verbleibenden Menge an Informationsobjekten ad-hoc zu sehen. Die mengenorientierte Analyse zeigt vor allem bei einer großen Anzahl von Informationsobjekten mit vielen Attributen Vorteile, wenn das Filtern der Attribute Gegenstand der Analyse ist. Die Selektion von Zielgruppen aus großen Kundendatenbeständen im Rahmen eines Kampagnenmanagements ist ein typisches Beispiel hierfür.

Die Vorgehensweise der mengenorientierten Analyse ist Vor- und Nachteil zugleich: In sehr hoher Geschwindigkeit können Daten und ihre Zusammenhänge gezeigt und gefiltert werden. Kommt es allerdings auf die Abbildung von Strukturen und der Modellierung von Zusammenhängen zwischen den Elementen an, z. B. bei der Definition von Aggregationsregeln im Rahmen des Kostencontrollings, dann bieten dimensionale Analysewerkzeuge Vorteile. Typische Vertreter der mengenorientierten Analyse sind CommaSoft Infonea, HumanIT InfoZoom, Qlik QlikView, Tibco Spotfire (Abb. 6.16).

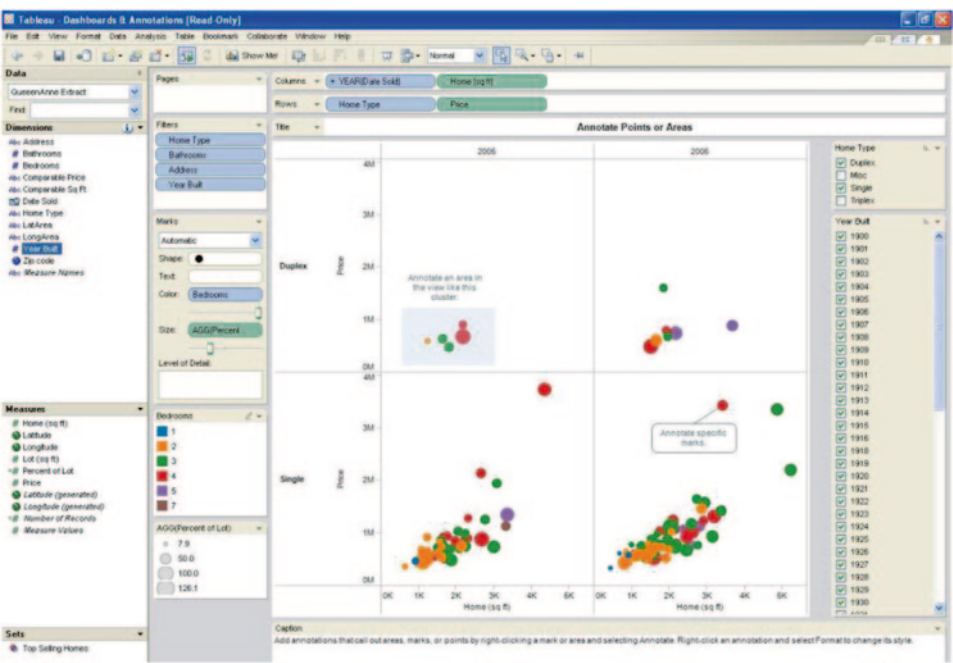


Abb. 6.15 Visuelle Analyse

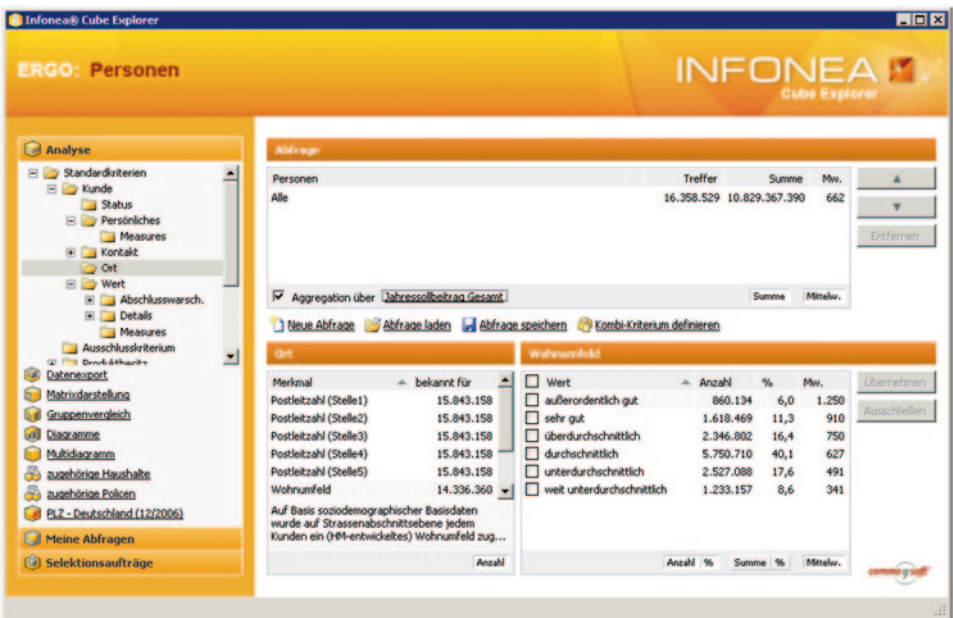


Abb. 6.16 Beispiel einer mengenorientierten Analyse, in der eine Treffermenge nach Wohnumfeld und Ort eingeschränkt wird

Dez per (Erf.)	Jan	Feb	Mrz	Apr
20.000,00	0,00	0,00	0,00	0,00
20.000,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
12.000,00	0,00	0,00	0,00	0,00
12.000,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00
0,00	0,00	0,00	0,00	0,00

Ausschneiden	Proportionale Verteilung...
Kopieren	Gleichwertige Verteilung...
Einfügen	Wiederholen...
Inhalte einfügen...	Löschen...
Zellen einfügen...	Prozent. Änderung...
Zellen löschen...	Gerade Linie...
Inhalte löschen	Wachstum %...
Zellen formatieren...	Relative proportionale Verteilung...
Dropdown-Auswahlliste...	Relative Prozentanpassung...
Übergangung hinzufügen	n-Elemente wiederholen...
Liste erstellen...	Gleichwertige Verteilung auf n-Elemente...
Hyperlink...	
Nachschlagen...	
Datenverteilung	
Blöcke	

Abb. 6.17 Angebot verschiedener Funktionen zur Verteilung aggregierter Planwerte auf Detailebenen in einem Planungswerkzeug

Unabhängig von der Analyseart sind Benutzerfreundlichkeit, gute Antwortzeiten und analytischer Funktionsumfang wesentliche Eigenschaften aller Analysewerkzeuge. Letztlich können neben dimensionaler und mengenorientierter Analyse auch weitere Varianten der Datenanalyse identifiziert werden. Von besonderer Bedeutung ist hier die automatisierte Datenmustererkennung in Datenbanken – das Data Mining (s. Kap. 5.7). Weitere, gerade simulative Analyseverfahren haben im Zusammenhang mit Planungsprozessen eine besondere Bedeutung (s. Kap. 5.5).

6.5.5 Planung und Simulation

Die Unterstützung von Planungs- und Budgetierungsprozessen erfordert eine Bereitstellung von Planungswerkzeugen für Dateneingabe, Datenverteilung, Forecasting und Simulation. Während im klassischen Berichtswesen Daten nur von Datenbanken zum Anwender in eine Richtung fließen, müssen Planungswerkzeuge auch ein Zurückschreiben von Daten aus dem Endanwenderwerkzeug ermöglichen. In kleinen und mittleren Unternehmen kann eine kostengünstige und schnelle Planungsunterstützung durch Standardwerkzeuge zur Planung erreicht werden. Diese bilden häufig einen großen Teil der betriebswirtschaftlichen Planungslogik durch eine vordefinierte Verknüpfung der Erfolgs-, Finanz- und Bilanzplanung bereits ab (Abb. 6.17).

Für größere Unternehmen mit vielen Beteiligten an den Planungs- und Budgetierungsprozessen erreichen Planungsplattformen mit ihren integrierten Funktionen zur Prozessunterstützung sowie die Flexibilität zur Anpassung an unternehmensspezifische Anforderungen häufig eine deutliche Effizienzsteigerung.

Neben den großen Anbietern IBM, Oracle, SAP und SAS bieten zahlreiche, häufig auch lokale Anbieter, Planungslösungen an, z. B. Arcplan, Bissantz, Board, CoPlanner, Corporate Planning, CUBUS, Denzhorn, elKom, Evidanza, Infor, LucaNet, Macs Software, Prevero, Prophix, Software4You, Tagetik, Thinking Networks u. v. a. m.

Grundsätzlich ist die Verbreitung von multidimensionalen Datenbanken als Plattform für Planungsanwendungen sehr hoch. Neben der Simulation ist auch eine hohe Performance bei der Datenverteilung ein besonderer Vorteil dieser Technologie.

6.5.6 Legale Konsolidierung

Die legale Konsolidierung gehört auch durch das engere Zusammenwachsen von externem und internem Rechnungswesen sowie Konsolidierung und Finanzplanung ebenfalls in eine umfassende Business-Intelligence-Systemlandschaft. Besondere Herausforderungen liegen hier in der Datenintegration aus heterogenen datenliefernden Einheiten, Prozessunterstützung bei der Abstimmung sowie bestimmten Buchungsmöglichkeiten. Im Back-End ist die Fähigkeit entscheidend, verschiedene Positionspläne und Rechnungslegungsstandards abzubilden und diese untereinander sowie mit der internen Sicht des Controllings verbinden zu können. Reporting- und Analyseanforderungen in der Konsolidierung werden durch Standard-Berichts- und Analysewerkzeuge oder eigenen Komponenten der Konsolidierungswerkzeuge umgesetzt. Gemeinsame Datenbasen für eine datenseitige Integration von Konsolidierung, Berichtswesen aber auch Planung sind hier ein nützliches Fundament.

Verbreitete Konsolidierungswerkzeuge sind Corporate Planning CP-CONS, Denzhorn BPS-ONE, elKomKons, IBM Cognos Controller, Infor PM Financial Consolidation, IDL Konsis, Lucanet Financial Intelligence Suite, Oracle Hyperion Financial Management, SAP BPC, SAP Financial Consolidation, Tagetik.

6.5.7 Data Mining

Die komplexe und ungerichtete Analyse von Datenbeständen zur Entdeckung von Strukturen und Mustern durch Verfahren der Statistik, des maschinellen Lernens und der künstlichen Intelligenz wird in Unternehmen regelmäßig nur in speziellen Gebieten eingesetzt. Im Zuge der Analyse von immer mehr Daten aus unterschiedlichsten („Big Data“) Quellen, der wachsenden Reife der Business Intelligence in Unternehmen und der Frage nach dem Mehrwert der BI-Lösungen für das Unternehmen erlebt Data Mining ein deutlich

wachsendes Interesse, neuerdings vor allem unter dem Begriff „Predictive Analytics“. Die Vorhersage von Werten (Prädiktion) ist neben der Beschreibung von Strukturen wie Segmenten oder Assoziationen die wichtigste Aufgabe von Data Mining. Mit wenigen Ausnahmen (z. B. Bissantz DeltaMaster oder SAP InfiniteInsight) sind die Werkzeuge für Experten (z. T. auch „Data Scientists“ genannt) ausgerichtet, die interessante Aussagen zu verschiedenen Aufgabentypen treffen können. Im Marketing- und Vertriebscontrolling zum Beispiel finden sich zahlreiche Anwendungsbereiche, insbesondere für Segmentierungsaufgaben (z. B. Kundensegmentierung), Klassifikationsaufgaben (z. B. Bonitätsbeurteilung/Scoring, Kündigungsvorhersage) oder Assoziationsaufgaben (z. B. Warenkorbanalyse im Handel).

Beispiele für Anbieter kommerzieller Data Mining Werkzeuge sind neben den Marktführern SAS und IBM SPSS, Angoss, Bissantz, Information Builders, Fico, Prudsys, StatSoft, SAP oder Tibco. Data-Mining-Funktionen sind inzwischen auch in die marktführenden Datenbanken integriert worden und stehen so bei IBM, Microsoft, Oracle und SAP zur Verfügung. Letztlich finden sich auch immer mehr Open Source Bibliotheken und Werkzeuge im Markt angeboten, insbesondere „R“ hat hier eine sehr weite Verbreitung, aber auch RapidMiner oder KNIME.

6.6 Fazit

Informationstechnologie kann zahlreiche ineffiziente Aufgaben der manuellen Datensammlung, -aufbereitung und -verteilung beschleunigen und ermöglicht so, Anwender bei der Analyse und Steuerung ihres Geschäftes zu unterstützen. Zum Einsatz kommen hier Softwarelösungen für das Datenmanagement, insbesondere für Datenintegration, Datenqualitätsmanagements und Datenspeicherung. Der Markt für diese Werkzeuge wächst sehr stark, was ihre Bedeutung gerade vor den Herausforderungen von „Big Data“, also wachsende Datenmengen, schnellere Integration und Einbeziehung polystrukturierter Daten unterstreicht. Der Grad an Transparenz des Unternehmensgeschehens wird durch den sinnvollen Einsatz von Business-Intelligence- und Performance-Management-Systemen deutlich erhöht. Die wesentlichen Anwendungsklassen hierbei sind Cockpits und Scorecards, Standard- und ad-hoc Berichtswesen, Datenanalyse und Data Mining sowie Planung und legale Konsolidierung. Ergebnis des Einsatzes dieser Werkzeuge auf einer passenden Datenmanagement-Infrastruktur ist eine schnellere Reaktionszeit auf Veränderungen, die entscheidende Wettbewerbsvorteile sichern kann. Insbesondere für die Datenanalyse ist allerdings weiterhin die Intelligenz des Anwenders unabdingbar, um aus ständig wachsenden Datenbergen die interessante Information herauszufiltern und Schlussfolgerungen abzuleiten und ggf. Entscheidungsmodelle zu kreieren.

Das Softwaresegment für analytische Informationssysteme ist einer der wenigen Bereiche des Softwaremarktes, der in den letzten Jahren kontinuierliche Zuwachsraten verzeichnen können. In den Bereichen der Standard-Funktionen wie Reporting und einfache

Datenanalyse hat eine Marktkonsolidierung begonnen. Dennoch gibt es eine große und wachsende Zahl von mehr als 250 Anbietern, die alleine in Deutschland ihre Systeme vertreiben und häufig auch innovative Ideen für mehr Business Intelligence einbringen. Für den Anwender ist daher eine Kenntnis der verschiedenen Marktsegmente und Anwendungsbereiche sowie eine sorgfältige Auswahl der passenden Lösung weiter entscheidend.

Literatur¹

- Bauer, A., Günzel, H. (Hrsg.): Data-Warehouse-Systeme, 4. Aufl. dpunkt.verlag, Heidelberg (2013)
- Bange, C., et al.: Der Markt für Business Intelligence in Deutschland 2013. BARC, Würzburg (2014)
- BARC.: Guide Business Intelligence 2014. BARC, Würzburg (2014a). www.barc.de/guide
- BARC.: Survey „Big Data Analytics“. BARC, Würzburg (2014b)
- Gluchowski, P.: Business intelligence. Konzepte, Technologien und Einsatzbereiche. HMD. **38**(222), 5–15 (Dezember 2001)
- Inmon, W.H.: Building the Data Warehouse. Wiley, New York (1996)
- Wang, R.Y., Strong, D.: Beyond accuracy: What data quality means to data consumers. JMIS. **12**(4), 5–33 (1996)

¹ Eine aktuelle Darstellung verfügbarer BARC-Studien mit Produktvergleichen ist unter www.barc.de verfügbar.

Teil II

Architektur und Technologien

Konzeptionelle Überlegungen zur Filterung,
Harmonisierung, Aggregation und Anreicherung im
Data Warehouse (DWH)

Hans-Georg Kemper und Ralf Finger

Inhaltsverzeichnis

7.1 Operative und dispositive Daten. 130

7.2 Transformation – Ein Schichtenmodell 131

 7.2.1 Filterung – Die Extraktion und Bereinigung operativer Daten. 133

 7.2.2 Harmonisierung – Die betriebswirtschaftliche Abstimmung gefilterter Daten . . . 137

 7.2.3 Aggregation – Die Verdichtung gefilterter und harmonisierter Daten 139

 7.2.4 Anreicherung – Die Bildung und Speicherung betriebswirtschaftlicher
 Kenngrößen. 141

7.3 Aktuelle Konzepte – Datentransformation wird nicht obsolet 142

7.4 Fazit. 143

Literatur 144

Zusammenfassung

Der folgende Beitrag beschäftigt sich mit der Problematik der Überführung operativer in dispositive Datenbestände, die direkt für Analytische Informationssysteme nutzbar sind. Die hierfür erforderlichen Transformationsprozesse werden in aufeinander aufbauende Filterungs-, Harmonisierungs-, Aggregations- sowie Anreicherungsaktivitäten unterschieden und durch entsprechende Praxisbeispiele illustriert.

H.-G. Kemper (✉)
Lehrstuhl für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik I,
Universität Stuttgart, Stuttgart, Deutschland
E-Mail: kemper@wi.uni-stuttgart.de

R. Finger
Information Works Unternehmensberatung und Informationssysteme GmbH, Cologne,
Deutschland
E-Mail: iw-info@information-works.de

7.1 Operative und dispositive Daten

Daten der operativen Systeme unterstützen gezielt die wertschöpfenden Geschäftsprozesse innerhalb eines Unternehmens. Sie sind demnach ausschließlich auf die Steuerung und Überwachung des Tagesgeschäftes ausgerichtet und meist transaktionsbezogen. Somit sind die Daten – je nach operativem Einsatzzweck – in ihren Begrifflichkeiten häufig nicht eindeutig vergleichbar und in ihrer Bewertung sowie Konsolidierung begründbar unterschiedlich. Weiterhin spiegeln sie aufgrund ihrer Transaktionsorientierung in aller Regel ausschließlich einen momentanen Informationsstand wider und werden durch Folgetransaktionen dynamisch im Zeitverlauf aktualisiert (Bauer und Günzel 2013).

Aus diesen Gründen bieten selbst Softwareanbieter integrierter ERP-Systeme – wie z. B. SAPTM – dedizierte dispositive Datenhaltungssysteme an, die als Datenreservoirs für die analytischen Anwendungssysteme eines Unternehmens herangezogen werden (Chamoni et al. 2005; Schütte et al. 2001).

Eine typische Architektur dispositiver Datenhaltungssysteme zeigt Abb. 7.1. Sie gliedert sich in *Core Data Warehouse* (C-DWH) und *Data Marts*. Dieses – auch als Hub-and-Spoke-Architektur bezeichnete – dispositive Datenkonzept versteht sich als themenbezogenes, integriertes Datenreservoir, bei dem das aus Sicht der Entscheidungsunterstützung benötigte, in aller Regel aggregierte Datenmaterial dauerhaft – also historienbildend – abgelegt wird (Jarke et al. 2003).

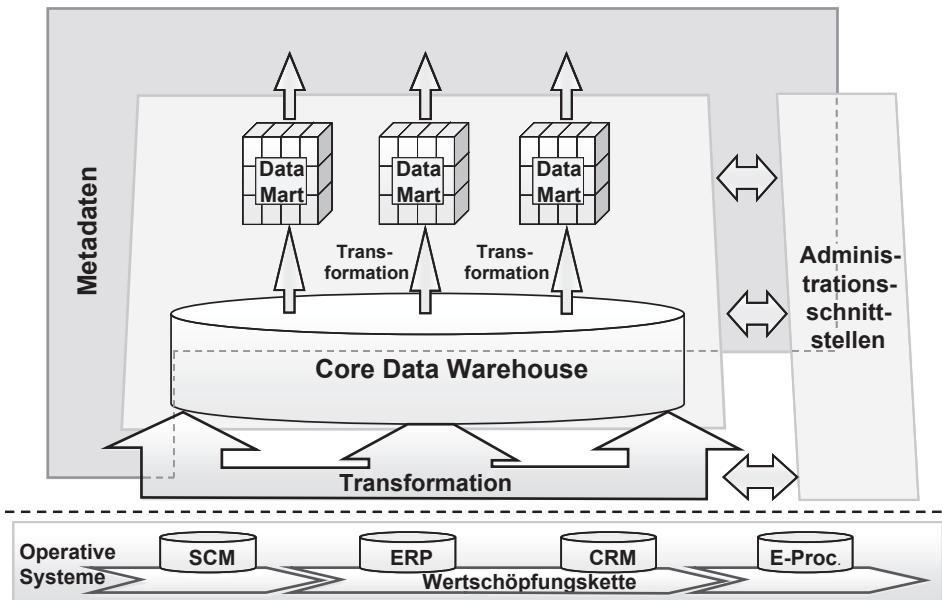


Abb. 7.1 Data Warehouse-Architektur

7.2 Transformation – Ein Schichtenmodell

Die mangelnde Eignung operativer Datenquellen stellt bei dem Aufbau und Betrieb von Data Warehouses eines der größten Probleme dar (von Maur und Winter 2002; Turban et al. 2010). Aus diesem – von vielen Analytikern als Manko empfundenen – Tatbestand jedoch auf die generelle Qualität des operativen Datenbestandes Rückschlüsse zu ziehen, verbietet sich. Vielmehr sind die operativen Daten auf die Anforderungen der mengen- und wertorientierten Abrechnungs-, Administrations- und Dispositionssysteme ausgerichtet und gewährleisten einen reibungslosen Ablauf des Tagesgeschäftes. Anforderungen aus Sicht der Entscheidungsunterstützung sind nicht die Kernaufgaben dieser Systeme, wurden entsprechend bei der Konzeption der operativen Datenhaltung meist nicht erhoben und können somit auch nicht abgedeckt werden. Demnach liegt es auf der Hand, dass lediglich mit Hilfe gezielter Umwandlungsaktivitäten aus den operativen Daten entscheidungsrelevante Informationen gewonnen werden können (Anandarajan et al. 2004; Lehner 2003).

Die für die betriebswirtschaftliche Interpretation der operativen Daten erforderlichen Prozesse werden im Weiteren unter dem Oberbegriff *Transformation* zusammengefasst und setzen sich aus den Sub-Prozessen der *Filterung*, *Harmonisierung*, *Aggregation* und *Anreicherung* zusammen.

- **Filterung** Unter der *Filterung* wird die Extraktion aus den operativen Daten und die Bereinigung syntaktischer oder inhaltlicher Defekte in den zu übernehmenden Daten verstanden.
- **Harmonisierung** Die *Harmonisierung* bezeichnet den Prozess der betriebswirtschaftlichen Abstimmung gefilterter Daten.
- **Aggregation** Die *Aggregation* ist die Verdichtung gefilterter und harmonisierter Daten.
- **Anreicherung** Die Bildung und Speicherung betriebswirtschaftlicher Kenngrößen aus gefilterten und harmonisierten Daten wird als *Anreicherung* bezeichnet.

Abbildung 7.2 zeigt die einzelnen Sub-Prozesse der Transformation. Deutlich wird, dass zu Beginn der Transformation die physische Extraktion der gewünschten Daten aus den operativen Beständen erfolgt. In aller Regel ist dieses Datenmaterial nicht vollkommen frei von Defekten. Daher können Mängel des operativen Datenmaterials im Rahmen der Extraktion deutlich werden, wie das Fehlen von Datenwerten oder die nicht korrekte Belegung von Datentypen aufgrund syntaktischer oder inhaltlicher Falscheingaben der Benutzer der operativen Systeme. Erst nach Abschluss von Reinigungsaktivitäten, die zum Teil automatisiert ablaufen können, häufig jedoch auch die Einbindung sachkompetenter Mitarbeiter erfordern, ist die Filterung des Datenmaterials abgeschlossen.

Um aussagefähige Informationen zu erhalten, sind die extrahierten und bereinigten operativen Daten anschließend einem Harmonisierungsprozess zu unterziehen. Hierbei werden Berichtswährungen abgeglichen, unterschiedliche Gebiets- bzw. Ressortgrenzen

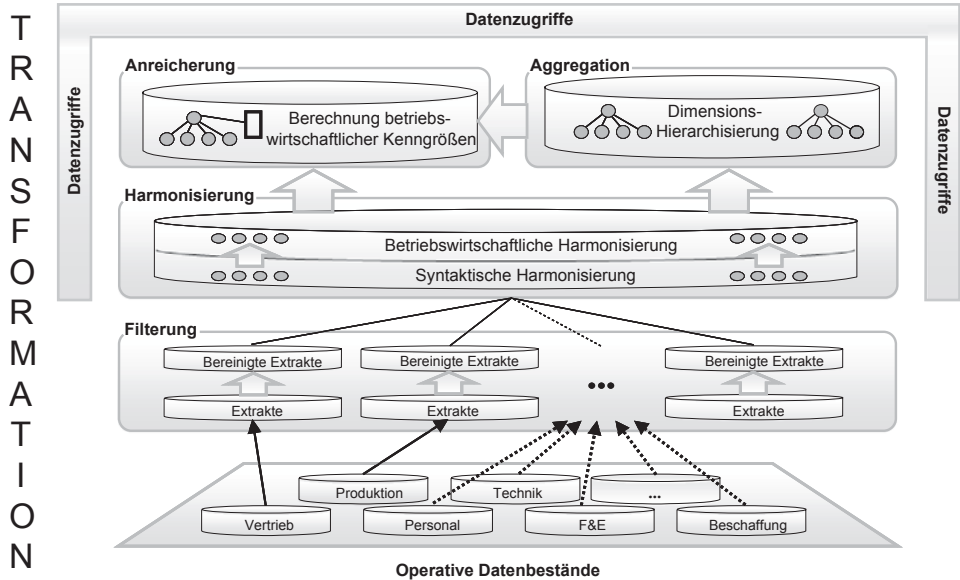


Abb. 7.2 Sub-Prozesse der Transformation. (Kemper et al. 2010)

zusammengeführt oder betriebswirtschaftlich nicht einheitlich abgegrenzte Begrifflichkeiten abgestimmt. Des Weiteren werden die Daten auf den von den späteren Nutzern des Data Warehouse gewünschten Detaillierungsgrad – auch Granularität genannt – transformiert (Inmon 2005). So werden beispielsweise nicht selten die Positionen der Einzelbelege zu tages-, wochen- oder monatsaktuellen Werten zusammengefasst.

Mit Hilfe der Aggregation werden die gefilterten und harmonisierten Daten, die sich in der gewünschten Granularität befinden, bewusst um Verdichtungsstrukturen erweitert, die sich an den Auswertungsdimensionen und antizipierbaren -hierarchien orientieren. Nicht selten werden hierbei Produkte zu Produktgruppen, Kunden zu Kundengruppen, Filialen zu Regionen und/oder Gebieten zusammengefasst.

Die Anreicherung erlaubt es, zusätzlich zu den Aggregatstrukturen einen betriebswirtschaftlichen Mehrwert in die Datenbasis zu integrieren. Berechnungen von Abweichungen oder die Ermittlung von Kennziffern – wie Deckungsbeitrag oder Umsatz – sollen hier nur exemplarisch für die große Anzahl von Möglichkeiten genannt werden.

Die folgenden Ausführungen beschreiben den Transformationsprozess anhand der oben skizzierten Sub-Prozesse auf einer konzeptionellen Ebene. Für den praktischen Einsatz innerhalb eines Data Warehouse wäre es erforderlich, die beschriebenen Designprinzipien systemtechnisch umzusetzen.

7.2.1 Filterung – Die Extraktion und Bereinigung operativer Daten

Die Filterung umfasst – wie oben dargestellt – Extraktions- und Bereinigungsaktivitäten.

Extraktion Erste Voraussetzung zur Übernahme operativer Datenquellen in eine dispositive Datenhaltung ist die Eröffnung der grundsätzlichen Zugriffsmöglichkeit auf operative Datenbestände. Diese sind in der Unternehmenspraxis häufig hinsichtlich Datenbanktechnologien und Betriebssystemen auf heterogenen Plattformen vorzufinden. Hier ist in enger Zusammenarbeit mit den Administratoren der operativen Systeme zunächst zu prüfen, ob direkte Durchgriffsmöglichkeiten oder Datenexportverfahren aus den operativen Systemen bereits verfügbar bzw. ob zusätzliche Schnittstellen zu erarbeiten sind. Weiterhin ist für eine Implementierung der Datenübernahme zu berücksichtigen, in welchen zyklischen Abständen ein „Snapshot“ der operativen Daten zu erfolgen hat (Kimball et al. 1998). Hierbei muss insbesondere beachtet werden, dass für das Auslesen operativer Datenbestände häufig lediglich begrenzte Zeitfenster zur Verfügung stehen. So konkurrieren die Extraktionsroutinen – wenn die Daten nicht Real- oder Near-time ausgelesen werden – meist mit anderen batchorientierten Anwendungen und werden in aller Regel in der Nacht abgewickelt, damit der operative Tagesbetrieb nicht durch ressourcenverbrauchende Lese- und Kopiervorgänge behindert wird.

Unterschiedliche Konzepte existieren im Hinblick auf die Abschätzung der Datenqualität. Ein erster Ansatz ist der Versuch, die Qualität operativer Datenfelder im Vorhinein abzuschätzen. Die Beantwortung der folgenden Fragen kann hierbei wertvolle Hilfen bieten:

- Handelt es sich bei dem zu extrahierenden Datenfeld um ein sog. „Muss-Eingabefeld“? In diesen Fällen kann davon ausgegangen werden, dass keine unbeabsichtigten inhaltslosen Leerfelder in der operativen Datenquelle existieren, da mit Hilfe programmtechnischer Funktionen – z. B. in Form spezieller Eingabemasken – in den operativen Systemen die Existenz eines Datenwertes sichergestellt wird.
- Sind Plausibilitätsprüfungen bei der manuellen oder automatisierten Datenerfassung in den operativen Systemen implementiert? Kann diese Frage bejaht werden, so ist sichergestellt, dass die zu extrahierenden Datenwerte – je nach Qualität der durchgeführten Plausibilitätsprüfungen in den operativen Systemen – keine unzulässige Datensyntax enthalten bzw. keine unerlaubten Datenausprägungen aufweisen.
- **Wird das Datenfeld im operativen Datenbestand tatsächlich seiner ursprünglichen Zwecksetzung folgend gepflegt?** In der Praxis ist es nicht unüblich, die ursprünglichen Feldvorgaben in operativen Datenbeständen im Bedarfsfalle zu modifizieren. Diese Vorgehensweise hat aus Sicht der operativen Datenverarbeitung den – durchaus diskussionswürdigen – Vorteil, dass bei Veränderungen der operativen Geschäftsabwicklung im Zeitverlauf keine oder lediglich geringfügige Modifikationen an den operativen Systemen vorzunehmen sind. Allerdings bewirkt eine solche Vorgehensweise, dass diese Daten mit Hilfe von analytischen Informationssystemen nur schwer auswertbar gemacht werden können. Um in einer solchen Situation abgestimmte Historien

aufbauen zu können, sind daher Maßnahmen während der Bereinigung zu ergreifen. Ein geeignetes Verfahren sind Umwandlungsroutinen, die über die Zeitpunkte der Modifikation gesteuert eine semantische Gleichheit der Dateninhalte über die Zeitachse gewährleisten.

- **Seit wann werden die einzelnen Felder der operativen Datenquelle erfasst?** Hierbei ist insbesondere zu klären, welche Datenfelder nachträglich in die operative Datenquelle integriert wurden. Auch diese Frage determiniert die Qualität des operativen Datenmaterials hinsichtlich der Erstübernahme in ein Data Warehouse, da sie bestimmt, in welchem Umfang Vergangenheitsdaten in Historien aufgebaut werden können.
- **Existieren bereits konkrete Pläne, die operativen Datenquellen zu verändern?** Im Gegensatz zu den Vorfragen tangiert die Beantwortung dieser Frage die zukünftige Entwicklung des Data Warehouse. Da die Datenübernahmen in aller Regel periodisch erfolgen und die neu extrahierten Daten den Altbestand im Data Warehouse ergänzen, sollte sichergestellt sein, dass die operativen Datenquellen in absehbarer Zukunft in ihrer Struktur als stabil betrachtet werden können. Ist dieses nicht der Fall, sind zukünftige Veränderungen in den zyklisch ablaufenden Extraktionsprogrammen sowie etwaige Migrationen der dispositiven Datenstrukturen zu planen.

Wie leicht nachvollziehbar ist, gibt die Beantwortung dieser und ähnlicher Fragen durch erfahrene Spezialisten der betreffenden operativen Systeme wichtige Hinweise auf die generelle Qualität der Datenquellen. So können diese Kriterien auch Anhaltspunkte dafür liefern, auf welche Datenquelle sinnvollerweise zurückgegriffen werden sollte, wenn ähnliche Inhalte aus unterschiedlichen Transaktionssystemen gewonnen werden können.

In Abb. 7.3 wird der Sachverhalt an einem umfassenden Beispiel illustriert.

Allerdings ist die Datenqualität – auch von erfahrenen Experten des betreffenden Transaktionssystems – nicht immer im Vorfeld gänzlich antizipierbar. Aus diesem Grunde empfiehlt sich nach der obigen Prüfung der Datenqualität ein weiterer Prüfungsschritt, der auf die tatsächlich in einem Datenfeld vorkommenden Ausprägungen abzielt. Moderne Extraktions- und Datenmanagementwerkzeuge erlauben es, auf einfache Weise Übersichten zu erzeugen, die je Feld Aufschluss geben über absolute und prozentuale Häufigkeiten des Vorkommens aller im Datenbestand vorhandenen Ausprägungen. Nicht selten findet man in solchen Übersichten systematische wie unsystematische Datenfehler oder auf den ersten Blick nicht interpretierbare Ausprägungen.

Bereinigung Im Rahmen der *Bereinigung* werden die extrahierten Daten sowohl von *syntaktischen* als auch von *semantischen* Mängeln befreit. Wie die Abb. 7.4 veranschaulicht, können hierbei mehrere Mängelklassen unterschieden werden.

1. Klasse: Automatisierbare Defekterkennung mit automatisierbarer Korrektur während des Extraktionsvorganges.
2. Klasse: Automatisierbare Defekterkennung mit manueller Korrektur nach dem Extraktionsvorgang.

Eine zu filternde Datentabelle einer operativen Quelle in einem Versicherungsunternehmen habe folgenden Satzaufbau:

PRÄMIE (KUNDEN_NR, KUNDEN_ART, PRÄMIE, BETREUER,...)

Eine Recherche ergibt:

- 1. **Muss-Felder?** KUNDEN_NR und PRÄMIE sind Muss-Felder. Alle anderen Attribute können somit Leerfelder aufweisen.*
- 2. **Plausibilitätskontrollen?** KUNDEN_NR unterliegt als Primärschlüssel einer Syntax- und Wertebereichsprüfung. Alle anderen Attribute können somit sowohl Syntaxfehler als auch inhaltliche Unstimmigkeiten aufweisen.*
- 3. **Konsistente Feldpflege?** Das Feld KUNDEN_ART hat im Zeitverlauf eine Interpretationsmodifikation erfahren. Ursprünglich wurden lediglich die Ausprägungen „0“ für „PRIVAT“ und „1“ für „GEWERBLICH“ akzeptiert. Zur Identifikation der Finanzkraft privater Kunden und der Branchenzugehörigkeit gewerblicher Kunden wurde nachträglich die Domäne des Attributes um die Ausprägungen „2“, „3“, „4“ erweitert, so dass nun die Ausprägungen „0“ und „1“ die privaten sowie „2“- „4“ die gewerblichen Kunden repräsentieren. Für die Übernahme des Attributes in das Data Warehouse sind demnach Transformationsregeln erforderlich, die - mit einem Zeitstempel versehen - die erweiterte Semantik zumindest für historische Auswertungen auf die ursprüngliche Differenzierung „PRIVAT“ und „GEWERBLICH“ zurückführen.*
- 4. **Zeitraum der Felderfassung?** Das Attribut BETREUER wurde nachträglich in den operativen Datenbestand integriert. Hiermit ergeben sich Restriktionen bzgl. der Auswertung der Historie.*
- 5. **Veränderungspläne?** Es ist geplant, zu Beginn des nächsten Jahres ein Re-design der operativen Systeme zur Prämienberechnung vorzunehmen, so dass eine Strukturmodifikation der operativen Datenhaltung erfolgen wird. Da die Stabilität der operativen Datenquelle demnach nicht gegeben ist, sollte geprüft werden, ob die neue Datenstruktur die vorgesehenen Extraktionen erlaubt bzw. welcher Aufwand mit der Anpassung der Extraktionsroutinen an die neue Struktur verbunden sein wird.*

Abb. 7.3 Abschätzung der Qualität operativer Daten

3. Klasse: Manuelle Defekterkennung mit manueller Korrektur nach dem Extraktionsvorgang.

Die erste Mängelgruppe beinhaltet die syntaktischen und semantischen Mängel, die automatisiert durch implementierte Routinen behoben werden können. Diese Gruppe stellt den Teil der Mängel dar, der vor der Implementierung der Extraktionsroutinen bekannt ist bzw. dessen Auftreten antizipierbar ist. So könnten interne Format-, Steuer- oder Sonderzeichen, die beispielsweise zur Dokumentation von Stornobuchungen in den operativen Daten verwendet werden, mit Hilfe von implementierten Mapping-Tabellen in den extrahierten Daten automatisiert erkannt und bereinigt werden. Als Beispiel einer automatisierten Behebung eines semantischen Mangels kann der Fall angesehen werden, dass fehlende

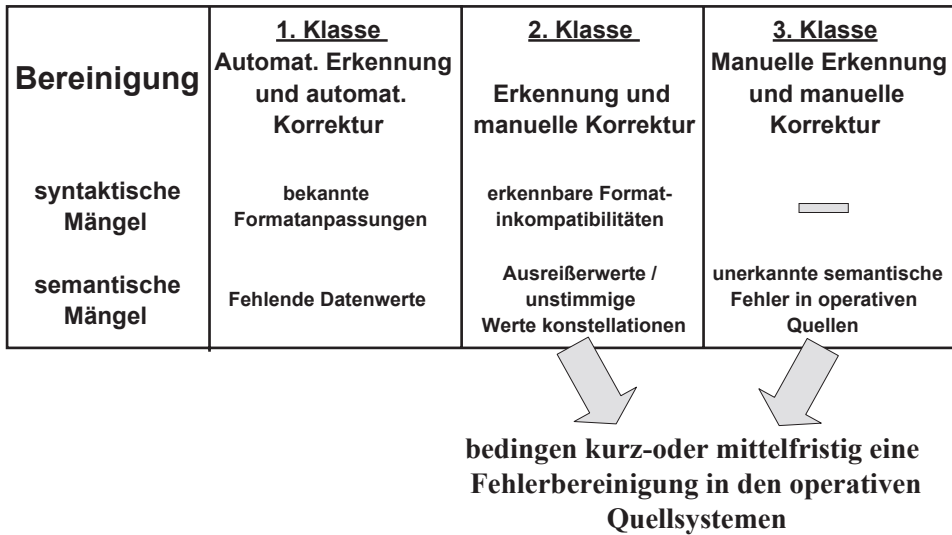


Abb. 7.4 Mängelklassifikation im Rahmen der Bereinigung. (Kemper 1999)

Ist-Werte in den operativen Daten – etwa verursacht durch unterbliebene Übertragung der Umsatzdaten einer Filiale – erkannt und mit Hilfe von vorab definierten Regeln durch Äquivalenzwerte – wie Planwerte des Monats, Ist-Werte des Vormonats, Ist-Werte des Vorjahresmonats – in den Extrakten ersetzt und entsprechend kommentiert werden, damit sinnvolle Verdichtungen auf der Basis des Datenmaterials durchgeführt werden können.

Der größte Teil der Bereinigungsaktivitäten kann jedoch nicht vollständig automatisiert durchgeführt werden. In der ersten Klasse dieser nicht vollständig automatisierbaren Bereinigungen sind lediglich die Mängel durch implementierbare Routinen erkennbar, die notwendigen Bereinigungen müssen hingegen manuell durch technische und betriebswirtschaftliche Fachkräfte durchgeführt werden. Bei den syntaktischen Mängeln können hierbei beispielsweise bislang unberücksichtigte Syntaxvarianten der operativen Datenquellen entdeckt werden, die durch technische Spezialisten manuell in den Extrakten berichtigt werden müssen. Auch ein Teil der semantischen Mängel kann automatisiert erkannt werden, indem mit Hilfe von *Plausibilitätskontrollen* und *Domänenüberprüfungen* fehlerhafte Datenfelder identifiziert werden können. Da diese Mängel in jedem Falle auf semantische Inkorrektheiten in den operativen Datenquellen zurückzuführen sind, sind kurz- oder mittelfristig Korrekturmaßnahmen in den operativen Quellsystemen erforderlich. Wenn die Fehler in den Quellsystemen nicht sofort korrigiert werden (können), sollten kurzfristig Bereinigungen in der Filterungsschicht des Data Warehouse durch betriebswirtschaftliche Fachspezialisten durchgeführt werden, damit zumindest im dispositiven Bereich die Fehler der operativen Quellsysteme nicht durchschlagen.

Da eine erforderliche Datensyntax stets vollständig beschrieben werden kann und somit Mängel immer automatisiert erkannt werden können, ist eine manuelle Mängelerkennung lediglich für semantische Mängel sinnvoll vorstellbar. In diesen Fällen ist es

denkbar, dass die extrahierten Daten inkorrekte Datenwerte beinhalten, die ausschließlich durch Analysen von betriebswirtschaftlichen Fachspezialisten erkannt werden können. Da es sich bei diesen semantischen Mängeln ebenfalls immer um Fehler in den operativen Datenquellen handelt, ist in diesem Falle wie oben beschrieben zu verfahren; das heißt, dass kurz- oder mittelfristig die operativen Quellsysteme zu berichtigen sind und bis zur Korrekturimplementierung eine Bereinigung der semantischen Fehler in der Filterungsschicht des Data Warehouse zu erfolgen hat.

7.2.2 Harmonisierung – Die betriebswirtschaftliche Abstimmung gefilterter Daten

Hat eine geeignete Filterung der Daten aus den operativen Systemen stattgefunden, so ist die Voraussetzung geschaffen, aus dem Datenbestand Informationen abzuleiten. Mit einer Übernahme und Bereinigung der operativen Daten ist jedoch noch keine Information mit Entscheidungsrelevanz entstanden, da eine themenbezogene Gruppierung der Daten noch nicht stattgefunden hat. Eine themenbezogene Gruppierung etwa nach Kunde, Produkt oder Organisationseinheit setzt allerdings das Zusammenführen verschiedenster gefilterter und bereinigter Extrakte operativer Datenbestände voraus. Vorbedingung für eine solche Zusammenführung der Daten ist

- die Abstimmung von Kodierungen, Synonymen und Homonymen,
- die Lösung des Problems der Schlüsseldisharmonien sowie
- die Vereinheitlichung betriebswirtschaftlicher Begriffsabgrenzungen.

Die Probleme der Abstimmung von Kodierungen, Synonymen und Homonymen sind anhand von Beispielen in Tab. 7.1 dargestellt.

Selbstverständlich wäre es müßig, über die Eignung der einzelnen Namensgebungen und Kodierungen zu diskutieren. Aus dispositiver Sicht ist vielmehr relevant, dass Feldnamen und Kodierungen inhaltlich harmonisiert werden, so dass bei Auswertungen eine einheitliche Interpretation der Dateninhalte möglich ist (Inmon et al. 2001). Dazu werden i. d. R. Tabellen implementiert, mit deren Hilfe die gefilterten – also die bereits extrahierten und bereinigten – Einzeldateien über Namensabgleichungen und Kodierungsabstimmungen zu themenorientierten Datenbeständen zusammengeführt werden.

Schwerwiegender, weil in der Praxis häufig nicht ohne weiteres lösbar, ist das Problem der Schlüsseldisharmonien. Um systemübergreifende Entitäten zusammenzuführen, ist ein gemeinsamer Primärschlüssel erforderlich. Dieser ist aber nicht immer vorhanden, so dass zunächst ein grundsätzliches Redesign der operativen Systeme erforderlich wäre. Um analytische Auswertungen trotz Schlüsseldisharmonien in den operativen Systemen durchführen zu können, kann eine Vorgehensweise gewählt werden, die eine sukzessive Lösung des Problembereiches erlaubt. Existiert beispielsweise kein einheitlicher Primärschlüssel für den Entitätstyp „KUNDE“ so könnte manuell eine Mapping-Tabelle erstellt

Tab. 7.1 Unterschiedliche Kodierung, Synonyme und Homonyme

	Beispiele			
	Charakteristika	Datenquelle 1	Datenquelle 2	Aktivität
Unter-schiedliche Kodierung	Gleiche Attributnamen; gleiche Bedeutung; unterschiedliche Domänen	Attribut: GESCHLECHT Domäne: (0,1)	Attribut: GESCHLECHT Domäne: (M,W)	Wahl einer Domäne
Synonyme	Unterschiedliche Attributnamen; gleiche Bedeutung; gleiche Domänen	Attribut: PERSONAL Inhalt: Name der Betriebs-angehörigen	Attribut: MITARBEITER Inhalt: Name der Betriebsangehörigen	Wahl eines Attributnamens
Homonyme	Gleiche Attributnamen; unterschiedliche Bedeutung; gleiche oder ungleiche Domänen	Attribut: PARTNER Inhalt: Name der Kunden	Attribut: PARTNER Inhalt: Name der Lieferanten	Wahl unterschiedlicher Attributnamen

Nach der Fusion zweier Versicherungsgesellschaften mit unterschiedlichem Angebotsschwerpunkt sollen geschäftssegmentbezogene Auswertungen für ein Berichtswesen nach Kundengruppen erarbeitet werden. Diese Anforderung ergibt sich unmittelbar aus der Geschäftspolitik, die zum Ziel hat, übergreifende Cross-Selling Potentiale auszumachen – also die Möglichkeit zu nutzen, durch Kenntnis des Kunden aus einem bestehenden Versicherungsvertrag weitere Vertragsangebote zu unterbreiten. Es zeigt sich, dass dieses Vorhaben praktisch undurchführbar ist, da die operativen Systeme zur Vertragserfassung und -abwicklung der beiden vormals eigenständigen Versicherungsgesellschaften über keine einheitlichen Schlüsselstrukturen verfügen. Um diese Problematik zu beheben, müsste ein Projekt zur Neunummerierung der Verträge aufgesetzt werden.

Abb. 7.5 Schlüsseldisharmonien

werden, welche für die Top 100 Kunden der in Abb. 7.5 genannten Versicherungsunternehmen eine 1:1-Schlüsselübersetzung zwischen den operativen Systemen herstellt und damit übergreifende Auswertungen ermöglicht. Sobald Massendaten zu verarbeiten sind, wie etwa im Privatkundengeschäft derselben Versicherungsunternehmen, ist eine manuelle Lösung natürlich nicht mehr möglich.

Gleiches gilt, wenn in den verschiedenen operativen Systemen nicht die notwendigen Schlüssel abgebildet sind, die für übergreifende Auswertungszwecke erforderlich wären. So wäre es beispielsweise für die Analyse von Gewährleistungsmaßnahmen eines Maschinenbauunternehmens wünschenswert, die kundenbezogene Sicht der Gewährleistungsfälle – d. h. der Kunde reklamiert ein schadhaftes Teil – mit der kundenneutralen Sicht der Beschaffung zu koppeln, um den Lieferanten der schadhaften Teile zu identifizieren. Falls im Garantiefallsystem jedoch kein Lieferantenschlüssel und im Beschaffungssystem kein

Kundenschlüssel existiert, ist das schadhafte Teil letztlich nicht eindeutig einem Lieferanten zuzuordnen, sofern mehrere Lieferanten ein baugleiches Teil liefern, das unter derselben internen Teilenummer geführt wird.

Häufig existieren sehr unterschiedliche Meinungen und Definitionen über betriebswirtschaftliche Begriffe und deren datentechnischer Abgrenzung in einem Unternehmen. Ein weiterer wichtiger Aspekt der Harmonisierung ist daher die Vereinheitlichung dieser betriebswirtschaftlichen Begriffe. Dies ist weniger aus technischer Sicht problematisch als vielmehr aufgrund des sich daraus ergebenden kontroversen Diskussionsprozesses, der in einem Unternehmen stattfinden muss.

Ist die Filterung und Harmonisierung abgeschlossen, so liegt im Data Warehouse ein bereinigter und konsistenter Datenbestand auf der Granularitätsebene vor, der bereits für Analytische Informationssysteme direkt nutzbar gemacht werden kann.

7.2.3 Aggregation – Die Verdichtung gefilterter und harmonisierter Daten

Typische Fragestellungen aus dem analytischen Bereich betreffen nicht ausschließlich die Daten der Granularität des Data Warehouse. Häufiger sind vielmehr Auswertungen auf der Basis von Summenstrukturen, die der Beantwortung von Fragen dienen wie: „Wie verlief der Umsatz mit dem Kunden A in der Produktgruppe B über die Monate des vergangenen Halbjahres?“ Diese Frage adressiert beispielsweise die drei Berichtsdimensionen „KUNDE“, „PRODUKT“ und „ZEIT“ auf der Basis aggregierter Größen.

Des Weiteren sind bestimmte betriebswirtschaftliche Kennzahlen überhaupt erst berechenbar, wenn Vorsummierungen stattgefunden haben. Um beispielsweise eine Ist-Plan-Abweichung für den Verkauf zu errechnen, sind die Ist-Verkaufsdaten, die in Form von Einzelartikeln auf Tagesbasis vorliegen können, auf Monats- und Produktgruppenniveau zu summieren, damit diese Ist-Werte für die Abweichungsberechnung mit einer Verkaufsplanung pro Produktgruppe und Monat kombiniert werden können.

Ziel der Aggregation ist demnach, auf Basis konsistenter Dimensionsdefinitionen Summenstrukturen zu erzeugen, die dann aus Gründen der Performanceoptimierung vorberechnet im Data Warehouse gehalten werden können. Typische Berichtsdimensionen sind meist hierarchisch angeordnet und besitzen mit dem Element „Alle“ oder „Gesamt“ eine Spitze, in der alle Einzelelemente aggregiert sind. Solche Dimensionen werden im Data Warehouse getrennt von den Bestandsdaten (gespeichert in sog. „fact tables“) in speziellen Strukturtabellen meist denormalisiert vorgehalten (sog. „dimension tables“) (Kimball und Ross 2013).

Dabei ist offensichtlich, dass die Dynamik der Unternehmenspraxis zu berücksichtigen ist. Veränderungen der Organisationsstruktur durch Unternehmenszukäufe, die Zusammenfassung von Teilmärkten oder die Umstrukturierung des Außendienstes können dazu zwingen, die Zugehörigkeitsbeziehungen in einer solchen Struktur mit einem Zeitstempel

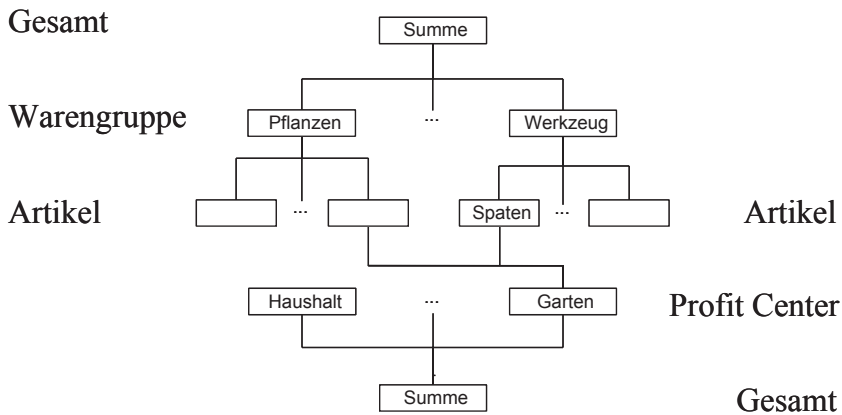


Abb. 7.6 Alternative Hierarchien

abzulegen. Dadurch wird es möglich, konsistente Auswertungen nach der jeweils aktuellen Struktur sowie nach der alten Struktur zu erstellen.

In der Regel werden verschiedene parallele Auswertungsdimensionen je nach Zielgruppe und Auswertungszweck erforderlich sein. Parallele Hierarchien können sich aus unterschiedlichen Sichtweisen auf die Daten ergeben. Zum Beispiel ist ein Kundenmanager eher an Summen auf der Basis von Kundengliederungen interessiert, während ein Produktmanager Summenstrukturen nach Produktgliederungen benötigt. Nicht immer sind die Betrachtungseinheiten der untersten Ebene – wie zum Beispiel das einzelne Produkt – eindeutig einer Hierarchie zugeordnet.

Abbildung 7.6 zeigt das Beispiel eines Warenhauses mit einer produktbezogenen Hierarchie, in der die einzelnen Artikel verschiedenen Warengruppen zugeordnet sind. So existiert beispielsweise die Warengruppe „Pflanzen“ wie auch die Warengruppe „Werkzeug“.

Zur Absatzsteigerung sind komplementäre Produkte in Profit Centers zusammengefasst und werden im Verkaufsbereich nebeneinander angeboten und beworben. In einer Matrixstruktur sind diese darüber hinaus durch einen Profit Center-Verantwortlichen organisatorisch verankert. Im Rahmen des Controllings sollen die Ergebnisse der Profit Centers nun verfolgt werden. Dazu sind diese als gesonderte Hierarchie abzubilden. Beispielsweise sind bestimmte Artikel aus den Warengruppen „Werkzeug“ und „Pflanzen“ dem Profit Center „Garten“ zugeordnet. Dadurch wird die Umsatzzahl des Artikels „Spaten“ sowohl in einem Bericht der Warengruppe „Werkzeug“ als auch in einem Bericht des Profit Centers „Garten“ erscheinen.

Bei einem zyklischen Update der Basisdatenbestände im Data Warehouse sind auch alle Aggregationen über die definierten Strukturen zu aktualisieren.

Insgesamt ist festzuhalten, dass es sich bei dieser Form der Vordefinition von Summenstrukturen um eine verstärkte Hinwendung zu einer Applikationsorientierung der Daten handelt. Applikationsorientiert sind die Daten insofern, als die Strukturen schon auf bestimmte Auswertungsklassen ausgerichtet und für diese optimiert sind. Dies durchbricht die klassische Vorgehensweise der applikationsneutralen Datenmodellierung der Normal-

formenlehre zugunsten einer stärker auf die antizipierbaren Anwendungen ausgerichteten Datenhaltung. Damit ist bei dieser Vorgehensweise die Applikationslogik teilweise bereits in den Daten selbst implementiert.

7.2.4 Anreicherung – Die Bildung und Speicherung betriebswirtschaftlicher Kenngrößen

Stellte die Aggregation bereits eine erste Lockerung des langjährig propagierten Paradigmas der strikten Trennung zwischen Daten und Logik dar, so wird mit der Anreicherung dieses Denkmuster fast vollständig aufgelöst. Mit Hilfe der Anreicherung werden die gefilterten, harmonisierten und teilweise aggregierten Daten um funktionale Aspekte erweitert. Der Grundgedanke der Anreicherung beruht darauf, Berechnungen durchzuführen und die Ergebnisse der Kalkulationen mit den übrigen Daten zu speichern. Hierbei werden bewusst betriebswirtschaftliche Kennzahlen gebildet, die für eine Reihe von potentiellen Informationsnachfragern von Relevanz sind. Die Vorteile dieser Vorgehensweise liegen vor allem

- in einem besseren Antwortzeitverhalten bei späteren Abfragen aufgrund der Vorberechnung,
- in der garantierten Konsistenz der kalkulierten Werte aufgrund der einmaligen Berechnung und
- in der Etablierung eines abgestimmten betriebswirtschaftlichen Instrumentariums.

Abbildung 7.7 verdeutlicht exemplarisch die Anreicherung verschiedener Datensichten. Der untere Teil der Abbildung zeigt einen Ausschnitt aus den gefilterten und harmonisierten Granularitätsdaten.

Diese Daten werden aus gefilterten und harmonisierten operativen Daten gebildet, die hier beispielsweise in Form von tagesaktuellen Rechnungspositionen vorliegen können und vor Übernahme ins Data Warehouse auf Monatsebene verdichtet werden müssen. Die mit breiten Pfeilen gekennzeichneten Datenwerte stellen in dieser Sicht die Anreicherungen dar. Hier im Beispiel wurden die Variablen Kosten und die Deckungsbeiträge ermittelt und in die Datenbasis integriert.

Im Gegensatz zu dem unteren Teil zeigt der obere Abschnitt der Abbildung eine Anreicherung in Form einer filial- und jahresbezogenen Deckungsbeitragsrechnung. Hier erfolgt die Anreicherung demnach auf verdichteten Daten, die auf der Basis der Granularitätsdaten des Data Warehouse im Rahmen der Aggregation gebildet wurden.

Filiale	Jahr	Sum(Umsatz)	Variable Kosten (Ist)	Fixkosten (Ist)	Deckungs- beitrag I
Berlin	2014	20.943.312 DM	16.406.772 DM	2.336.400 DM	2.200.140 DM
Dresden	2014	15.927.384 DM	12.466.512 DM	2.317.200 DM	1.143.672 DM
Essen	2014	18.478.512 DM	14.465.856 DM	2.366.400 DM	1.646.256 DM
....
....

Verdichtete Datenwerte auf Filial- und Jahresebene
(gebildet aus Daten der untersten Granularität)

Anreicherung:
Berechnete Daten

Produkt	Monat	Umsatz	Variable Kosten (Ist) pro ME	Menge	Variable Kosten (Ist)	Deckungs- beitrag
15W-A-BOXEN	Januar 14	67.290 DM	29	1.843	53.447 DM	13.843 DM
15W-A-BOXEN	Februar 14	44.651 DM	29	1.226	35.554 DM	9.097 DM
15W-A-BOXEN	März 14	68.700 DM	31	1.764	54.684 DM	14.016 DM
....
....

Datenwerte der untersten Granularitätsstufe
(z.B. gebildet aus tagesaktuellen Rechnungsbelegen)

Anreicherung:
Berechnete Daten

Abb. 7.7 Anreicherung durch Berechnung betriebswirtschaftlicher Kennziffern

7.3 Aktuelle Konzepte – Datentransformation wird nicht obsolet

Immer wieder werden Diskussionen laut, ob im Lichte aktueller Konzepte und neuer technologischer Ansätze noch ein Bedarf für Datentransformationen existiert. An dieser Stelle sollen einige dieser Überlegungen aufgegriffen und kritisch hinterfragt werden.

In-Memory Überzeugte Verfechter relationaler In-Memory-Datenbanken führen an, durch die Nutzung von In-Memory-Datenbanken werde die Datenüberleitung in eine dispositive Datenhaltung aus Performancegründen obsolet. Dieses Argument ist nicht ganz von der Hand zu weisen. So zeigen aktuelle Ansätze von ERP-Anbietern, dass Online-Auswertungen auf operativen Systemen mit akzeptabler Performance möglich sind, wenn das operative System auf einer In-Memory-Datenbank betrieben wird (z. B. Operational Reporting mit Hana Live der SAP AG). Bei genauer Betrachtung dieser Ansätze wird jedoch deutlich, dass die Auswertungen keineswegs direkt auf den zugrundeliegenden Tabellen des ERP ausgeführt werden, sondern vielmehr auf einem System geschachtelter Views, die eine fachliche Transformation der Auswertungsdaten In-Memory durchführen. Die o. a. Argumentation zur Ableitung entscheidungsrelevanter Informationen gilt also weiterhin. Der Unterschied besteht nur darin, dass keine Speicherung von

Zwischenergebnissen mehr stattfindet. Auch werden in größeren Organisationen stets eine Mehrzahl operativer Systeme auswertungsrelevante Daten beinhalten. Solange diese nicht datentechnisch vollständig integriert sind, kann von einer relationalen Datenbank – sei sie In-Memory oder nicht – keine integrierte Auswertbarkeit dieser Daten erwartet werden.

Ferner ist zu bedenken, dass gerade die Zwischenspeicherung von Transformationsergebnissen explizit notwendig sein kann, etwa wenn historische Wertdarstellungen in aufsichtsrechtlich relevantem Reporting dauerhaft nachweisbar bleiben müssen.

Big Data Big Data wurde als Begriffe vom McKinsey Global Institute geprägt (McKinsey 2011). Im Kern werden darunter Datenmengen verstanden, die mit herkömmlichen Datenbanken nicht mehr verarbeitet werden können. Welche Datenbanktechnologien als „herkömmlich“ gelten, bleibt dabei unscharf.

In der Praxis werden mit Big Data solche Datenbanktechnologien bezeichnet, die besonders auf die Speicherung und zeitnahe Abfrage großer, poly-strukturierter Datenbestände optimiert sind (z. B. Suchoperationen nach Texten auf Basis von Hadoop). Harmonisierte betriebswirtschaftliche Daten bieten diese Speicherungsformen nicht an. Big Data Stores sind insofern als Eingangsschicht für ein Data Warehouse durchaus denkbar. Sie beinhalten jedoch keine konzeptionelle Lösung für eine sinnvolle Filterung, Harmonisierung und Anreicherung.

Cloud Computing Unter Cloud Computing wird die Nutzung von IT-Diensten verstanden, die als virtualisierte Ressourcen in Form von IT-Infrastrukturen (Infrastructure as a Service – IaaS), Anwendungsplattformen (Platform as a Service – PaaS) oder Anwendungssoftware (Software as a Service – SaaS) zur Verfügung gestellt werden. Je nach Offenheit der Cloud-Dienste werden hierbei Private Clouds, Public Clouds und – als Zwischenform – Hybrid Clouds unterschieden.

Konzepte des Cloud Computing bieten somit Infrastrukturen an, mit deren Hilfe unterschiedliche Partner einer Wertschöpfungskette unter Verwendung gemeinsamer Ressourcen (z. B. in Form einer Cloud für eine geschlossene Benutzergruppe) intensive Austauschbeziehungen umsetzen können. So ist es auf diese Weise möglich, dass cloud-basierte DWHs von einer Vielzahl Quellsystemen unterschiedlichster Partner direkt bedient werden (Baars und Kemper 2011). Es liegt auf der Hand, dass hierbei aufgrund der potentiellen syntaktischen und semantischen Heterogenität der Daten den Prozessen der Transformation – insbesondere denen der Filterungs- und Harmonisierungsschicht – wesentlich größere Beachtung zukommen muss als im Kontext traditioneller DWHs.

7.4 Fazit

Wie dargestellt, sind Transformationsprozesse zur Umwandlung von operativen in entscheidungsrelevante Daten zwingend erforderlich, um aussagekräftige Informationen für das Management generieren zu können (Grothe und Gentsch 2000). Auch neue Ansätze –

wie In-Memory-Datenbanken, Big-Data-Ansätze oder Cloud Computing – machen diese Aktivitäten nicht obsolet. Vielmehr erhöhen diese Technologien sogar die Komplexität der Transformationsprozesse, da sie die Vielfalt der Datenformate, die Formen der Datenpersistierung und die Datenlokalität entscheidend verändern. Aus diesem Grunde ist insbesondere die Professionalisierung der unternehmensübergreifenden Data Governance entscheidend, die die Hoheitsrechte zur Festlegung und Durchsetzung methodischer, organisatorischer und technologischer Maßnahmen zur Vereinheitlichung des Datenmanagement festlegt (Mosley und Brackett 2009; Finger 2013).

Weiterhin sollte auch der Metadatenverwaltung ein besonderer Stellenwert eingeräumt werden, da ihr aufgrund der o. a. Symbiose aus Logik und Datenhaltung eine weitaus höhere Bedeutung zukommt als den Data Dictionaries in traditionellen Datenhaltungskonzepten (von Maur und Winter 2003). Die Metadatenverwaltung sollte als zentrales Dokumentations- und Steuerungswerkzeug des Data Warehouse die Navigation innerhalb des Systems erleichtern und detaillierte Informationen zu sämtlichen Filterungs-, Harmonisierungs-, Aggregations- und Anreicherungsaktivitäten benutzeradäquat zur Verfügung stellen. Insbesondere muss für den Anwender des Data Warehouse erkennbar sein, aus welchen Quellen sich die Daten zusammensetzen, welche betriebswirtschaftlichen Kenngrößen abgebildet werden und wie diese Kenngrößen betriebswirtschaftlich zu interpretieren sind. Zusätzlich sind Auswertungsdimensionen sowie deren Hierarchien zu dokumentieren und die Existenz zusätzlicher betriebswirtschaftlicher Anreicherungen auf allen relevanten Verdichtungsausprägungen darzulegen.

Literatur

- Anandarajan, M., Anandarajan, A., Srinivasan, C.A. (Hrsg.): Business Intelligence Techniques. Springer, Berlin (2004)
- Baars, H., Kemper, H.-G.: Ubiquitous Computing – An Application Domain for Business Intelligence in the Cloud? Proceedings of the 17th Americas Conference on Information Systems (AMCIS), 04.08.-07.08.2011, Detroit, USA (2011).
- Bauer, A., Günzel, H. (Hrsg.): Data Warehouse Systeme, Architektur – Entwicklung – Anwendung, 4. Aufl. dpunkt, Heidelberg (2013)
- Chamoni, P., Gluchowski, P., Hahne, M.: Business Information Warehouse. Springer, Berlin (2005)
- Finger, R.: Data Governance und Business Intelligence – eine Einordnung. BI Spektrum. (Feb 2013)
- Grothe, M., Gentsch, P.: Business Intelligence. Aus Informationen Wettbewerbsvorteile gewinnen. Addison-Wesley, München (2000)
- Inmon, W.H.: Building the Data Warehouse, 4. Aufl. John Wiley & Sons, Indianapolis u. a. (2005)
- Inmon, W.H., Imhoff, C., Sousa, R.: Corporate Information Factory, 2. Aufl. Wiley, New York u. a. (2001).
- Jarke, M., Lenzerini, M., Vassiliou, Y., Vassiliadis, P.: Fundamentals of Data Warehouses, 2. Aufl. Springer, Berlin u. a. (2003)
- Kemper, H.-G.: Architektur und Gestaltung von Management-Unterstützungs-Systemen. Stuttgart, Leipzig (1999)

- Kemper, H.-G., Baars, H., Mehanna, W.: Business Intelligence – Grundlagen und praktische Anwendungen, 3. Aufl. Vieweg, Wiesbaden (2010).
- Kimball, R., Ross, M.: The Data Warehouse Toolkit, 3. Aufl. John Wiley & Sons, New York (2013)
- Kimball, R., Reeves, L., Ross, M., Thornthwaite, W.: The Data Warehouse Lifecycle Toolkit. Expert Methods for Designing, Developing and Deploying Data Warehouses. John Wiley & Sons, New York (1998)
- Lehner, W.: Datenbanktechnologie für Data-Warehouse-Systeme – Konzepte und Methoden. dpunkt, Heidelberg (2003)
- McKinsey: Big Data: The Next Frontier for Innovation, Competition and Productivity. McKinsey Global Institute (May 2011)
- Mosley, M., Brackett, M. (Hrsg.): The DAMA Guide to The Data Management Body of Knowledge, 1 Aufl. DAMA International (2009)
- Schütte, R., Rothowe, T., Holten, R.: Data Warehouse Managementhandbuch. Springer, Berlin u. a. (2001)
- Turban, E., Aronson, J.E., Liang, T.-P., Sharda, R.: Decision Support and Business Intelligence Systems, 9 Aufl. Prentice Hall, New Jersey (2010)
- von Maur, E., Winter, R. (Hrsg.): Vom Data Warehouse zum Knowledge Center. Springer, Heidelberg (2002)
- von Maur, E., Winter, R. (Hrsg.): Data Warehouse Management. Springer, Berlin Heidelberg (2003)

Michael Hahne

Inhaltsverzeichnis

8.1 Einleitung 148

8.2 Traditionelle Business-Intelligence-Architekturen 149

8.2.1 Stove-Pipe-Ansatz 149

8.2.2 Data Marts mit abgestimmten Datenmodellen 151

8.2.3 Core Data Warehouse 151

8.2.4 Hub-and-Spoke-Architektur 153

8.2.5 Data-Mart-Busarchitektur nach Kimball 155

8.2.6 Corporate Information Factory nach Inmon 156

8.2.7 Architekturvergleich Kimball und Inmon 158

8.3 Core-Data-Warehouse-Modellierung in Schichtenmodellen 158

8.3.1 Aufgaben und Komponenten in Multi-Layer-Architekturen 159

8.3.2 Eignungskriterien für Methoden der Core-Data-Warehouse-Modellierung 162

8.4 Star-Schema-Modellierung im Core Data Warehouse 164

8.4.1 Granulare Star-Schemata im Core Data Warehouse 164

8.4.2 Bewertung dimensionaler Core-Data-Warehouse-Modelle 167

8.5 Normalisierte Core-Data-Warehouse-Modelle 168

8.5.1 Core-Data-Warehouse-Modellierung in 3NF 168

8.5.2 Historisierungsaspekte von 3NF-Modellen 168

8.5.3 Bewertung 3NF-Modellierung im Core Data Warehouse 170

8.6 Core Data Warehouse mit Data-Vault 171

8.6.1 Hub-Tabellen 172

8.6.2 Satellite-Tabellen 173

8.6.3 Link-Tabellen 176

8.6.4 Zeitstempel im Data Vault 179

8.6.5 Harmonisierung von fachlichen Schlüsseln 180

M. Hahne (✉)
Hahne Consulting GmbH, Bretzenheim, Deutschland
E-Mail: info@hahneconsulting.de

8.6.6	Agilität in Data-Vault-Modellen	181
8.6.7	Bewertung der Data-Vault-Methode	182
8.7	Zusammenfassung	183
	Literatur	184

Zusammenfassung

In der aktuellen Diskussion um den Aufbau effizienter Business-Intelligence-Architekturen mit der Notwendigkeit, die Heterogenität an Strukturen, Daten und Anforderungen zu beherrschen, sind zunehmend Aspekte der Agilität -im Sinn von Time to Market- sowie der Kostensenkung für Entwicklung und Betrieb relevant. Die Datenbewirtschaftungsprozesse im Data Warehouse sind vermehrt im so genannten 24x7 Modus zu betreiben und bedürfen einer Entzerrung unter expliziter Berücksichtigung von Performance-Aspekten auch in Real-Time-Umgebungen.

Durch die zunehmende Globalisierung von Unternehmen entstehen auch immer globalere Data Warehouse Lösungen, in denen eine stringente Trennung in Online-Zeit und Batch-Zeit nach herkömmlichen Mustern über alle Bereiche hinweg nicht mehr anzutreffen ist. Eine Entzerrung der Prozesse der Datenintegration für unterschiedliche Service Level und heterogene Lademuster ist daher gefordert.

Konzepte der Virtualisierung und des Self-Service BI sind in diesem Zusammenhang aktuelle Möglichkeiten, den Herausforderungen zu begegnen. Vor diesem Hintergrund werden in diesem Aufsatz Architektur und Modellierung für Business-Intelligence-Systeme diskutiert.

8.1 Einleitung

Heutige Data-Warehouse-Architekturen sind in mehrfacher Hinsicht mit neuen Anforderungen konfrontiert, die eine klare Architektur sowie adäquate Methoden des Entwurfs fordern. Neben den klassischen Erwartungen an die Kostenstrukturen für den Aufbau und den Betrieb von BI-Systemen sind zunehmend gerade die geschäftlichen Anforderungen, sogenannte Business Requirements, die aufgrund der stetig zunehmenden Dynamik des Geschäftslebens einer starken Volatilität unterworfen sind, wesentliche Treiber für neue BI-Projekte. Gefordert sind Architekturen und Konzepte, die ein schnelles Reagieren auf neue Herausforderungen ermöglichen und somit die Time-to-Market von BI-Applikationen drastisch senkt. Die konzeptionelle Antwort hierauf sind mehrschichtige Data-Warehouse-Architekturen, die als logische Klammer zu verstehen sind (vgl. Hahne 2014).

Eine zukunftsorientierte Architektur für die dispositive Informationsversorgung ist nach heutigem allgemeinen Verständnis im Wesentlichen durch eine mehrschichtige unternehmensweit ausgerichtete Data Warehouse Struktur gegeben (vgl. Hahne und Böttiger 2006; Haupt und Hahne 2007). Im Acquisition Layer erfolgt die Aufnahme der nicht transformierten Rohdaten aus den Quellsystemen und den externen Datenquellen. Ziel

der Transformations- und Harmonisierungsprozesse ist die Ablage im Integration Layer, in dem die aufbereiteten Daten in ihrer vollständigen Historie vorliegen. Die Aggregation hinsichtlich der betriebswirtschaftlichen Anforderungen erfolgt auf dem Weg in den Reporting Layer (vgl. Hahne 2011, S. 60 ff.).

Dabei haben sich in der Praxis verschiedenste Varianten mehrstufiger Architekturen ausgebildet. Je nach Art und Komplexität der Aufgaben der Harmonisierung und Transformation kommen durchaus sehr viele unterschiedliche Stufen der Datenveredelung zum Einsatz. Insofern kann bei Multi-Layer-Architekturen von einem logischen Konzept gesprochen werden, das jeweils unternehmensindividuell auszuprägen ist.

Die aufgeführten Modelle zur Implementierung der Business Rules müssen dabei nicht notwendigerweise persistent ausgeprägt sein. Aktuelle Architektur- und Modellierungsansätze wie die Data-Vault-Methode sehen hier explizit eine Virtualisierung vor (vgl. dazu [Linstedt 2010] und Hultgren 2012). Nicht allein durch diese, aber doch wesentlich von diesen gefördert, ermöglichen die Ansätze der Virtualisierung auch eine bessere Unterstützung von Self-Service-BI.

8.2 Traditionelle Business-Intelligence-Architekturen

Die Architektur von BI-Systemen dient der Beschreibung der wesentlichen Komponenten mit ihren Eigenschaften und Funktionen sowie deren Beziehung untereinander. Dabei sind in der Praxis sehr unterschiedliche Formen und Ausgestaltungen anzutreffen, die nicht immer aufgrund proaktiver Entscheidungen etwa aus einer BI-Organisation heraus entstehen, sondern oft das Ergebnis historisch gewachsener Landschaften sind. Jedoch zeigt sich, dass eine saubere Architektur als Grundlage für die BI-Systeme in Unternehmen Vorteile für Entwicklung und Betrieb mit sich bringt. Dies drückt sich auch in der zunehmend bedeutenden Rolle des BI-Architekten aus.¹

Bekannte Architekturvarianten unterscheiden sich deutlich hinsichtlich der Anzahl der Komponenten, der gesamten Komplexität, des Aufwands für Entwicklung und Betrieb sowie der Performance und Skalierbarkeit. Aber auch die Fähigkeit, effizient und agil mit neuen Anforderungen umzugehen und den Wandel zu unterstützen, ist ein wesentliches Erfolgskriterium für verschiedene Ansätze.²

8.2.1 Stove-Pipe-Ansatz

In den Fällen, in Auswertungen erforderlich sind, kann eine dezentrale Architektur mit unabhängigen Data Marts wie in Abb. 8.1 dargestellt durchaus sinnvoll sein. Bei diesem

¹ Für eine Übersicht der verschiedenen Architekturvarianten siehe auch KeBM14 S. 21 ff. Zu Aspekten der Agilität siehe auch (Göhl und Hahne 2011, S. 12 ff.).

² Die Übertragung der Methoden agiler Softwareentwicklung auf die Domäne Business Intelligence wird eingehend in (Hughes 2008) diskutiert.

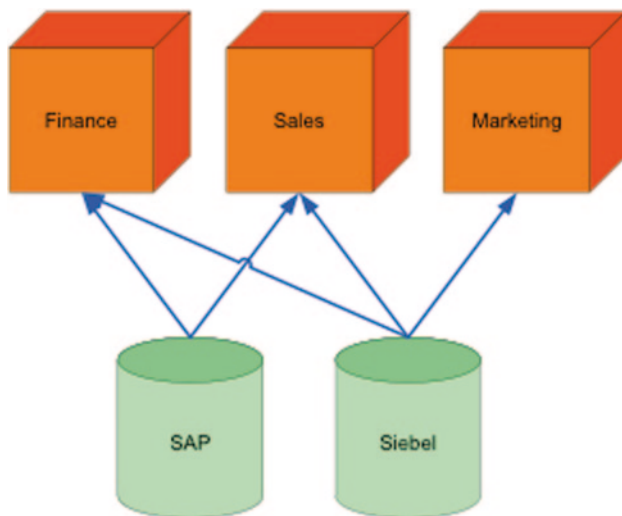


Abb. 8.1 Unabhängige Data Marts

Ansatz, auch unter dem Namen „Stove Pipe“ (Ofenrohr) bekannt (Kimball et al. 2008, S. 249), entstehen einzelne Silos bzw. Inseln, da die Daten für jeden Anwendungsbereich isoliert aus den Quellsystemen extrahiert und aufbereitet werden (Kemper et al. 2014, S. 22 f.). Dabei erfolgt die Transformation der Daten redundant, was auch zu entsprechenden Aufwänden und potenziellen Inkonsistenzen im Falle einer Änderung führt.

Diese Datensilos sind nur eingeschränkt in einem anderen Kontext nutzbar und bieten damit eine sehr schlechte Unterstützung für bereichsübergreifende Auswertungen. Für autonome Organisationseinheiten kann dieser Ansatz aufgrund der leichteren Berücksichtigung fachlicher Anforderungen jedoch durchaus geeignet sein (Sinz und Ulbrich-vom-Ende 2010, S. 189 f.). Oftmals handelt es sich aber um eine rein historisch gewachsene Struktur, die den im Zeitablauf gestiegenen Anforderungen nicht mehr gerecht wird.

Der Übergang von einer derartigen Struktur mit unabhängigen Data Marts zu einer besser geeigneten Architektur mit logischer Integration der Daten gestaltet sich im Allgemeinen recht schwierig, da es auch keine standardisierten bewährten Migrationskonzepte gibt. Dies wird auch durch empirische Untersuchungen untermauert.³

Der Stove-Pipe-Ansatz ist oftmals historisch bedingt und liefert keine Integration, sondern isolierte Data Marts.

³ In einer Studie aus 2006 wurde die Bedeutung und Verbreitung einzelner Architekturformen untersucht. Demzufolge ist die Hub-and-Spoke Architektur mit knapp 40% am weitesten verbreitet, unabhängige Data Marts kamen nur bei gut 10% der befragten Unternehmen zum Einsatz (Ariyachandra und Watson 2006).

8.2.2 Data Marts mit abgestimmten Datenmodellen

Eine erste Möglichkeit zur Entschärfung der Probleme des Stove-Pipe-Ansatzes besteht in der Abstimmung der Data-Mart-Datenmodelle bzw. Dimensionsstruktur (conformed dimensions, siehe Abb. 8.2). Diese erleichtern die Gewährleistung von Konsistenz und Integrität der dispositiven Daten. Neben den Dimensionen sind auch die Kennzahlen abgestimmt, man spricht hier von conformed facts.

Die Abstimmung und der Aufbau eines konsolidierten Datenbestands erfolgt dabei virtuell durch die Abstimmung und Koordination zwischen den Unternehmensbereichen ohne den Aufwand für dessen Entwicklung. Andererseits geht dies einher mit einem erhöhten Aufwand für die Abstimmung (Sinz und Ulbrich-vom-Ende 2010, S. 191).

Auch bei dieser Gestaltungsalternative erfolgen die Transformationen redundant, sodass die Risiken möglicher Inkonsistenzen sowie erhöhte Aufwände im Fall einer Änderung bestehen bleiben.

Die Ausprägung der Data Marts erfolgt typischerweise kontextbezogen, sodass sich diese hinsichtlich der Granularität in allen Dimensionen unterscheiden. Des Weiteren berücksichtigen diese Data Marts ggf. unterschiedliche betriebswirtschaftliche Anreicherungen und basieren auf verschiedenen Aggregationsniveaus in den Dimensionshierarchien. Somit bleiben übergreifende Auswertungen oftmals mit Informationsverlusten verbunden.

8.2.3 Core Data Warehouse

Sind für die analytischen Anwendungen nur Daten aus einer Anwendungsdomäne zu berücksichtigen, kann der Verzicht auf Data Marts eine Alternative sein. Stattdessen ist ein zentrales Core Data Warehouse aufzubauen, auf dem die Analysen direkt erfolgen. Neben

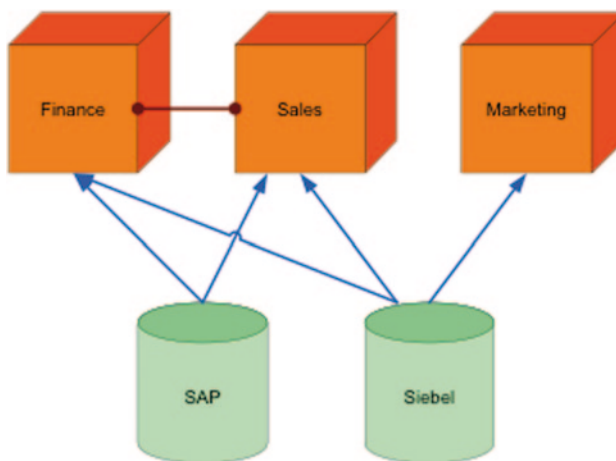


Abb. 8.2 Abgestimmte Data-Mart-Modelle (Conformed)

dem Aspekt der Analyse hat ein Core Data Warehouse auch eine Sammel- und Integrationsfunktion. Es gewährleistet dadurch die Qualitätssicherung und hat eine Distributionsfunktion.

Dieser in Abb. 8.3 visualisierte Architekturansatz stößt hinsichtlich der Anzahl der Benutzer schnell an seine Grenzen und ist kritisch bei einem größeren Datenvolumen, auf dem direkt Auswertungen stattfinden (Sinz und Ulbrich-vom-Ende 2010, S. 188). Aufgrund der fehlenden anwendungsspezifischen Aufbereitung dispositiver Daten ist dieser Ansatz nicht für verschiedene ggf. zu integrierende Anwendungsdomänen geeignet, da es an der kontextbezogenen Aggregation betriebswirtschaftlicher Daten mangelt.

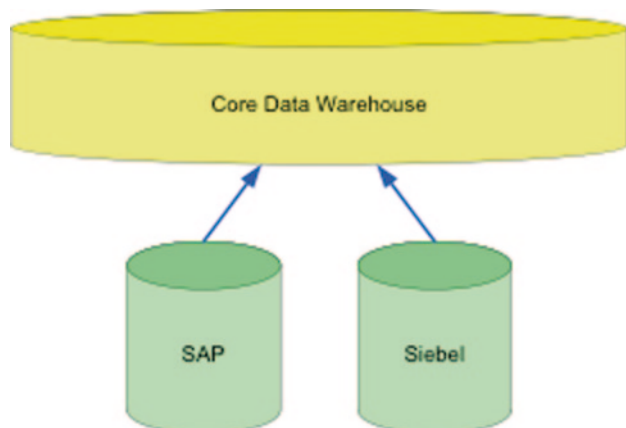
Auf Basis der zentralen Architektur eines Core Data Warehouse sind Betrieb und Pflege des Systems zunächst leichter zu realisieren als bei einer Silo-Architektur. Aufgrund der wenigen Komponenten sind auch Änderungen in den Datenstrukturen direkt für alle Anwendungen sichtbar. Dadurch stehen Änderungen im Rahmen eines Change-Prozesses relativ schnell zur Verfügung. Komplexere Lösungen stoßen aber aus Performance- und Administrationsgründen schnell an ihre Grenzen (Kemper et al. 2014, S. 23).

In einem Core Data Warehouse erfolgt die Speicherung dispositiver Daten nach ersten Transformationsschritten der Bereinigung und Harmonisierung für unterschiedlichste Auswertungszwecke für eine Vielzahl von Benutzern und weist daher einen hohen Grad von Mehrfachverwendbarkeit der Daten zusammen mit einer starken Detaillierung auf.

Gerade die Integrationsfunktion eines Core Data Warehouse ermöglicht Analysen auf abgestimmten harmonisierten Datenbeständen. Jedoch stößt dies bei mehreren Geschäftsfeldern mit stark divergierenden Geschäftsprozessen an seine Grenzen, da eine Integration auf allen Ebenen für alle Bereiche oftmals nicht sinnvoll mit vertretbarem Kostenaufwand realisierbar ist.

In diesen Fällen, in denen einzelne Geschäftseinheiten durch unterschiedliche Produkt- oder Marktstrukturen gekennzeichnet sind, ist der Einsatz mehrerer autarker Core Data Warehouse sinnvoll, die sich jeweils auf die Anforderungen einer strategischen Einheit fokussieren. Dies ist exemplarisch in Abb. 8.4 dargestellt.

Abb. 8.3 Zentrales Core Data Warehouse



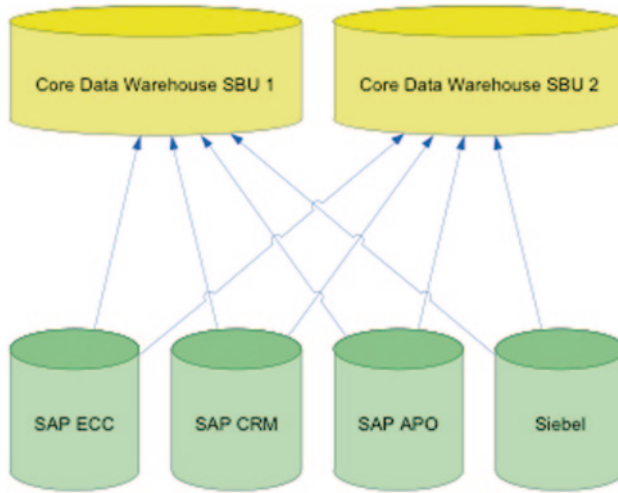


Abb. 8.4 Mehrere Core Data Warehouses

Eine solche Architektur mit mehreren Core Data Warehouse findet sich typischerweise bei Konzernen und spartenorientierten Unternehmen.

In Konzernen und Unternehmen mit sehr unterschiedlichen Geschäftsprozessen sind oftmals einzelne Core Data Warehouse für jede strategische Geschäftseinheit vorzufinden.

8.2.4 Hub-and-Spoke-Architektur

Ein Core Data Warehouse ist eine zentrale Architekturkomponente in vielen Varianten von BI-Architekturen, da es sehr gut geeignet ist, um die folgenden Funktionen zu erfüllen (Kemper et al. 2014, S. 39 f.):

- Sammlung und Integration von dispositiven Daten,
- Distribution, also Verteilung der abgestimmten Daten an nachgelagerte Komponenten wie etwa Data Marts,
- Qualitätssicherung, da die syntaktische und semantische Stimmigkeit der dispositiven Daten durch die Integration und Harmonisierung gesichert ist.

Die Distributionsfunktion kommt insbesondere in der um Data Marts erweiterten Architektur zum Tragen, denn aus dem Core Data Warehouse erfolgt die Bewirtschaftung der Data Marts auf Basis geeigneter Aggregations- und Transformationsprozesse. Oftmals wird diese Form daher auch als Hub-and-Spoke-Architektur bezeichnet.

Die Data Marts sind dabei im Regelfall immer noch unterschiedlich hinsichtlich ihrer Granularität⁴ in allen Dimensionen, der Verwendung unterschiedlicher Formen der Aggre-

⁴ Zum Begriff der Granularität siehe auch [Hahn05, S. 23].

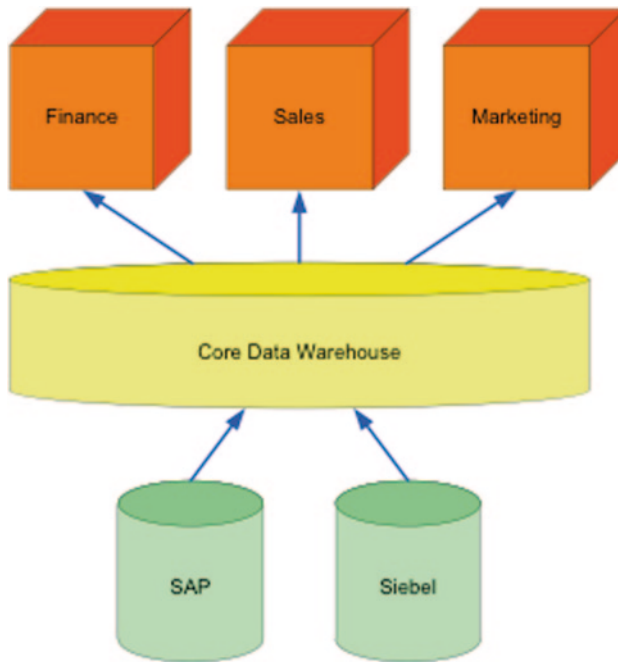


Abb. 8.5 Core Data Warehouse mit abhängigen Data Marts

gation und Nutzung verschiedener Dimensionshierarchien sowie auch bezogen auf die betriebswirtschaftliche Anreicherung. Da sich die Data Marts aus dem Core Data Warehouse ableiten, wird auch von abhängigen Data Marts gesprochen (vgl. GIGD08, S. 129 f.).

Die Vorteile einer solchen Architektur können direkt aus Abb. 8.5 abgeleitet werden, denn es treten viel weniger Schnittstellen auf, sodass die Logik zur Transformation nicht redundant vorzufinden ist. Ein weiterer Vorteil der Architektur liegt in der zentralen Datenintegration und Aufbereitung. Im Wesentlichen ist dies auch die Grundlage des Architekturverständnisses nach Inmon mit der Corporate Information Factory (CIF) (siehe Inmon et al. 2001).

In einer Hub-and-Spoke-Architektur dient das Core Data Warehouse als Hub und erfüllt die Aufgabe der Integration, Qualitätssicherung und Datenverteilung an die Data Marts als Spokes, die einen hohen Grad an Anwendungsorientierung und vordefinierte betriebswirtschaftliche Anreicherungen und Aggregationen aufweisen.

In der Praxis findet sich jedoch häufig ein Mix verschiedener Architekturansätze wie in Abb. 8.6 exemplarisch dargestellt, der teilweise historisch entstanden oder aber das Ergebnis bewusster Gestaltung sein kann.

In dem Beispiel in Abb. 8.6 ist erkennbar, dass teilweise auch virtuelle Data Warehouses (Competition) eingesetzt werden. Diese ermöglichen einen direkten Zugriff auf die Daten des Core Data Warehouse. Dieser Aspekt wird zunehmend kontrovers diskutiert, und die Frage, ob das Core Data Warehouse direkt abgefragt werden darf, ist nicht eindeu-

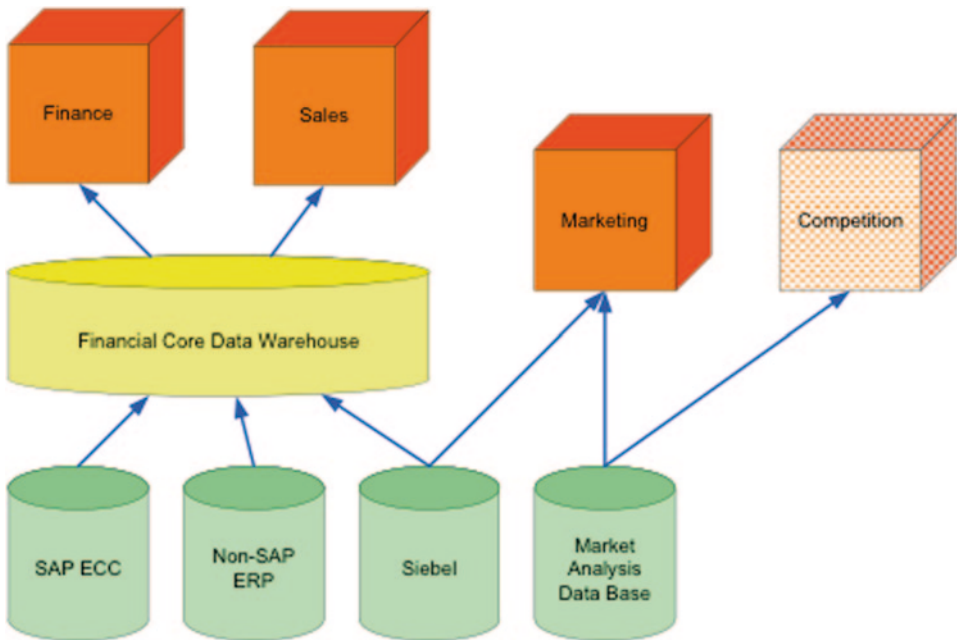


Abb. 8.6 Gemischte Architektur

tig zu beantworten. Aufgrund der negativen Erfahrungen der Vergangenheit findet dieser Ansatz aber immer weniger Zuspruch.

Grenzen der Analyse direkt auf dem zentralen Datenbestand ergeben sich unter anderem durch die zunehmend großen Datenvolumina und die Gefahr von Abfragen, die sehr langsam sind und damit das System sehr stark beanspruchen.

8.2.5 Data-Mart-Busarchitektur nach Kimball

In der Sichtweise nach Kimball sollte ein Core Data Warehouse dimensional modelliert sein (Kimball und Ross 2002, S. 10 ff.). Er nennt es Dimensional Data Warehouse. Es handelt sich hierbei um ein Repository, das sehr wohl für Auswertungen genutzt werden soll bzw. kann. Die einzelnen Data Marts dieser sogenannten Data-Mart-Busarchitektur werden auch Subject Areas genannt.

Bei dieser Architektur gibt es demzufolge ein zentrales Repository, das wie in Abb. 8.7 verdeutlicht, zwar im Sinne eines Hubs die Data Marts bedient, jedoch selbst schon ein mehrdimensionales Modell der integrierten atomaren Daten darstellt. Daher auch der Name Dimensional Data Warehouse für das Repository, das ebenfalls atomare Granularität aufweist.

Wichtig ist in dieser Sichtweise, dass alle Modelle ausschließlich in dimensionaler Form vorliegen und die Auswertungen auf diesem zentralen Repository stattfinden. Zu-

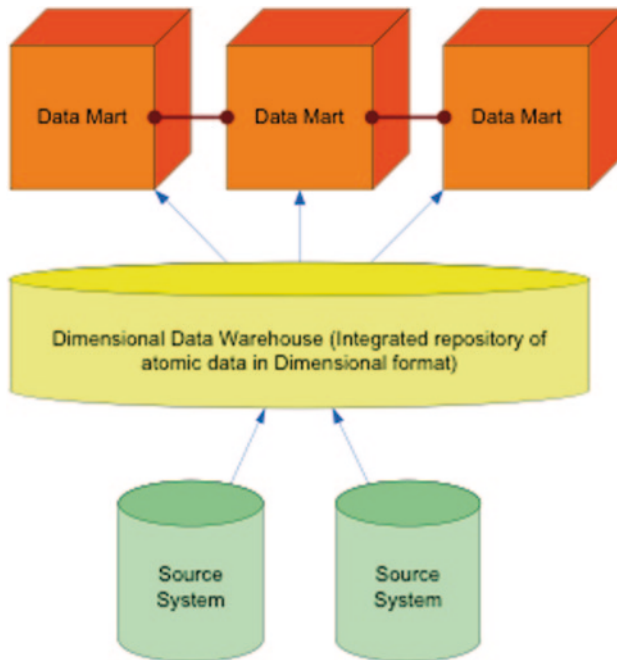


Abb. 8.7 Data-Warehouse-Architektur nach Kimball. (Kimball und Ross 2002)

sätzliche Data Marts müssen nicht notwendigerweise persistiert vorliegen, eine Speicherung ist also optional, die Modelle sind aber alle untereinander abgestimmt und nutzen das Prinzip der „conformed dimensions and facts“ im Sinne einer Enterprise Bus Architecture (siehe Kimball et al. 2008, S. 114 ff.).

In einer Architektur nach Kimball (Kimball und Ross 2002) sind alle Modelle dimensional strukturiert, insbesondere auch das Core Data Warehouse als zentrales granulares Repository.

8.2.6 Corporate Information Factory nach Inmon

Während nach Kimball die dimensionale Modellierung für das gesamte Data Warehouse verpflichtend ist, sieht Inmon deren Nutzen nur auf der Ebene der Data Marts, für die er auch immer den dimensionalen Ansatz empfohlen hat („... if I had to design a data mart tomorrow, I would not consider using any other approach.“, Inmon 2000).

Es sind aber nur die Departmental Data Marts, für die der Ansatz der dimensionalen Modellierung zum Tragen kommt. Diese implementieren ja gerade die abteilungs- oder funktionsbezogene Sichtweise für die Endbenutzerzugriffe. Dabei basieren die zwingend physisch gespeicherten Data Marts auf einheitlichen Strukturen zur Aggregation und Anreicherung (vgl. Inmon et al. 2001, S. 110 ff.).

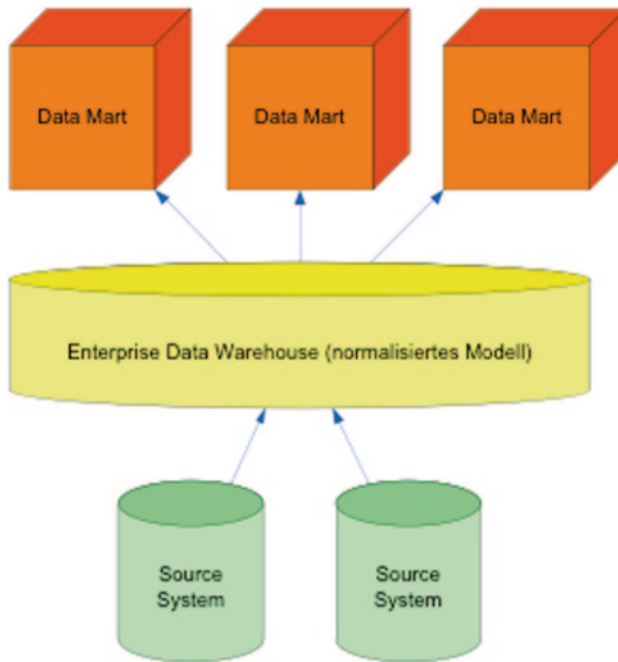


Abb. 8.8 CIF nach Inmon bzw. DW-2.0-Architektur

Eine zentrale Komponente der Corporate Information Factory (CIF) ist das Enterprise Data Warehouse, ein auf einem normalisierten Modell basierendes zentrales Repository von granularen Daten, das als Basis im Sinne eines Hubs innerhalb einer Hub-and-Spoke-Architektur fungiert und auch nicht für direkte Analysen des Endanwenders genutzt werden soll (siehe Abb. 8.8). Die dimensionale Modellierung ist nach Inmon für diesen, wie auch für die meisten anderen Bereiche in der Architektur ungeeignet (vgl. Inmon et al. 2008, S. 18 ff.).

In der Architektur nach Inmon et al. (2008) kommt die dimensionale Modellierung nur für die Data Marts zum Einsatz. Das Core Data Warehouse ist normalisiert modelliert und soll eine möglichst flexible granulare Basis darstellen.

Das Architekturverständnis von Inmon erfuhr im Zeitablauf eine Erweiterung u. a. um Aspekte des Umgangs mit unstrukturierten bzw. semistrukturierten Daten. Diese Architektur wurde als DW 2.0 eingeführt (Inmon et al. 2008).

Aspekte des sogenannten Cross Media Storage Managers und des Nearline Storage wurden zu einem umfassenden Konzept verschiedener Sektoren zur Speicherung basierend auf einem Konzept des Information Lifecycle Management (ILM) ausgebaut (siehe hierzu auch Hahne 2007, 2011). Durch die Einführung verschiedener Sektoren kann den unterschiedlichen Anforderungen an die Service-Levels der Daten im Laufe des Informationslebenszyklus besser Rechnung getragen werden. Neben dem Onlinesektor treten der Nearlinesektor und der Archivsektor hinzu. Aspekte des Information Lifecycle Management im BI-Kontext gehen auf die Inmon-Architektur zurück (Inmon et al. 2008, S. 55 ff.).

8.2.7 Architekturvergleich Kimball und Inmon

Obwohl die öffentlich geführten Diskussion, sicherlich auch durch die beiden Protagonisten selbst bewusst verursacht, sehr kontrovers und polarisierend geführt wird, haben beide Vorstellungen von einer guten Data-Warehouse-Architektur auch Gemeinsamkeiten. Die beiden Architekturvarianten mit ihren zentralen Eigenschaften sind in der Tab. 8.1 gegenüber gestellt (vgl. Adamson 2011, S. 24 ff.).

Zu den Gemeinsamkeiten gehört auch die Einsicht, dass eine logische Integrationsschicht sinnvoll und notwendig ist. In beiden Sichtweisen ist diese Anforderung über ein zentrales Repository abgedeckt. Des Weiteren soll dieses in beiden Fällen atomare Granularität aufweisen. Unterschiedlich ist hingegen die Speicherform dieses zentralen Datenpools. Während in der Kimball-Architektur die dimensionale Modellierung ein zwingendes Kriterium ist, sieht Inmon hier im Wesentlichen normalisierte Modelle vor. Auch der Zugriff, der auf das Dimensional Data Warehouse der Normalfall ist, wird im Fall des Core Data Warehouse nach Inmon im Regelfall ausgeschlossen, da dies durch die dahinterliegenden Data Marts abgedeckt ist.

Ein weiteres Unterscheidungsmerkmal ist die Persistenz der Data Marts. Denn in der Kimball-Sichtweise ist deren physische Speicherung optional. Bei Inmon ist die Persistierung der Data Marts immer vorgesehen, um schlechte Performance bei Zugriffen auf das Core Data Warehouse per se auszuschließen.

8.3 Core-Data-Warehouse-Modellierung in Schichtenmodellen

Zur Beurteilung der Eignung verschiedener Modellierungsparadigmen für die Gestaltung eines Core Data Warehouse ist es notwendig, die einzelnen Komponenten einer Business-Intelligence-Architektur im Kontext der entlang der Datenveredelung stattfindenden Transformations- und Integrationsschritte zu bewerten. Die Schritte der Aufbereitung der Daten in Business-Intelligence-Systemen von der Quelle bis zur Datenstruktur für Reporting und Analyse sind im weitesten Sinne Aufgaben der Datenintegration.

Tab. 8.1 Gegenüberstellung Kimball- und Inmon-Architektur

Architektur	Begriffe	Charakteristika	Mehrdimensionalität
Inmon	Corporate Information Factory, Enterprise Data Warehouse	EDW ist integriert und atomar EDW nicht im direkten Zugriff Data Marts basieren auf EDW und sind physisch getrennt	Nur für Data Marts
Kimball	Dimensional Data Warehouse, Data Mart Bus Architecture	Dimensional Data Warehouse: integriert auf Basis von „conformed dimensions“ im Star Schema Subject Areas im Dimensional Data Warehouse heißen Data Marts Data Marts müssen nicht separat Sein	Alle Daten sind mehrdimensional organisiert

Die im Rahmen der Data-Warehouse-Bewirtschaftung stattfindenden Aufgaben sind gut in das Raster der Teilprozesse der Transformation mit den Phasen Filterung, Harmonisierung, Aggregation und Anreicherung einzuordnen (Kemper und Finger 2010).

In der Phase der Extraktion bzw. Filterung erfolgt die Bereinigung der operativen Daten um syntaktische und inhaltliche Defekte. Das Ergebnis der Harmonisierung sind dann betriebswirtschaftlich abgestimmte gefilterte Daten. In der Phase der Aggregation erfolgt weiter die Verdichtung gefilterter, harmonisierter Daten mit dem Ziel einer wiederverwendbaren Dimensionshierarchisierung. Die Bildung und Speicherung betriebswirtschaftlicher Kennzahlen aus gefilterten, harmonisierten Daten stehen im Fokus der Anreicherung (Kemper et al. 2014).

Die aufgeführten Schritte der Datenintegration sind gut auf das Schichtenmodell übertragbar. Dabei ist es interessant zu differenzieren, in welcher Schicht gewisse Transformationsschritte typischerweise stattfinden sollten (Hahne 2014).

8.3.1 Aufgaben und Komponenten in Multi-Layer-Architekturen

Die heute immer wieder aufgeführten Anforderungen an ein Data Warehouse bzw. dessen wichtigsten Eigenschaften sind:

- Integration und Harmonisierung,
- Nicht-Volatilität und Historie,
- Granularität,
- Prüffähigkeit (Auditability),
- Skalierbarkeit,
- Flexibilität und
- Real-Time-Fähigkeit.

Die Schaffung einer inhaltlich widerspruchsfreien, harmonisierten, integrierten Datensammlung aus den unterschiedlichen operativen und externen Quellen ist eine komplexe Aufgabe. Sie ist aber notwendig, denn ohne die Integration ist ein Data Warehouse hinsichtlich Umfang (Scope) und Wertschöpfung limitiert.

Im Gegensatz zu operativen Systemen sind Daten im Data Warehouse langfristig aufzubewahren, sie werden nicht gelöscht oder geändert. Im Falle der Löschung oder Änderung von Daten im Data Warehouse geht die Prüffähigkeit (Audit) und damit die Möglichkeiten des Abstimmens mit alten Analysen und Quellen verloren. Auf die langfristige Speicherung kann daher nicht verzichtet werden. Darüber hinaus ermöglicht die Aufbewahrung insbesondere granularer Daten die Erkennung komplexer Abhängigkeiten im Zeitablauf und den Einsatz von predictive Analytics. Die Aggregation der Daten auf dem Weg ins Data Warehouse ist eine Einbahnstraße. Die Daten sollten möglichst granular gespeichert vorliegen.

Können alle Data-Warehouse-Daten bis zu ihrer Quelle nachverfolgt werden, unterstützt dies Data-Governance- und Compliance-Anforderungen. Ohne diese Prüffähigkeit ist das Vertrauen in die Informationen limitiert.

Das Data Warehouse muss in seiner Größe hinsichtlich der Anzahl zu integrierender Systeme, Datenvolumen und Nutzer beliebig skalieren können, ansonsten ist die Erfüllung von zukünftigen Anforderungen nur schwer möglich. Auch hinsichtlich der Aktualität ist Skalierbarkeit ein Thema, da zunehmend Anforderungen an eine Datenversorgung in real-time, oder right-time relevant sind.

Die operative Systemlandschaft ändert sich häufig und mit zunehmender Geschwindigkeit. Anforderungen ändern sich immer schneller, daher ist ein Data Warehouse hochgradig volatil. Modelle und Architekturen müssen sich flexibel anpassen können. Ist ein Data Warehouse nicht anpassungsfähig, veraltet es im Zeitablauf und hat einen immer geringeren Nutzen.

Basis der weiteren Ausführungen ist eine erweiterte mehrschichtige Referenzarchitektur, die den Aspekten der verschiedenen Typen von Aufgaben der Transformation und Harmonisierung gerecht wird. Gängig ist eine Betrachtung der Komponenten in der Reihenfolge, in der die Daten von der Quelle kommend diese passieren. In Abb. 8.9 ist dieser Fluss der Daten nach der in Europa gängigen Praxis von unten nach oben dargestellt. In der englischen Literatur wird dies allerdings als *Downstream* bezeichnet. *Upstream* ist demzufolge die Richtung hin zur Quelle und *Downstream* geht in Richtung Data Marts. Dies ist die gängige Beschreibung für die Richtung eines Prozesses, z. B. *Downstream* für die Richtung zum Kunden und *Upstream* für die Herkunft eines Produktes.

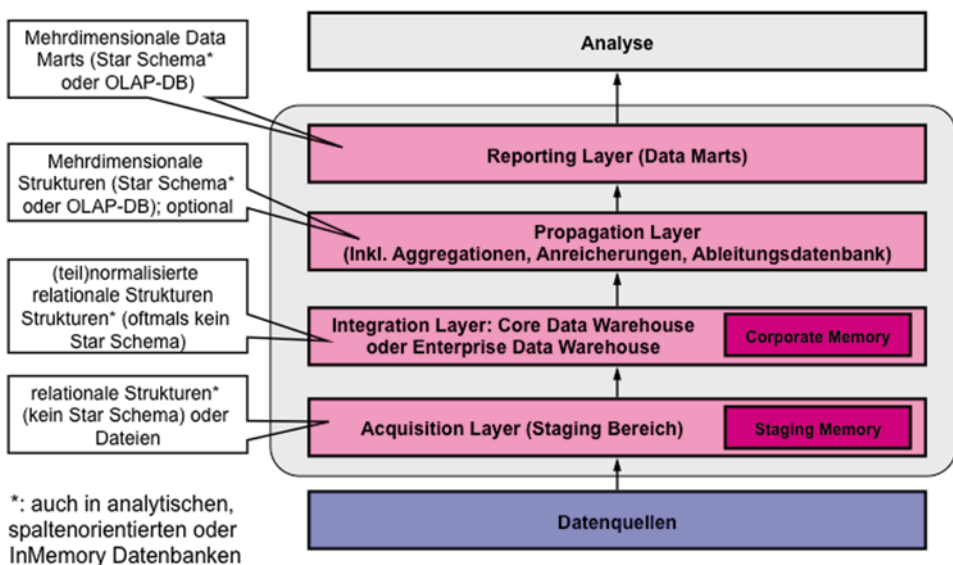


Abb. 8.9 Erweiterte konzeptionelle Multi-Layer-Architektur

Erste Komponente entlang der Datenveredelung ist der Staging-Bereich, dessen Aufgabe die temporäre Zwischenspeicherung extrahierter Rohdaten vor der Integration ist. Oftmals ist es notwendig, den Bedarf einer erneuten Extraktion aus den Quellsystemen zu vermeiden, da diese u. U. nicht zeitnah zu erledigen ist oder sogar die benötigten Daten in der Quelle bereits gelöscht sind. Um dies abzufangen, kann ein sogenanntes *Corporate Staging Memory* als Extraktionshistorie eingeführt werden.

Weitgehende Einigkeit herrscht in der Literatur darüber, dass ein Staging-Bereich nicht für Abfragen seitens der Endanwender gedacht ist. Für Auswertungen zur Datenqualität und Datenabstimmung sind aber technische Abfragen möglich.

Die wichtigsten Aufgaben eines Staging-Bereichs sind die Vermeidung evtl. erneuter Extraktion aus den Quellen, die schnelle Datenaufnahme im Data Warehouse (Durchlauf-erhitzer) sowie eine möglichst geringe bzw. gar keine Transformation oder Veränderung der Quelldaten.

Im Core Data Warehouse werden die Daten auf ein festgelegtes Niveau der Granularität überführt. Nach heute allgemein akzeptiertem Verständnis wird dabei im Regelfall die Beleggranularität gefordert. Da in dieser Komponente die gesamte relevante Historie der Unternehmensdaten des Data Warehouse gespeichert werden, kann dies ein recht großes Datenvolumen ergeben. Dieser Single Point of Truth (SPOT) ist oftmals unternehmensweit ausgerichtet, bei Konzernen und spartenorientierten Strukturen aber eher nur über eine strategische Geschäftseinheit (Business Unit) hinweg.

Die wichtigsten Aufgaben des Core Data Warehouse sind die Bereitstellung der gesamten relevanten Unternehmensdaten in granularer vollständiger integrierter Form (Sammel- und Integrationsfunktion), die Distributionsfunktion als Datenlieferant (Hub) für nachgelagerte Schichten sowie die Qualitätssicherungsfunktion.

Eine weitere wichtige Aufgabe des Core Data Warehouse ist die Historisierung der Daten, also die Möglichkeit der Nachverfolgbarkeit der Änderungen von Hierarchien, Attributen oder Ausprägungen im Zeitablauf.

In einigen Branchen hat sich mittlerweile die bitemporale Historisierung durchgesetzt, der zufolge die Zeiten der fachlichen Gültigkeit von der technischen Gültigkeit zu differenzieren sind, oftmals auch als Validity Time (VT) und Transaction Time (TT) bezeichnet (s. Abelló and Martín 2003). Dies findet sich insbesondere in den Bereichen, in denen das operative System eine fachliche Gültigkeit pflegt und diese dort auch rückwirkend oder in die Zukunft hin verändert werden kann. Ein gutes Beispiel hierfür sind Verträge eines Versicherungsunternehmens.

Die Strukturen im Core Data Warehouse erhalten dann nicht nur eine zeitliche Zuordnung sondern neben die fachliche Gültigkeit kommt die technische Gültigkeit aus der Perspektive der Kenntniserlangung im Data Warehouse hinzu. Formal ist damit auch eine dritte zeitliche Zuordnung möglich, die dann den Änderungszeitpunkt im operativen System widerspiegelt, zuweilen als tritemporale Historisierung bezeichnet.

Ist nur die Zeit der Kenntniserlangung im Data Warehouse gegeben, wird hieraus im Regelfall die fachliche Gültigkeit abgeleitet. Die Grundprinzipien der bitemporalen His-

torisierung im Data Warehousing entstanden in Anlehnung an das Konzept bitemporaler Datenbanken (vgl. dazu Snodgrass 2000).

In den Fällen, in denen die Quelle die fachliche Gültigkeit liefert, kann diese von der zeitlichen Ausprägung der Kenntniserlangung im Data Warehouse abweichen. Insbesondere rückwirkende Änderung sowie Änderungen in die Zukunft sind dabei anzutreffen.

Wesentlich ist in diesem Modell, dass es zwei zeitliche Perspektiven gibt und dass es demzufolge möglich ist, einerseits Berichte und Analysen zu jedem Zeitpunkt eines Kenntnisstands im Data Warehouse zu reproduzieren sowie andererseits davon entkoppelt die fachliche Gültigkeit von Attributausprägungen zu betrachten (vgl. dazu Hahne 2014, S. 172 ff.).

Die Anforderungen an eine bitemporale Historisierung sind nicht von den herkömmlichen Implementierungsmustern der Slowly-Changing-Dimensions abgedeckt, da diese ja nur eine zeitliche Dimension abdecken (s. Johnston und Weis 2010, S. 42 ff.). Das Verfahren der vollständigen Zeitstempelung ist jedoch entsprechend erweiterbar. Dadurch können die Anforderungen an eine Historisierung unter Berücksichtigung der beiden zeitlichen Dimensionen Fachliche Gültigkeit und Technische Gültigkeit antizipiert werden.

Der Propagation Layer ist die Ebene der Datenbereitstellung, der insbesondere dann notwendig ist, wenn eine einheitliche Dimensionshistorisierung und einheitliche Kennzahlenberechnung für die mehrfache Verwendung in verschiedenen Data Marts erreicht werden soll. In diesem Schritt erfolgt auch die Implementierung der Geschäftslogik bzw. der fachlichen Regeln (*Business Rules*) mit dem Ziel abgeleiteter Merkmale und Kennzahlen.

Die zentrale Kapselung der Anwendung von Geschäftsregeln bzw. Business Rules ist eine zentrale Aufgabe des Propagation Layer. Die Daten werden dort in Form eines mehrdimensionalen Modells für Analyse und Auswertungen bereitgestellt.

Der Reporting Layer stellt die Data Marts bereit. Die notwendige Agilität in der Erstellung und Änderung von Data Marts als Folge von Änderungen geschäftlicher Anforderungen oder Sichtweisen ist dabei berücksichtigt, sodass auch von Data Marts on Demand gesprochen werden kann. Insbesondere sind hierdurch Aspekte des Self Service Business Intelligence im Kontext volatiler fachlicher Anforderungen abgedeckt.

Im Reporting Layer werden als wichtigste Aufgabe die für Analyse und Reporting notwendigen mehrdimensionalen Strukturen unter Berücksichtigung von Agilitätsaspekten bei sich ständig ändernden Anforderungen bereitgestellt.

8.3.2 Eignungskriterien für Methoden der Core-Data-Warehouse-Modellierung

Die möglichen Ansätze unterschiedlicher Herangehensweisen an die Gestaltung des zentralen Repositoriums folgen entweder dem Star-Schema-Ansatz oder basieren auf einer Form der Modellierung in einer Normalform, wobei die klassische dritte Normalform mit ihren 3NF-Modellen durchaus häufig anzutreffen ist. Damit bieten sich mehrere Optionen

zur Gestaltung eines Core Data Warehouse, die im spezifischen Einzelfall abzuwägen sind.

Die Modellierungsoptionen für ein Core Data Warehouse und die damit verbundenen Architekturvarianten sollen anhand der folgenden Kriterien beurteilt werden:

- Deckt das Core Data Warehouse bereits die Aufgaben einer Staging Area mit ab?
- Sind die wesentlichen Aufgaben eines Core Data Warehouse vollständig abgedeckt?
- Sind mit dem Core Data Warehouse auch die Anforderungen eines Propagation Layer bereits berücksichtigt?
- Erfolgt die Abdeckung der Aufgaben des Reporting Layer ebenfalls im Core Data Warehouse?
- Die im Data Warehouse anzutreffenden Change-Prozesse stellen Anforderungen an die Flexibilität im Umgang damit. Ist die Methodik der Core-Data-Warehouse-Gestaltung für einen effizienten Umgang mit diesen geeignet?⁵

Das Staging-Modell aus Abb. 8.10 dient als Grundlage der Bewirtschaftung eines Core Data Warehouse und damit auch der Analyse der Eignung einer Methodik für die Modellierung dieser Schicht. Dieses Modell bildet die Fachlichkeit von Einkaufsbelegen ab, die in ihre Kopf- und Positionsdaten aufgeteilt sind. Daneben gibt es noch weitere Entitäten zur Berücksichtigung des Lieferanten, der in drei Rollen mit dem Einkaufsbeleg in einer Beziehung steht, sowie dem Material als Attribut einer Einkaufsbelegposition.

Die Kennzahlen des Modells sind dabei auf Ebene des Belegkopfs (Anzahl Rechnungen und Anzahl Lieferungen) und auch auf Ebene der Position (z. B. Bestellmenge und Nettopreis) abgelegt. Weitere Attribute sind darüber hinaus in allen Entitäten zu finden.

Aspekte der Flexibilität im Umgang mit Änderungen im Sinne eines Change-Prozesses betreffen im Data Warehouse immer zwei Seiten. Neben den Änderungen, die aus neuen Business Requirements resultieren, sind insbesondere im Core Data Warehouse Frage-

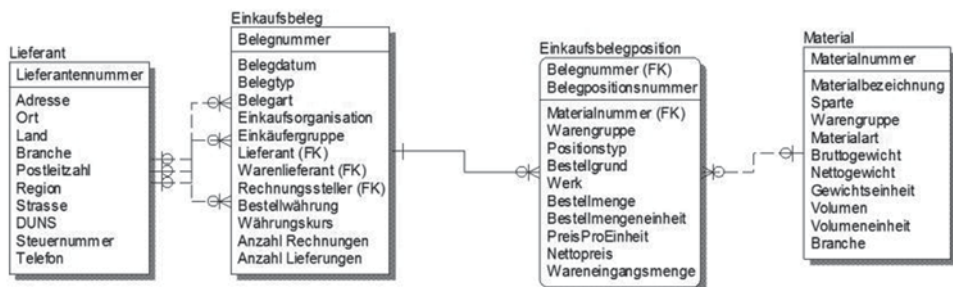


Abb. 8.10 Staging-Modell für Einkaufsbelege

⁵ Eine ähnliche Betrachtung in (Imhoff et al. 2003 S. 389 ff.) betont die Aspekte der Komplexität und des Aufwands für Wartung und Weiterentwicklung.

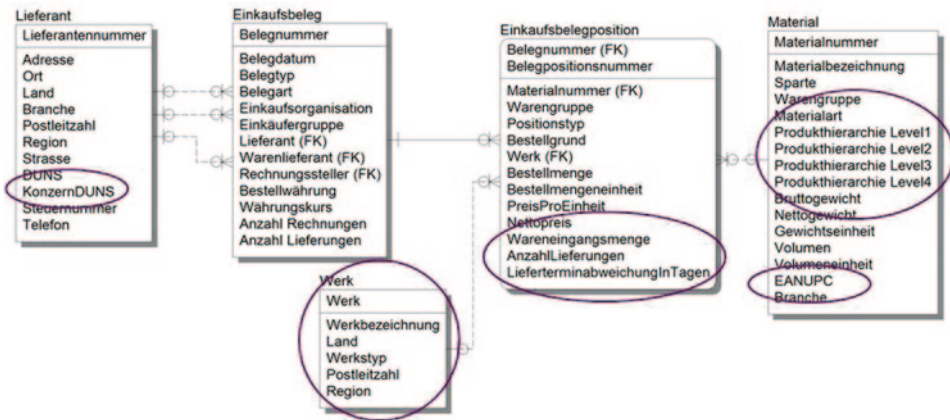


Abb. 8.11 Geändertes Staging-Modell für Einkaufsbelege

stellungen interessant, die sich mit dem Change-Prozess bei Änderungen auf der Quellsystemseite ergeben. Dies soll anhand des geänderten Staging-Modells aus Abb. 8.11 für die einzelnen Optionen der Core-Data-Warehouse-Gestaltung betrachtet werden.

Dieses geänderte Staging-Modell hat einerseits neue Attribute und andererseits eine neue Tabelle verknüpft über eine Beziehung. Die Tabelle Lieferant hat das neue Attribut KonzernDUNS sowie das neue Attribut Steuernummer. Die Tabelle Material beinhaltet mehrere neue Attribute u. a. aus der Produkthierarchie. Alle Änderungen sind in Abb. 8.11 hervorgehoben.

8.4 Star-Schema-Modellierung im Core Data Warehouse

Im Verständnis der Kimball-Architektur eines Data Warehouse ist ein Star Schema – Best Practice bei relationalen Data Marts – auch die Grundlage für den Aufbau eines zentralen Repositoriums, dem Core Data Warehouse, das auch Dimensional Data Warehouse oder auch Data-Mart-Busarchitektur genannt wird.

8.4.1 Granulare Star-Schemata im Core Data Warehouse

Für eine zentrale Architekturkomponente, die als Single Point of Truth fungiert, ist es dabei wichtig, keine unnötigen Aggregationsschritte vorwegzunehmen und dadurch die weitere Verwendung der harmonisierten Daten einzuschränken. Somit ist auch ein dimensional aufgebautes Core Data Warehouse granular bzw. belegorientiert aufzubauen. Dies ist stets eine Frage des Abwägens, Best Practice ist jedoch, allgemein im Core Data Warehouse die granularste Ebene abzubilden.

Übertragen auf ein Star-Schema bedeutet dies, dass die Modelle atomar sein sollten (s. Kimball et al. 2008, S. 134) und damit zumindest eine Belegdimension sowie ggf. eine

weitere Dimension für die Positionen der Belege haben müssen. Diese können nach den bereits ausgeführten Regeln der Gestaltung als degenerierte Dimension ausgeprägt sein. Dies hat sich in der Praxis auch gut bewährt (s. Kimball und Ross 2002, S. 50 f.).

Die im Rahmen einer normalisierten Modellierung sich ergebenden abhängigen Attribute eines Belegs sind im Star-Schema jedoch nicht als Attribut einer solchen Dimension aufgebaut, sondern finden vielmehr Eingang in die Definition von weiteren Dimensionen.

Aus der Diskussion der Berücksichtigung von Aspekten der Historisierung ist bekannt, dass die Verknüpfung einer Dimension zur Faktentabelle eine *as-posted*-Sicht der granularen Ebene der Dimension darstellt. In der Konsequenz muss bei der Bewirtschaftung der Faktentabelle für jeden Satz eine Dimensionsausprägung für alle Dimensionen zugeordnet werden. In Abb. 8.12 ist ein Star-Schema für die Einkaufsbelege unseres Falls nach der Änderung des Staging-Modells dargestellt. Dieses beinhaltet die Dimensionen Einkauf, Währung, Werk, Datum, Belegart, Lieferant in den Rollen Lieferant, Warenlieferant und Rechnungssteller sowie noch die degenerierte Dimension mit der Belegnummer. Dies ist noch nicht die unterste Stufe der Granularität, denn die Positionsebene der Belege kommt in einem weiteren Modell getrennt hiervon hinzu. Diese sind im Regelfall getrennt abzubilden, da eine Vermischung von Kennzahlen auf Belegebene mit Kennzahlen auf der Positionsebene das Risiko falscher Abfrageergebnisse mit sich bringt.

Ändert sich im Zeitablauf in dem Staging-Modell ein Satz z. B. in der Form, dass der Warenlieferant geändert wird, muss der Satz im Star-Schema korrigiert werden. Es gibt

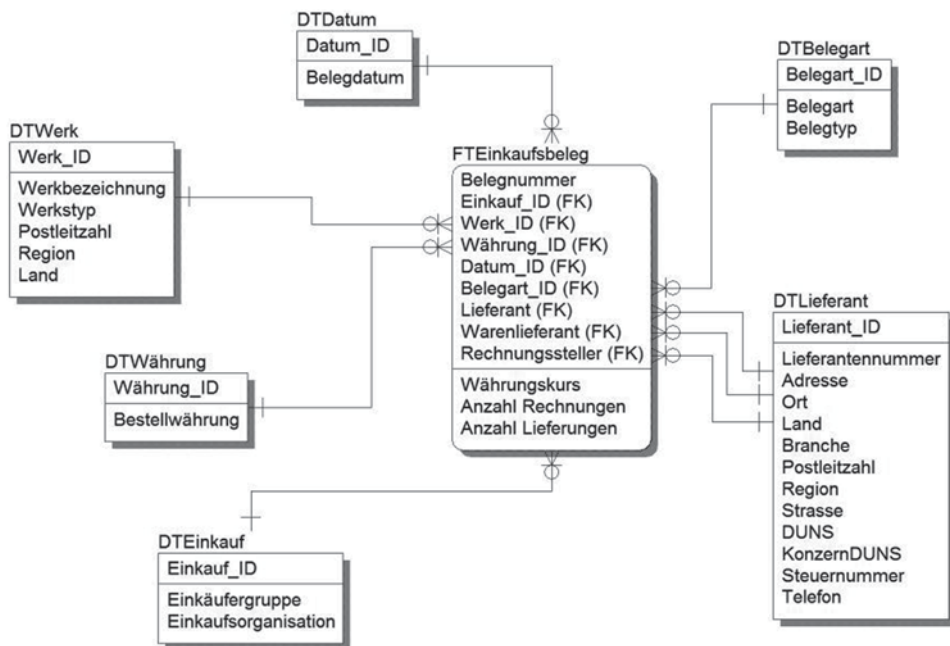


Abb. 8.12 Einkaufsbelege im Star-Schema-Core-Data-Warehouse

also Situationen, in denen Änderungen der Daten im Staging-Bereich auch zu Änderungen im Star-Schema führen, die nicht alleine durch Hinzufügen neuer Sätze abzubilden sind.

Die Abbildung der Ebene der Positionen des Einkaufsbelegs ist in Abb. 8.13 zusammengefasst. Die Dimensionen des Modells für die Belegkopfdaten können einfach übertragen werden. Zusätzlich gibt es noch die Dimensionen Bestellgrund und Positionstyp, die sich direkt aus den Attributen der Einkaufsbelegposition im Staging-Modell ergeben. Die weiteren Änderungen betreffen im Wesentlichen die Faktentabellen da ganz andere Kennzahlen auf Ebene der Position abzubilden sind. Diese sind z. B. die Anzahl Lieferungen und die Wareneingangsmenge. Dies sind Größen, die sich als Attribute der Position im Zeitablauf ändern, was sich im Staging-Modell durch einen neuen Satz ausdrückt, der nun im Core Data Warehouse zu adaptieren ist. Dies führt gemeinhin zu dem Bedarf einer Änderung von Faktensätzen im Star-Schema.

Um die granular strukturierten Star-Schema-Modelle im Core Data Warehouse weniger anfällig gegenüber Datenänderungen auszugestalten, ist es möglich, Aspekte der Historisierung von Fakten in Form einer Ladezeitdimension zu berücksichtigen. Dies erhöht jedoch die Komplexität der Bewirtschaftung und der Abfrage solcher Modelle.

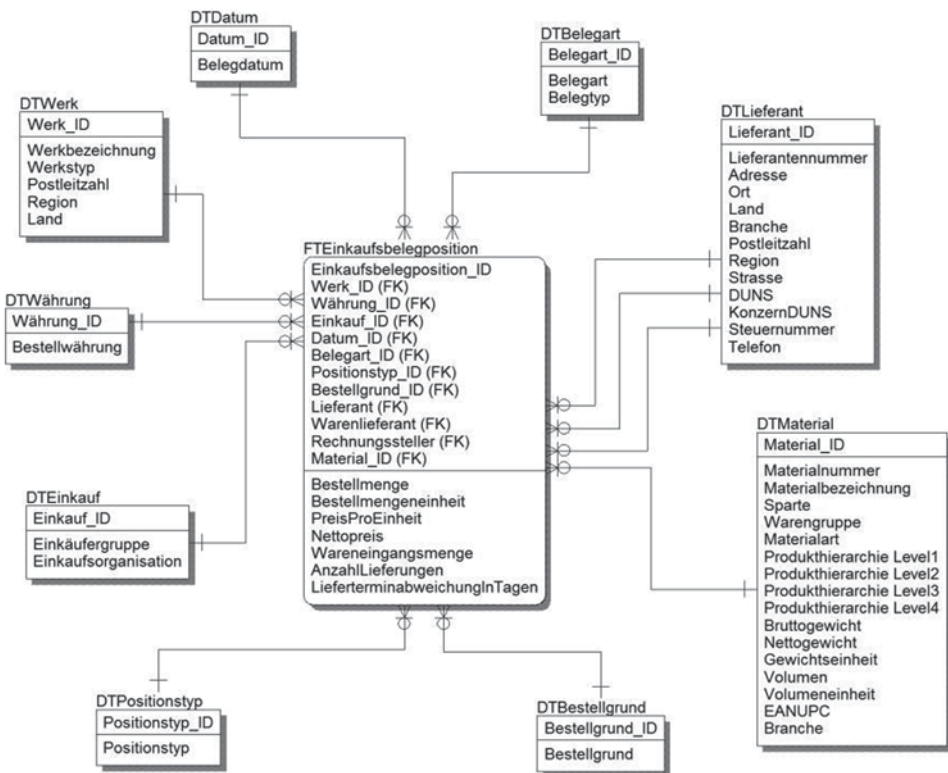


Abb. 8.13 Einkaufsbelegpositionen im Star-Schema-Core-Data-Warehouse

8.4.2 Bewertung dimensionaler Core-Data-Warehouse-Modelle

In einem Star-Schema sind die Daten bereits umfassenden Transformationen und verschiedenen Schritten der Datenaufbereitung unterzogen worden, sodass die Anforderungen an einen Staging-Bereich nicht mit abgedeckt werden können. Vielmehr ist es notwendig, dass dieser Staging-Bereich mit einer persistenten Extraktionshistorie ausgestattet ist, damit die erneute Extraktion aus den Quellsystemen vermeidbar ist.

Die Aufgaben eines Core-Data-Warehouse-Modells mit den Funktionen der Sammlung, Harmonisierung und Distribution im Sinne eines SPOT sind durch derartige Modelle erfüllbar, insofern ist der Ansatz der Star-Schema-Modellierung für ein Core Data Warehouse für dessen Kernaufgaben als geeignet anzusehen.

Ein wichtiger Aspekt in der Beurteilung der Eignung eines Modellierungsparadigmas für ein Core Data Warehouse ist der Umgang mit Änderungen insbesondere aufgrund geänderter geschäftlicher Anforderungen oder Quellstrukturen. Die Änderungen fachlicher Anforderungen sind oftmals durch eine geänderte Businesslogik abzudecken. Im Star-Schema bedeutet dies bereits ein Modelländerung, da die Businesslogik Bestandteil der Transformationsschritte zum dimensionalen Modell ist.

Aber auch die Änderung von Quellstrukturen führt im Regelfall zu umfassenden Änderungen des Star-Schema-Modells da z. B. neue Attribute hinzukommen oder auch neue Tabellen im Staging-Modell entstehen. Auch für die Prozesse zur Datenintegration zwischen dem Staging-Modell und dem Core-Data-Warehouse-Modell hat dies weitreichende Änderungen zur Konsequenz. Daher impliziert eine Änderung der Quellmodelle auch immer weitreichende Änderungen im Core Data Warehouse und dessen Bewirtschaftung. Es entsteht somit ein großer Aufwand für Entwicklung und Test, da alle bereits produktiv gesetzten ETL-Prozesse potenziell betroffen sind und daher recht viele Dinge erneut zu testen sind.

Da in einem Star-Schema ein Großteil der Businesslogik bereits berücksichtigt ist, ergibt sich auf der anderen Seite ein geringer Anpassungsaufwand für die Prozesse der Datenintegration in Richtung Reporting-Layer.

Die Aspekte der unitemporalen Historisierung sind durch die Ansätze der Zeitstempelung und allgemein durch die Verfahren der Implementierung von Slowly-Changing-Dimensions abgedeckt. Bitemporale Historisierung ist nicht Kernbestandteil des Star-Schema-Ansatzes und bedarf besonderer Modellierungsvarianten.

Ein auf Basis des Star-Schema-Ansatzes gestaltetes Core Data Warehouse hat tendenziell einen geringen Grad an Agilität im Umgang mit Change-Prozessen. Auch die damit einhergehenden Aufgaben für Entwicklung und Test weisen eine recht hohe Komplexität auf.

8.5 Normalisierte Core-Data-Warehouse-Modelle

Aspekte der Normalisierung sind im Kontext der Snow-Flake-Schemata Standard der relationalen Data Mart Gestaltung. Dabei hat sich aber gezeigt, dass die Normalisierung im Star-Schema und die sich daraus ergebenden Modelle oftmals weniger sinnvoll sind. Dies ist jedoch gänzlich anders im Fall der Gestaltung eines Core Data Warehouse.

8.5.1 Core-Data-Warehouse-Modellierung in 3NF

Die Modellierung des Core Data Warehouse nah an einer dritten Normalform entsprechend der Lehre zur Datenmodellierung impliziert eine Ausrichtung auf die Kerninformationsobjekte und führt damit zu einem fachlichen Modell mit den wesentlichen Objekten und deren Beziehungen zueinander. Es fördert somit das Verständnis für die Zusammenhänge zwischen den fachlichen Objekten.

Für das Beispielmodell mit Einkaufsbelegen und deren Positionen, noch vor den Änderungen, ergibt sich im ersten Ansatz das Modell aus Abb. 8.14 mit seinen zwei Kernentitäten für die Kopf- und Positionsdaten.

Die Attribute sind nicht weiter aus diesen beiden Entitäten herausgebrochen. Eine weitere daran anknüpfende Entität in dem Modell könnte dann z. B. der Artikel sein.

8.5.2 Historisierungsaspekte von 3NF-Modellen

Auf Basis eines 3NF-Modells können die meisten Anforderungen an ein Core Data Warehouse sehr gut abgebildet werden. Jedoch führt die Berücksichtigung von Aspekten der Historisierung zu Modellen, die sehr komplex und schlecht zu lesen sind, da dies insbesondere many-many-Beziehungen an Stellen impliziert, wo fachlich lediglich eine klassische 1-zu-n-Beziehung existiert.

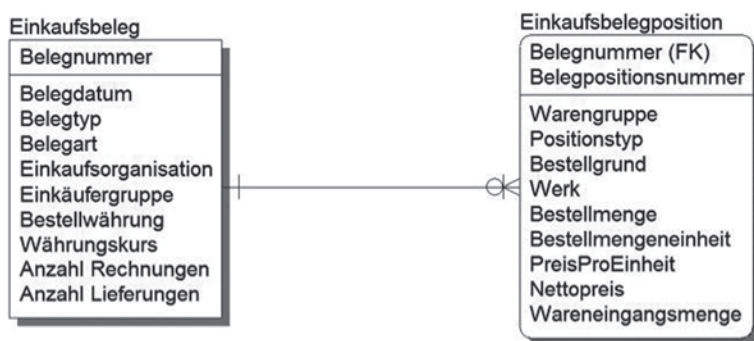


Abb. 8.14 3NF-Beziehung zwischen Beleg und Position

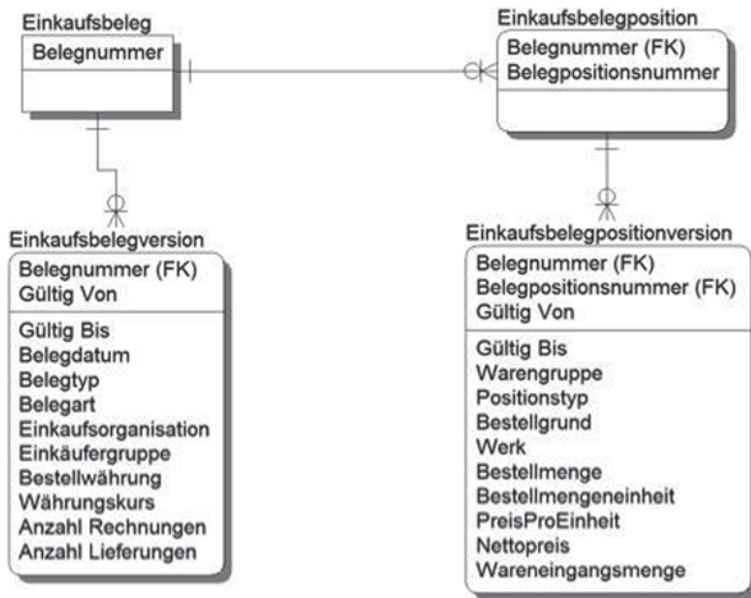


Abb. 8.15 Abbildung von Versionen zur Historisierung im 3NF-Modell

Es hat sich eine Modellierungsvariante bewährt, die diese negativen Aspekte ausgleicht. Dabei ist ein wesentlicher Aspekt in der Trennung der fachlichen Beziehung von den Aspekten der Historisierung zu sehen. Abb. 8.15 zeigt diese unter dem Stichwort Ankermodellierung bekannten Ansatz für das Beispiel der Einkaufsbelege.⁶

Die zugrunde liegende fachliche Beziehung zwischen Beleg und Belegposition ist vom Typ der 1:n-Kardinalität direkt ersichtlich und besteht lediglich aus den beiden Entitäten ohne weitere Attribute. Diese Kernentitäten heißen auch Anker.

An die Anker angehängt sind die Entitäten für die Versionen, also die einzelnen zeitlich abgegrenzten Attributausprägungen. Für die Versionen des Einkaufsbelegs ergibt sich dann als Schlüsselattribut neben der Belegnummer selbst in der Versionsentität auch noch ein Zeitstempel, in diesem Fall das *Gültig Von*-Datum. Somit sind verschiedene Ausprägungen der Attribute eines Einkaufsbelegs über die zeitliche Gültigkeit abgegrenzt. Für die Positionen des Einkaufsbelegs ergeben sich in der Version die Schlüsselattribute *Belegnummer*, *Belegpositionsnummer* und *Gültig Von*.

Zwischen dem Anker und seinen Versionen existiert immer eine 1:n-Beziehung. Diese hätte bei einer Auflösung hin zur tragenden Entität, im Beispiel Einkaufsbeleg und Einkaufsbelegposition, die Konsequenz einer sich daraus ergebenden many-many-Beziehung

⁶ Die Ankermodellierung ist nicht mit dem Anchor-Modeling zu verwechseln. Beide sind zwar in der Domäne Business Intelligence angesiedelt, jedoch ist Anchor-Modeling eine Variante der Gestaltung, die zu Modellen in der sechsten Normalform führt und in der vereinfacht ausgedrückt jedes Attribut in einer eigenen Entität abgebildet ist.

zwischen den fachlichen Entitäten. Dies ist somit eine der wichtigen Eigenschaften des Ansatzes der Ankermodellierung, denn hierdurch bleibt die fachliche Beziehung mit ihrer Kardinalität erhalten und verliert nicht aufgrund der Historisierung ihre Aussagekraft.

An die Ankertabellen können ohne Probleme weitere Tabellen neben der eigentlichen Versionstabelle angehängt werden. Insbesondere ist es in diesen Modellen möglich, Attribute in verschiedene Versionstabellen zu gruppieren und auf diese Art z. B. originäre Attribute, die unverändert von der Quelle übernommen wurden, sowie abgeleitete Attribute, die als Bestandteil der ETL-Prozesse zur Bewirtschaftung des Core Data Warehouse berechnet bzw. bestimmt werden, voneinander zu trennen. Dies ist zugleich ein wichtiges Mittel zur Steuerung verschiedener Service Level oder, allgemeiner formuliert, verschiedener Domänen hinsichtlich Latenz, Change Management und Release Management.

Neben der dargestellten unitemporalen Historisierung im Modell mit Ankertabellen und Versionstabellen ist grundsätzlich auch eine multitemporale Zeitstempelung möglich. Für jede zeitliche Dimension ist dann in der Versionstabelle ein eigener Zeitstempel notwendig. In Abb. 8.16 ist dies für den Fall der bitemporalen Historisierung mit den zeitlichen Perspektiven der fachlichen und technischen Gültigkeit dargestellt.

8.5.3 Bewertung 3NF-Modellierung im Core Data Warehouse

Aufgrund der granularen Struktur kann ein Core Data Warehouse in der 3NF-Form oftmals bereits die Anforderungen an einen Staging-Bereich abdecken. Dazu müssen die unverändert übernommenen Attribute aus den Quellen Bestandteil des Core-Data-Warehouse-Modells sein. In diesen Fällen ist der Staging-Bereich temporärer Art und es besteht keine Notwendigkeit, diesen dauerhaft aufzubewahren.

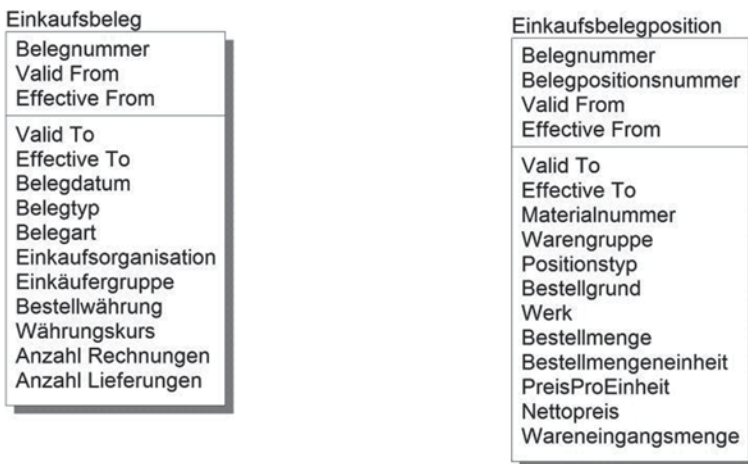


Abb. 8.16 Bitemporale Versionierung im 3NF-Modell

Da in dem 3NF-Modell die granularen Daten gesammelt und integriert sind und für die weitere Distribution bereitstehen, sind damit die wesentlichen Kernaufgaben eines Core Data Warehouse als zentrale Komponente in einer Hub-and-Spoke-Architektur erfüllt.

Da ein 3NF-Modell im Core Data Warehouse die fachlichen Beziehungen repräsentiert und sich nicht an den einzelnen Zyklen der Weiterentwicklung orientiert, haben geänderte Anforderungen im Regelfall, insofern die Art der Harmonisierung oder die extrahierten Daten davon betroffen sind, auch eine Auswirkung auf das Datenmodell. Dadurch entsteht demzufolge ein Aufwand für Anpassung und Test im Core Data Warehouse. Die Art der Änderung die hier bereits erledigt ist, braucht nicht mehr auf dem Weg zu den Data Marts erfolgen, insofern handelt es sich zu einem Teil nur um eine Verlagerung des Aufwands der mit einem Change verbunden ist.

Die Aspekte der Historisierung sind in dem Paradigma der 3NF-Modellierung des Core Data Warehouse ebenso berücksichtigt wie die Anforderungen an deren Reproduzierbarkeit. Auch multitemporale Strukturen sind sehr gut in 3NF-Modellen abbildbar. Dabei hat sich die Trennung der Kernentität von den Versionsentitäten mit den Attributen bewährt.

Aus der Klassifikation der Aufgaben der Datenintegration ergibt sich, dass nur die Aufgaben der Filterung und Bereinigung im ETL-Prozess zwischen Staging-Bereich und Core Data Warehouse liegen. Die Aggregation und Anreicherung findet im Regelfall erst auf dem Weg zum Reporting Layer bzw. zu den Data Marts statt. Der Einsatz einer Ableitungsdatenbank zur Kapselung mehrfach verwendeter Transformationen ist dabei eine sinnvolle Zwischenschicht.

Ein auf Basis der 3NF-Modellierung gestaltetes Core Data Warehouse hat tendenziell einen geringen Grad an Agilität im Umgang mit Change-Prozessen. Auch die damit einhergehenden Aufgaben für Entwicklung und Test weisen eine recht hohe Komplexität auf, sind jedoch geringer als bei dem Einsatz eines Star-Schemas.

8.6 Core Data Warehouse mit Data-Vault

Die bisherigen Ausführungen in den vorherigen Abschnitten haben gezeigt, dass die Gestaltung eines Core Data Warehouse mit der Star-Schema Methodik und der strikten Modellierung konform zur dritten Normalform auch Nachteile mit sich bringt.

Die Grenzen der Star-Schema-Modellierung im Core Data Warehouse sind dabei durch die folgenden Eigenschaften gegeben:

- Die Granularität ist festgelegt und kann sich in einem Modell nicht ändern.
- Die Dimensionierung der Star-Schemata ist festgelegt.
- Die Daten sind redundant abgelegt.
- Near-Realtime-Ladeprozesse können mit Problemen der Synchronisierung von Dimensions- und Faktentabellen einhergehen.

Aber insbesondere der Prozess des Change Managements und der damit verbundene Aufwand für Entwicklung und Test ist ein gewichtiger Grund, um nach Alternativen zu suchen.

Bei der Modellierung in der dritten Normalform hat sich als wesentlicher Kritikpunkt gezeigt, dass die Zeitstempel im Primärschlüssel komplexe Update-Prozesse entlang der Parent-Child-Beziehungen implizieren. Die zur Umgehung notwendigen Ankermodelle mit fachlichen Entitäten und Versionstabellen erhöhen hingegen die Komplexität wiederum.

Beide Ansätze tun sich schwer mit Modelländerungen, etwa aufgrund von erweitertem Scope, der Berücksichtigung neuer Systeme oder neuer Anforderungen, sowie der geforderten Flexibilität im Umgang damit. Darüber hinaus kombinieren beide Modelle Schlüssel, deskriptive Attribute und Beziehungen in den gleichen Tabellen.

Auf dieser Grundlage der Argumentation entstand im Jahr 2000 die Data-Vault-Methode zur Modellierung eines Core Data Warehouse, von dem Erfinder Dan Linstedt nunmehr in (Linstedt 2010) publiziert.⁷ Data Vault ist eine Methode zur Data-Warehouse-Modellierung und stellt damit eine Alternative zu 3NF-Modellen und zum Star-Schema dar. Ausgangslage waren für Linstedt die Annahmen, dass einerseits 3NF-Modelle sich hervorragend für operative Systeme eignen sowie Star-Schema-Modelle Best Practice für relationale Data Marts sind, jedoch andererseits beide nicht optimal auf die Bedürfnisse eines Core Data Warehouse zugeschnitten sind. Entgegen den anderen beiden Ansätzen speichert ein Data-Vault-Modell Schlüssel, deskriptive Attribute und Beziehungsinformationen konsequent in getrennten Tabellen.

Ein Data-Vault-Modell hat die folgenden drei Hauptkomponenten:

- **Hub:** Tabelle mit natürlichen Schlüsseln wie etwa Rechnungsnummer, Mitarbeiternummer, Kundennummer, SKU (Stock Keeping Unit)
- **Satellite:** Tabelle mit beschreibenden Informationen zu einem Hub oder Link, etwa Kundenname, Produktbezeichnung, Belegart, Positionsart etc.
- **Link:** Tabelle zur Abbildung der Beziehung zwischen Hubs und Links (Beziehungen zwischen Links sind erlaubt, vgl. dazu (Linstedt 2010, S. 96 ff.)

8.6.1 Hub-Tabellen

In einem Hub sind keine Beziehungsinformationen und auch keine deskriptiven Attribute (Details) abgebildet. Ein Hub besteht aus dem künstlichem Schlüssel Hub-ID, dem unternehmensweit eindeutigen Businessschlüssel, dem Ladezeitstempel und der Quelle.

Der Businessschlüssel in einem Hub ist nach Möglichkeit unternehmensweit ausgerichtet. Im Beispiel in Abb. 8.17 ist der Businessschlüssel die Kundennummer. Im Hub erfolgt das Mapping auf die künstliche Hub-Id mit der weiteren Angabe, wann und aus

⁷ Eine sehr umfassende Darstellung der Data-Vault-Methode gibt (Hultgren 2012).

Abb. 8.17 Hub-Tabelle
Kunde

The diagram shows a green rectangular box representing a table structure. It is titled 'HUB Kunde' in a dark green header. Below the header, the table is divided into four rows, each with a single column: 'HUB ID', 'Kundennummer', 'LOADDATE', and 'SOURCE'.

HUB Kunde			
HUB ID			
Kundennummer			
LOADDATE			
SOURCE			

welchem Quellsystem dieser fachliche Schlüssel das erste Mal ins Data Warehouse geladen wurde. Im Folgenden sollen Hub-Entitäten dunkel hinterlegt dargestellt werden.

Eine mögliche Ausprägung zu diesem Hub des Beispiels ist in Abb. 8.18 dargestellt. In der Tabelle repräsentiert das Ladedatum bzw. der Zeitstempel (hier ein Tagesdatum) den Zeitpunkt, wann dieser Businessschlüssel (die Kundennummer) erstmals in das Data Warehouse geladen wurde. Die Source beschreibt, aus welchem System der Businessschlüssel entstammt.

Auf die Harmonisierung über mehrere Quellsysteme hinweg wird später in diesem Abschnitt noch eingegangen. Der Hub definiert somit die Liste der eindeutigen Business-Keys und deren Zuordnung (Mapping) zu künstlichen Schlüsseln (sogenannten Surrogate-Keys, die daher auch oft mit SID für Surrogate Identifier bezeichnet sind). Daher darf ein Hub auch nicht erneut aufgebaut werden und einmal eingefügte Sätze bleiben in dieser Tabelle und werden nicht gelöscht. Ausnahmen können allerdings dann gegeben sein, wenn beispielsweise in ein B2B-Kundenobjekt versehentlich B2C-Kundendaten geladen wurden. Dies ist eine erlaubte Korrektur, wobei aber abhängige Daten in den anderen Tabellen des Modells ebenfalls zu korrigieren sind.

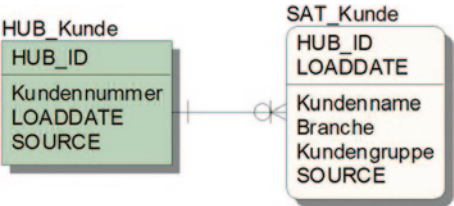
8.6.2 Satellite-Tabellen

Die Satellite-Tabellen umfassen alle beschreibenden Attribute aller Quellsysteme und bilden die vollständige Historie im Data Warehouse ab. Eine Satellite-Entität besteht aus dem Fremdschlüssel zum Hub, dem Ladezeitstempel, den Attributen und der Datenquelle (Source).

HUB ID	Kundennummer	LOADDATE	SOURCE
1	1003456	12.04.2013	ECC Europe
2	1003457	12.04.2013	ECC Europe
3	1003458	13.04.2013	ECC Europe
4	1003459	13.04.2013	ECC Europe

Abb. 8.18 Ausprägungen im Hub Kunde

Abb. 8.19 Satellite-Tabelle



Ein Hub kann dabei mehrere Satellites haben. Die Satellites werden z. B. nach der Art der Daten, der Änderungshäufigkeit und der Quelle aufgeteilt. Passend zum Hub Kunde ist in Abb. 8.19 eine Beziehung zu einer Satellite-Tabelle (hell hinterlegt) dargestellt.

Formal modellseitig besteht zwischen diesen beiden Tabellen eine 1:n-Beziehung, die aber nicht in das Datenbankmodell im Sinne der referentiellen Integrität generiert werden muss. Demzufolge brauchen in dem Datenbankmodell die Fremdschlüsselbeziehungen nicht implementiert werden, die Konsistenzprüfung wird dann seitens des ETL-Prozesses abgefangen.

Während es in einem Hub nur eine Zeile für jede Kundennummer, also den Business-schlüssel, gibt, ist dies in einem Satellite nicht so. Mit jedem Ladelauf ist zu prüfen, ob sich an den Attributen im Satellite etwas geändert hat und wenn dies der Fall ist, erfolgt die Einfügung eines neuen Satzes in dem Satellite mit dem Zeitstempel des Ladens. Somit ergeben sich, wie in Abb. 8.20 ersichtlich, zu einer HUB-ID, also einer Kundennummer, mehrere Zeilen im Satellite. In dem Beispiel ist der Kundenname des Kunden mit der Nummer 1003459 von Baby Waltz auf Baby Walz geändert worden. Diese Änderung führt zu dem neuen Satz mit Ladedatum (LOADDATE) 14.04.2013 in Abb. 8.20.

Da diese Daten für die Weitergabe an die nachfolgenden Schichten normalerweise zusammen zu übergeben sind, müssen die Hub-Tabellen mit den angehängten Satellite-Tabellen verknüpft werden.

The diagram shows two tables. The top table is the Hub table with columns HUB ID, Kundennummer, LOADDATE, and SOURCE. It contains four rows. The bottom table is the Satellite table with columns HUB ID, LOADDATE, Kundenname, ..., and SOURCE. It contains five rows. A purple arrow points from the fourth row of the Hub table (HUB ID 4, Kundennummer 1003459) to the fourth and fifth rows of the Satellite table, which both have HUB ID 4, indicating that multiple satellite records can exist for a single hub record.

HUB ID	Kundennummer	LOADDATE	SOURCE
1	1003456	12.04.2013	ECC Europe
2	1003457	12.04.2013	ECC Europe
3	1003458	13.04.2013	ECC Europe
4	1003459	13.04.2013	ECC Europe

HUB ID	LOADDATE	Kundenname	...	SOURCE
1	12.04.2013	Rofu	...	ECC Europe
2	12.04.2013	Baby One	...	ECC Europe
3	13.04.2013	Kinderland	...	ECC Europe
4	13.04.2013	Baby Waltz	...	ECC Europe
4	14.04.2013	Baby Walz	...	ECC Europe

Abb. 8.20 Ausprägungen in der Satellite-Tabelle

Die Abfragen an diese Modelle sind recht komplex, denn es muss für jeden Schlüssel aus dem Hub jeweils in einem Subselect das maximale Ladedatum im Satellite ermittelt werden. Gibt es zu einem Schlüssel im Hub gar keinen Eintrag in dem Satellite, wird dieser Satz nur durch einen Outer-Join in der Ergebnismenge berücksichtigt (Hahne 2014). Hier haben sich Verfahren bewährt, die einerseits jeweils eine vollständige Abbildung einer Version in allen beteiligten Tabellen gewährleisten und andererseits auch durch zusätzliche Versionen eine vollständige Historie darstellen.

Die Änderungen von Attributen in der Satellite-Tabelle werden über neue Zeilen abgebildet. Um viele Duplizierungen von nicht geänderten Attributen zu vermeiden (z.B. nur der Name ändert sich, andere Attribute bleiben), ist die Trennung in mehrere Satellite-Tabellen sinnvoll. In Abb. 8.21 wird das Beispiel des Kunden erweitert um die Tabellen für die Attribute Branche und Kundengruppe in einem Satellite und Adressattribute in einer weiteren Satellite-Tabelle.

Offenbar ist die Art der Aufteilung der Attribute auf verschiedene Satellites ein Freiheitsgrad der Modellierung, der aber die Aspekte der Bewirtschaftung genauso adressiert wie Aspekte der Datenredundanz. Dem Ansatz des Anchor-Modeling (Rönnebeck et al. 2010) folgend müsste sogar jedes Attribut in einer eigenen Tabelle gespeichert werden. In der Praxis ist es ein Abwägen basierend auf Eigenschaften der Änderungshäufigkeit, dem Quellsystem und fachlichen Kombinationen von Attributen.

Wenn die Attribute auf mehrere Satellites verteilt sind, können diese in der Konsequenz auch unterschiedlichen Änderungszyklen unterworfen sein. So entstehen etwa Einträge in den Satellites mit teilweise unterschiedlichen Zeitstempeln (LOADDATE).

In Abfragen mit mehreren Satellites an einem Hub sind die Zeitstempel pro Hub einzeln zu ermitteln. Somit ist dann pro Hub eine Subselect-Klausel notwendig, um die korrekten Zeitstempel zu ermitteln. Dies kann durch den Einsatz sogenannter Point-In-Time-Tabellen seitens der Abfrage vereinfacht werden (Hahne 2014).

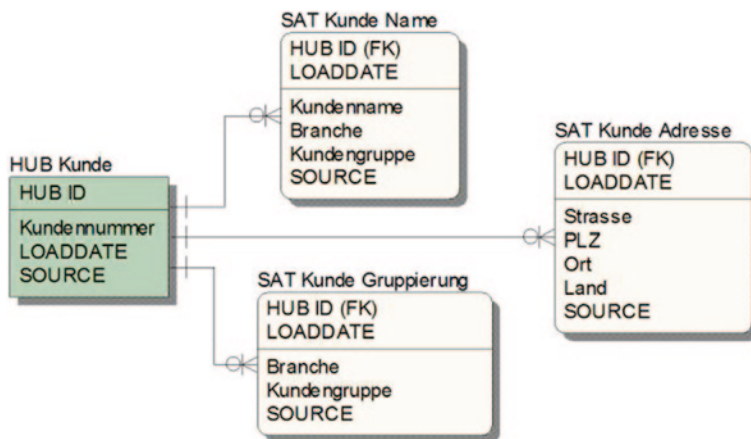


Abb. 8.21 Hub mit mehreren Satellite-Tabellen

Der Einsatz von Point-In-Time-Tabellen hat einerseits den Vorteil einer etwas besseren Performance bei Abfragen, da nur noch ein Subselect benötigt wird. Andererseits muss diese Tabelle aber auch im Rahmen der Bewirtschaftung eines Data-Vault-Modells mit gefüllt werden. Daraus können sich potenziell Inkonsistenzen ergeben, zumindest aber gibt es eine weitere Abhängigkeit für eine konsistente Abfrage an das Modell.⁸

8.6.3 Link-Tabellen

Zu einem Data-Vault-Modell gehören neben den bisher dargestellten Hub- und Satellite-Tabellen auch noch die Link-Tabellen über die Beziehungen zwischen Hub-Tabellen abgebildet werden. Grundprinzip bei Data Vaults ist dabei, dass ausschließlich many-many-Beziehungen berücksichtigt werden.⁹ Dies erfolgt auch dann, wenn es sich um eine fachliche 1:n-Beziehung handelt. Dies hat den Vorteil, dass eine Änderung der Kardinalität einer Beziehung nicht zu kaskadierenden Modelländerungen führt. Dies ist auch konform zu der Aufteilung von Attributen, Schlüsseln und Beziehungen auf verschiedene Tabellen. Dies wird am Beispiel der Kundenaufträge im Data-Vault-Modell in Abb. 8.22 deutlich.

Ein Link ist eine Tabelle zur Abbildung einer Beziehung vom Grad größer oder gleich zwei wobei alle beteiligten Tabellen zum Link in einer 1:n Beziehung stehen. Die Beziehung zwischen den beteiligten Hub-Tabellen ist damit eine many-many-Beziehung.

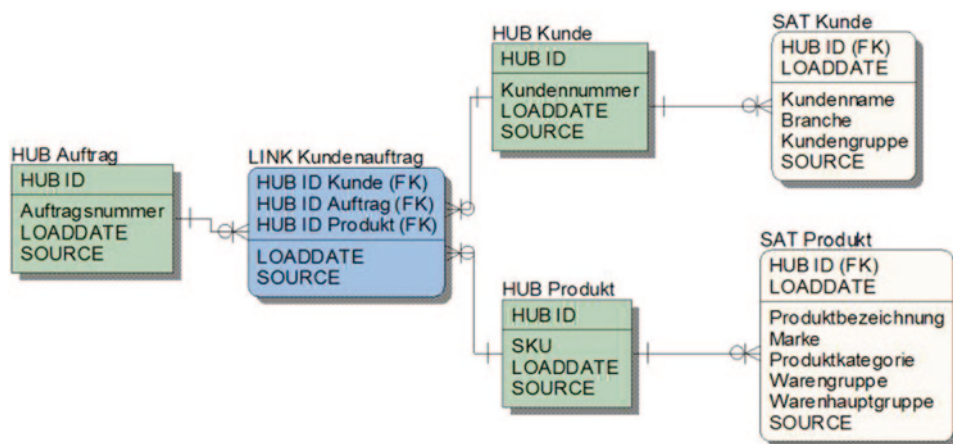


Abb. 8.22 Link-Tabelle im Data Vault

⁸ Dies gilt sinngemäß auch für die so genannten *Bridge-Tabellen* im Data Vault, die ebenfalls zur Verbesserung der Performance Verknüpfungen von Link-Tabellen mit weiteren Tabellen persistieren.

⁹ Eine detailliert Darstellung verschiedener Formen von Link-Tabellen, u. a. auch für rekursive Beziehungen, ist bei (Hultgren 2012, S. 93 ff.) aufgeführt.

Der Link Kundenauftrag in Abb 8.22 hat den zusammen gesetzten Primärschlüssel, der aus den Fremdschlüsseln zu den in der Beziehung beteiligten Hubs Kunde, Produkt und Auftrag besteht. Mit dem ersten Auftreten der Beziehung in den Quelldaten wird diese Kombination von Schlüsseln in die Link-Tabelle geschrieben. Dies geschieht zusammen mit dem Zeitstempel des Auftretens und der Quelle, aus der diese Information stammt. Eine explizite Historisierung von dieser Beziehung auf Basis einer aus dem Quellsystem übermittelten Gültigkeit der Beziehung ist dabei nicht vorgesehen.

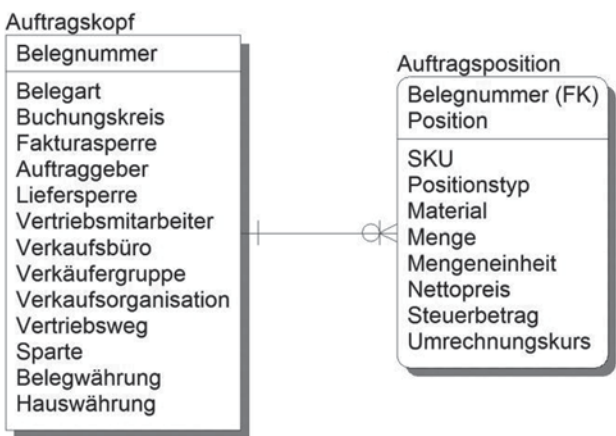
Der Primärschlüssel der Link-Tabelle ist demzufolge ein zusammengesetzter Schlüssel der aus den Fremdschlüsseln zu den Hub-Tabellen besteht. Sollen auch Attribute in Form von Satelliten an diese Beziehung gebunden werden, ist dazu ein künstlicher Primärschlüssel notwendig. Hierzu dient eine Link-ID.

In der Logik eines Data-Vault-Modells sind für dessen Gestaltung zunächst die Businesssschlüssel zu identifizieren, damit hieraus die Hub-Tabellen gebildet werden können. Aus den Beziehungen zwischen Businesssschlüsseln ergeben sich dann die Link-Tabellen. Alle weiteren Attribute finden Eingang in die Satellite-Tabellen. Dieser Prozess soll anhand des Staging-Modells aus Abb. 8.23 verdeutlicht werden.

In dem ER-Modell der Auftragsdaten in Abb. 8.23 gibt es eine Master-Detail-Beziehung zwischen dem Auftragskopf und den Auftragspositionen, die in der gewohnten Form als schwache Entität ausgeprägt ist, also existenziell abhängig ist vom Auftragskopf. Dies drückt sich auch dadurch aus, dass das Primärattribut in der Entität Auftragskopf in den Schlüssel der Position migriert und die Belegnummer als Primärschlüssel des Auftragskopfs auch Bestandteil des Schlüssels der Auftragsposition ist.

Da die Belegnummer ein fachlicher wichtiger Schlüssel ist, ergibt sich daraus bereits ein erster Hub. Ebenso formt auch die Position eines Auftrags einen Businesssschlüssel, der Schlüssel ist aber der zusammengesetzte Schlüssel aus der Belegnummer und der Position (vgl. Hahne 2014). Generell führen Transaktionen aus den operativen Prozessen immer zu Hubs (s. Hultgren 2012, S. 294). In der Data-Vault-Modellierung wird diese Ab-

Abb. 8.23 Staging-Modell Auftragsdaten



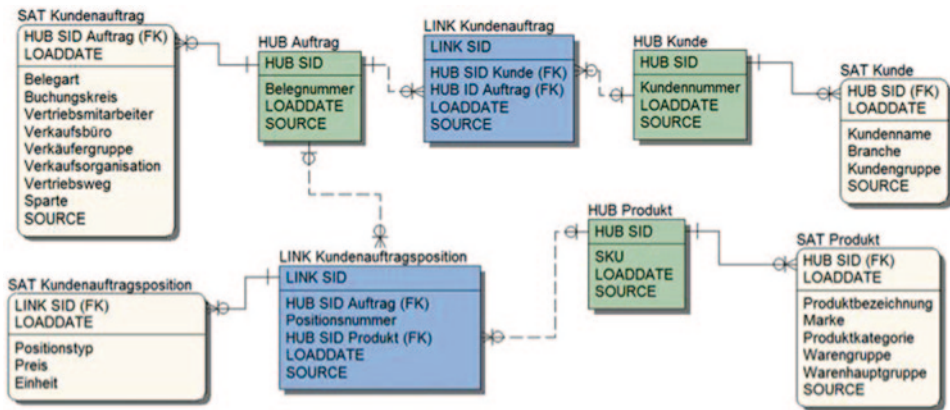


Abb. 8.24 Variante mit schwachen Hubs und degenerierten Feldern

hängigkeit, die sich aus der schwachen Entität für die Position ergibt, in Abhängigkeit von der Modellierungsphilosophie gerne vermieden. Daher ist in dem Modell in Abb. 8.24 die Position als so genannter *schwacher Hub* ausgeprägt und die Positionsnummer findet sich als *degeneriertes Feld* in dem Link wieder.

Weitere Businessschlüssel in dem Modell sind der Kunde (Auftraggeber) und das Produkt (SKU). Darüber hinaus sind noch weitere Felder potenziell als fachliche Schlüssel einzustufen, wobei sich dies ggf. erst später im Ablauf eines Projektes ergibt.

Zu den identifizierten Businessschlüsseln des Beispiels gibt es einige sich direkt aus dem Modell ergebende Beziehungen die in entsprechende Link-Tabellen münden. Unmittelbar ersichtlich ist die Beziehung zwischen dem Auftragskopf und seinen Positionen. Die passende Link-Tabelle ist in Abb. 8.24 als Link Kundenauftragsposition benannt und beinhaltet als zusammengesetzter Schlüssel die beiden Schlüssel der Hubs Auftrag und Auftragsposition, letzteren als degenerierten Hub.

Die beiden anderen Link-Tabellen in dem Beispiel verdeutlichen die Variante, wenn zusätzlich Satellite-Tabellen an den Link angebunden sein sollen, es also Attribute der Beziehung gibt. Im konkreten Beispiel ist dies die Beziehung zwischen dem Kunden und dem Auftrag, die Link-Tabelle dazu hat den Namen Link Kundenauftrag. An dieser Tabelle ist zu erkennen, dass nunmehr der Primärschlüssel ein künstlicher ist (LINK SID).¹⁰ Die Fremdschlüssel wandern in den abhängigen Teil, stellen aber gemeinsam einen alternativen Schlüssel der Tabelle dar. Dieser Schlüssel LINK SID dient nun in dem angehängten Satelliten zur Aufnahme der Attribute des Auftragskopfs. Streng genommen ist es genauso möglich, in diesem Fall diesen Satellite auch an den Hub Auftrag zu hängen, da die Attribute wie z. B. Verkaufsbüro oder Belegart bereits eindeutig für den Auftragskopf sind. Es gibt aber durchaus viele Beziehungen, deren Attribute nur in dem Zusammenspiel der beteiligten Hub-Tabellen eindeutig sind und daher zwingend diese Vorgehensweise benötigen.

¹⁰ SID steht hier für Surrogate Identifier, ein künstlicher Schlüssel.

8.6.4 Zeitstempel im Data Vault

Die bisher genutzten Beispiele basieren alle auf einer einfachen Angabe des Ladezeitpunkts als Zeitstempel in den Tabellen. Dies hat den Vorteil, dass die Operationen zur Bewirtschaftung eines Data-Vault-Modells reine Insert-Prozesse sind. Es werden also jeweils nur neue Sätze geschrieben, jedoch keine existierenden Sätze gelöscht oder aktualisiert. Ausnahmen davon stellen lediglich kontrollierte Korrekturen dar, die z. B. aufgrund von korrupten Datenlieferungen entstanden und inhaltlich falsch sind. Der Nachteil dieser Vorgehensweise ist bei den Skripten für Abfragen an ein Data-Vault-Modell deutlich geworden, es sind dabei komplexe Sub-Select-Statements notwendig, um die passenden Zeitstempel in den Tabellen zu filtern.

Mit dem Zeitstempelverfahren können Abfragen die korrekte Version anhand der Attribute Gültig Von und Gültig Bis bereits mit einem Between-Statement ermitteln. Dieses Verfahren kann auf ein Data-Vault-Modell angewendet werden und führt zu Strukturen in den Satelliten-Tabellen wie in Abb. 8.25 dargestellt. Die Gültigkeitsfelder sind dabei als Loaddate Begin und Loaddate End aufgeführt. Dies verdeutlicht auch, dass hierfür im Wesentlichen die Zeit der Verarbeitung im Data Warehouse, also die technische Gültigkeit, herangezogen wird.

In der Literatur findet sich häufig die Empfehlung, den Zeitstempel der das Ende der Gültigkeit festlegt zunächst auf NULL zu setzen (s. Linstedt 2010, S. 49 f.). Die Erfahrungen in der Zeitstempelung im Star-Schema sprechen eher dafür, hier ein technisches Maximaldatum zu wählen, damit Abfragen direkt mit der Between-Klausel arbeiten können.

In der Tabelle ist wieder das Attribut Kundenname neben weiteren aufgeführt. Die Auswirkungen einer Änderung des Kundennamens im Modell sind in Abb. 8.25 grau hervorgehoben. Es gibt nun nicht nur einen neuen Satz, der die Information trägt, ab wann die neue Ausprägung im Data Warehouse bekannt geworden ist, sondern zusätzlich muss ein Datensatz in Form eines Updates auf diesen Umstand angepasst werden, in dem der schließende Zeitstempel (Loaddate End) aktualisiert wird. Im Beispiel wird das Ende des alten Satzes terminiert auf den Vortag der Beladung (14.04.2013) und der neue Satz wird mit dem aktuellen Ladetag (15.04.2013) als Zeitstempel eingefügt.

Eine bitemporale Historisierung kennt das Data-Vault-Modell in dieser Form nicht direkt, kann jedoch diesbezüglich angepasst werden, etwa durch eine Erweiterung des Schlüssels um einen Zeitstempel der fachlichen Gültigkeit. Bekannt ist auch das Konzept

HUB ID	LOADDATE Begin	LOADDATE End	Kundenname	...	SOURCE
1	12.04.2013	NULL	Rofu	...	ECC Europe
2	12.04.2013	NULL	Baby One	...	ECC Europe
3	13.04.2013	NULL	Kinderland	...	ECC Europe
4	13.04.2013	14.04.2013	Baby Waltz	...	ECC Europe
4	15.04.2013	NULL	Baby Walz	...	ECC Europe

Abb. 8.25 Gültig von und bis Zeitstempelung im Satellite

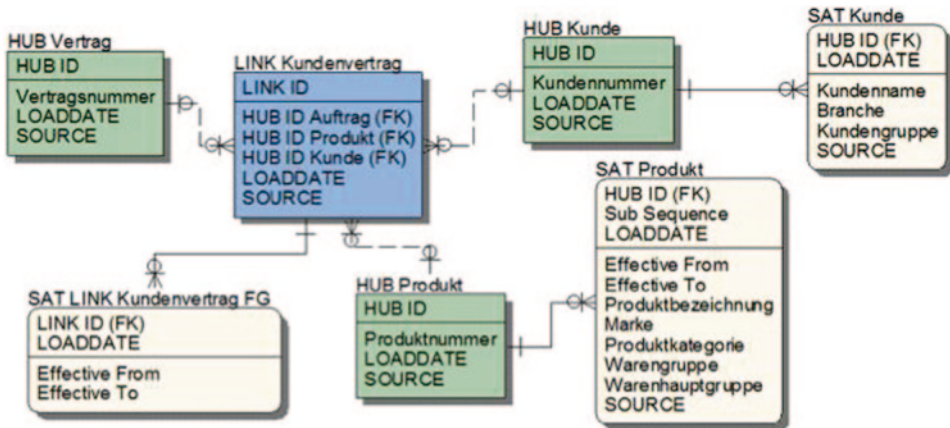


Abb. 8.26 Bitemporale Historisierung im Data Vault

der *multi-active-satellites* wie in Abb. 8.26 für den Produkt-Satelliten dargestellt. Dadurch können zu einem gegebenen technischen Gültigkeitszeitpunkt auch mehrere fachliche Versionen gültig sein. Dies ist für eine vollständige Reproduzierbarkeit der Historie notwendig.

Die Historisierung von Beziehungen (Link Kundenauftrag) erfolgt durch Satelliten, in denen die Zeitstempel der fachlichen Gültigkeit abgelegt sind. Diese sind die so genannten *effective-satellites*.

8.6.5 Harmonisierung von fachlichen Schlüsseln

Die Hub-Tabellen bilden die wesentlichen konstituierenden Bausteine eines Data-Vault-Modells und ergeben sich aus den fachlichen Schlüsseln. Alle in operativen Systemen genutzten Businesschlüssel sind dafür geeignete Kandidaten. Vielfach ist es jedoch so, dass ein fachlicher Schlüssel wie z. B. eine Kundennummer sich nicht nur aus einem Quellsystem ergibt, sondern in mehreren Quellsystemen entsteht. Sind dabei die auftretenden Nummernkreise disjunkt, können diese problemlos in den gleichen Hub gehen, da dort bereits die Zuordnung zum Quellsystem grundsätzlich verankert ist. Diese ist aber nicht Bestandteil des Schlüssels. Es handelt sich somit nicht um einen zusammengesetzten fachlichen Schlüssel bestehend etwa aus dem Quellsystem und der Kundennummer (siehe Abb. 8.27).

In einem Hub können daher überlappende Nummernkreise nicht direkt abgebildet werden. Im Konzept der Data-Vault-Modellierung dient dazu der sogenannte Business Vault, ein abgeleitetes Data-Vault-Modell, in dem diese Schritte der fachlichen Harmonisierung verschiedenster auch überlappender Schlüsselbereiche aus verschiedenen operativen Systemen erfolgt (s. Linstedt 2010, S. 40 ff.).

Abb. 8.27 Harmonisierung von Businessschlüsseln

HUB ID	Kundennummer	LOADDATE	SOURCE
1	1003456	12.04.2013	ECC Europe
2	2004132	12.04.2013	ECC Asia
3	2004133	12.04.2013	ECC Asia
4	1003458	13.04.2013	ECC Europe
5	1003459	13.04.2013	ECC Europe
6	2004134	13.04.2013	ECC Asia
7	2004135	13.04.2013	ECC Asia

8.6.6 Agilität in Data-Vault-Modellen

Eine wesentliche Eigenschaft eines Data-Vault-Modells ist dessen Stabilität in Bezug auf Änderung. Dies postulierte Dan Linstedt bereits als notwendige Bedingung während der Entstehung der Methode und soll hier an dem obigen Beispiel der Einkaufsbelege verdeutlicht werden.

Die in dem Staging-Modell identifizierten fachlichen Schlüssel führen zu einem Data-Vault-Modell mit den Hub-Tabellen Lieferant, Material, Einkaufsbeleg und dem degenerierten Hub Einkaufsbelegposition. Für die Beziehung zwischen den Kopfdaten und den Positionsdaten kommt in dem Modell wieder ein Link mit einem künstlichen Schlüssel (LINK_SID) zum Tragen. Dies ist hier allerdings optional, da es keine angehängten Satelliten zu der Link-Tabelle gibt. Die Attribute der Position sowie des Belegkopfs sind jeweils über eigene Satelliten abgebildet.

Interessant ist die dreifache Beziehung zwischen dem Beleg und dem Lieferanten, denn dabei handelt es sich ja um die drei Rollen Warenlieferant, Lieferant und Rechnungssteller. Für jede der drei Rollen ist in dem Modell eine eigene Link-Tabelle modelliert. Dies verdeutlicht damit auch die fachliche Beziehung. Eine alternative Herangehensweise ist der Aufbau einer Beziehung zwischen dem Beleg, dem Lieferanten und einem Hub für die Rolle. Diese ist allerdings nicht wirklich ein Businessschlüssel. Die dargestellte Form der Abbildung ist daher zu bevorzugen.

Kommt es nun zu Änderungen des Staging-Modells, sind die Auswirkungen auf das Modell des Core Data Warehouse zu betrachten. In einem Data Vault gilt das Grundprinzip, dass einmal erstellte Tabellen sich nicht mehr ändern. Änderungen werden nur durch neue Tabellen aufgefangen. Dadurch ist gewährleistet, dass bereits in Produktion befindliche Tabellen und deren Bewirtschaftung in Form von ETL-Prozessen auch weiterhin Bestand haben und auch gar nicht erneut getestet werden müssen, da diese von den Änderungen nicht betroffen sind.

Für die Änderungen bedeutet dies, dass diese im Data-Vault-Modell neue Tabellen für Hubs, Satelliten und Links implizieren können. Die bestehenden Komponenten bleiben davon aber unberührt, denn nur wenn alle bisherigen Modellbestandteile vollständig erhalten bleiben, kann von einer Unanfälligkeit gegenüber Änderungen gesprochen werden. Im Sinne eines Reengineering ist es jedoch durchaus möglich, gelegentlich das Modell

aufzuräumen und hinsichtlich des Schnitts der Satelliten etwa zu optimieren. Dies sollte aber nur sehr dosiert erwogen werden.

Für den Aufbau eines Core Data Warehouse auf Basis der Data-Vault-Methode wird von Dan Linstedt die folgende Vorgehensweise empfohlen (s. Linstedt 2010):

- Die Art der Zeitstempelung ist festzulegen.
- Modellierung der Hubs: Businesssschlüssels und deren Nutzung
- Modellierung der Links: Beziehungen zwischen den Businesssschlüsseln basierend auf den Geschäftsprozessen
- Modellierung der Satelliten: Kontext wird zu den Hubs und deren Beziehungen hinzugefügt, differenziert nach Änderungshäufigkeit, Quelle und Art der Information
- Modellierung von Hilfstabellen (z. B. Point-In-Time-Tabellen)

Aufgrund der klaren Festlegungen, wie die Tabellen zu gestalten sind und in welchen Beziehungen diese stehen können, ergeben sich einige Integritätsbedingungen für Data-Vault-Modelle (siehe dazu auch Hahne 2014).

8.6.7 Bewertung der Data-Vault-Methode

Nach den bisherigen Ausführungen wird deutlich, dass die Data-Vault-Methode sehr gut geeignet ist, um ein Core Data Warehouse zu modellieren. Mit einem Fokus auf die schnelle Beladung und nahezu keinen Änderungsbedarf im Falle einer Änderung auf Quellsystemseite liegt die Stärke des Ansatzes einerseits darin, Quellsysteme mit hoher Ladegeschwindigkeit anzubinden. Andererseits soll dies mit relativ geringem Aufwand u. a. durch Mechanismen der Generierung von ETL-Prozessen geschehen.

Diese Vorteile führen aber auch zu Strukturen, durch die Abfragen an Data-Vault-Modelle komplex werden und daher nicht für den Endanwender geeignet sind. Ein Core Data Warehouse auf Basis von Data-Vault-Modellen sollte daher nicht direkt für Abfragen und Analysen genutzt werden.

Ein Data-Vault-Modell im Core Data Warehouse deckt die Anforderungen an einen Staging-Bereich im Regelfall mit ab; daher ist dieser nur temporärer Art und bedarf keiner besonderen Aufbewahrung. Da die Daten im Sinne eines SPOT auch gesammelt, integriert und dann weiter verteilt werden, sind die Aufgaben des Core Data Warehouse selbst auch erfüllt.

Im Falle der Änderungen der Quellsysteme oder der fachlichen Anforderungen wird das Core-Data-Warehouse-Modell nur erweitert. Dadurch ist der Aufwand für Test und Anpassung sehr gering. Jedoch sind dafür die Anpassungen in den nachgelagerten Schichten etwa im Reporting Layer gegeben und auch größer als beispielsweise in der Situation eines Star-Schema-Modells als zentrales Paradigma für das Core Data Warehouse. Der Aufwand wird also teilweise nachgelagert trotzdem anfallen.

Aspekte der Historisierung auf Basis der technischen Kenntniserlangung im Data Warehouse sind umfänglich gegeben. Für eine echte fachliche Historisierung, die auch Bestandteil der gelieferten Daten von den Quellsystemen sein kann, ist das Modell geringfügig zu verändern, um auch bitemporale Anforderungen abzudecken.

Konform zur Klassifikation der Aufgaben der Datenintegration sind nur die Phasen Filterung und Bereinigung im ETL-Prozess zwischen Staging und Core Warehouse Modell.

Ein auf Basis der Data-Vault-Methode gestaltetes Core Data Warehouse hat tendenziell einen hohen Grad an Agilität im Umgang mit Change-Prozessen. Auch die damit einhergehenden Aufgaben für Entwicklung und Test weisen eine tendenziell geringe Komplexität auf, haben jedoch die Notwendigkeit nachgelagerter Anpassungen in Richtung Reporting Layer mit den Star-Schema-Modellen.

8.7 Zusammenfassung

Heutige Data-Warehouse-Architekturen sind stark mit neuen Anforderungen hinsichtlich der Kostenstrukturen für den Aufbau und den Betrieb von BI-Systemen sowie hinsichtlich der sich zunehmend schneller ändernden geschäftlichen Anforderungen, den sogenannte Business Requirements, konfrontiert. Gefordert sind Architekturen und Konzepte, die die Heterogenität an Strukturen, Daten und Anforderungen beherrschbar machen und die eine Reduktion der Time-to-Market von BI-Applikationen ermöglichen.

Die Lösung für eine zukunftsorientierte Architektur für die dispositive Informationsversorgung stellen mehrschichtige unternehmensweit ausgerichtete Data-Warehouse-Architekturen dem Hub-and-Spoke-Ansatz folgend dar. Diese sind aufgrund der vorzufindenden heterogenen Art und Komplexität der Aufgaben zur Harmonisierung und Transformation jeweils unternehmensindividuell ausgeprägt.

Durch die zunehmende Globalisierung von Unternehmen entstehen auch immer globalere Data Warehouse Lösungen, in denen eine stringente Trennung in Online-Zeit und Batch-Zeit bereichsübergreifend nach herkömmlichen Mustern nicht mehr anzutreffen ist. Die Prozesse der Datenintegration im Data Warehouse sind vermehrt im so genannten 24x7 Modus zu betreiben und bedürfen einer Entzerrung für unterschiedliche Service Level und Lademuster.

Die unterschiedlichen Optionen zur Ausgestaltung von Business-Intelligence-Architekturen dem Paradigma von Inmon oder Kimball folgend sowie auf Basis der Data-Vault-Methodik haben dabei unterschiedliche Stärken und Schwächen. Dennoch haben sich aber einige Best Practices ergeben. So ist etwa ein Trend hin zu weniger redundanter Datenhaltung unter anderem durch Techniken der Virtualisierung klar erkennbar. Diese ist bei Data-Vault-Architekturen sogar explizit vorgesehen, da dort die Modelle zur Implementierung der Business Rules nicht notwendigerweise persistent ausgeprägt sein müssen. Dies fördert auch eine bessere Unterstützung von Self-Service-BI.

Eine weitere Entwicklung besteht in der Trennung und separaten Handhabung von Geschäftsregeln verschiedener Bereiche auch mit unterschiedlichen Service Levels in getrennten Modellteilen oder Architekturkomponenten und dies auch tendenziell näher am Benutzer, also in Schichten oder Komponenten näher am Business als am Core Data Warehouse.

Literatur

- Abelló, A., Martín, C.: A bitemporal storage structure for a corporate data warehouse, Technical Paper, Universitat Politècnica de Catalunya (2003)
- Adamson, C.: Star Schema – The Complete Reference. Wiley Publishing, New York (2011)
- Ariyachandra, T., Watson, H.J.: Which data warehouse architecture is most successful? *Bus. Int. J.* **11**(1), 4–6 (2006)
- Gluchowski, P.; Gabriel, R.; Dittmar, C.: Management Support Systeme und Business Intelligence. 2. Auflage. Springer-Verlag, Berlin, Heidelberg (2008)
- Göhl, R., Hahne, M.: Bessere Architektur, Organisation und Methodik, *BI-Spektrum*, 6. Jg., Heft 2, S. 10–15 (2011)
- Hahne, M.: SAP Business Information Warehouse – Mehrdimensionale Datenmodellierung. Springer-Verlag, Berlin, Heidelberg (2005)
- Hahne, M.: Lifecycle management – Neue Architektur für das data warehouse. In: Gluchowski, P., et al. (Hrsg.) *Schlaglichter der Wirtschaftsinformatik*, GUC-Verlag, Chemnitz 2007, S. 157–166 (2007)
- Hahne, M.: Architektur und Modellierung für skalierbare flexible Data Warehouse Systeme. In: Felden, C., Krebs, S., Stock, S. (Hrsg.) *Perspektiven der Business Intelligence – Festschrift anlässlich des 60. Geburtstages von Peter Chameni*, München (2011)
- Hahne, M.: Modellierung von Business-Intelligence-Systemen – Leitfaden für erfolgreiche Projekte auf Basis flexibler Data-Warehouse-Architekturen, Edition TDWI. dpunkt, Heidelberg (2014)
- Hahne, M., Böttiger, W.: Push with Enterprise Data Warehousing – Improving performance and Stability, *BI-Spektrum*, 1. Jg., Heft 1, S. 34–36 (2006)
- Haupt, J., Hahne, M.: Next generation enterprise data warehousing, E-3, S. 82–84 (Mai 2007)
- Hughes, R.: Agile Data Warehousing – Delivering World-Class Business Intelligence Systems Using Scrum and XP. Morgan Kaufmann, Bloomington (2008)
- Hultgren, H.: Modeling the Agile Data Warehouse with Data Vault. New Hamilton, Stockholm (2012)
- Imhoff, C., Gallema, N., Geiger, J.G.: Mastering Data Warehouse Design – Relational and Dimensional Techniques. John Wiley & Sons, Indianapolis (2003)
- Inmon, W.H.: The Problem with Dimensional Modeling, *DM Review* (May 2000)
- Inmon, W.H., Imhoff, C., Sousa, R.: Corporate Information Factory, Wiley, 2nd ed. New York (2001)
- Inmon, W.H., Strauss, D., Neishloss, G.: DW 2.0 – The Architecture for the Next Generation of Data Warehousing. Morgan Kaufmann, Amsterdam (2008)
- Johnston, T., Weis, R.: Managing Time in Relational Databases. Morgan Kaufmann, Burlington (2010)
- Kemper, G., Finger, R.: Transformation operativer Daten – Konzeptionelle Überlegungen zur Filterung, Harmonisierung, Aggregation und Anreicherung im Data Warehouse. In: Chameni, P., Gluchowski, P. (Hrsg.) *Analytische Informationssysteme – Data Warehouse, On-Line Analytical Processing, Data Mining*, 4. Aufl., Springer, S. 159–174. Berlin (2010)
- Kemper, H.-G., Baars, H., Mehanna, W.: Business Intelligence – Grundlagen und praktische Anwendungen. 4. Aufl. Vieweg, Wiesbaden (2014)

- Kimball, R., Ross, M.: *The Data Warehouse Toolkit – The Complete Guide to Dimensional Modeling*, 2. Aufl. Wiley, John Wiley & Sons, New York (2002)
- Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B.: *The Data Warehouse Lifecycle Toolkit – Practical Techniques for Building Data Warehouse and Business Intelligence Systems*, John Wiley & Sons, 2. Aufl. New York (2008)
- Linstedt, D.: *Super Charge Your Data Warehouse*. Dan Linstedt-Verlag (2010)
- Rönnebeck, L., Regardt, O., Bergholtz, M., Johannesson, P., Wohed, P.: *Anchor Modelling – Agile Information Modeling in Evolving Data Environments*, Stockholm University (2010)
- Sinz, E.J., Ulbrich-vom-Ende, A.: *Architektur von Data-Warehouse-Systemen*. In: Chamoni, P., Gluchowski, P. (Hrsg.) *Analytische Informationssysteme*, 4. Aufl., S. 175–196. Springer, Heidelberg (2010)
- Snodgrass, R.T.: *Developing Time-Oriented Database Applications in SQL*. Morgan Kaufmann, San Francisco (2000)

Tobias Knabke und Sebastian Olbrich

Inhaltsverzeichnis

9.1	Einleitung und Motivation	188
9.2	Grundlagen In-Memory-Datenbanken	189
9.2.1	Aktuelle Entwicklungen im Hauptspeicherbereich	189
9.2.2	Datenorganisation	191
9.3	Konsequenzen für Business Intelligence und Business Analytics	193
9.3.1	Klassische Business Intelligence-Architekturen	193
9.3.2	Zukünftige Business Intelligence-Architektur als semi-virtuelles Data-Warehouse	195
9.3.3	Konsequenzen für die Informationslandschaft	197
9.4	Aktuelle Beispiele aus der betrieblichen Praxis	198
9.4.1	Steigende Datenmenge und zunehmende Integration unstrukturierter Daten erhöhen die Komplexität von Business Intelligence	198
9.4.2	Potentiale von In-Memory-Technologie am Beispiel von Handelsunternehmen	199
9.4.3	Unterstützung von Geschäftsprozessen durch In-Memory-Technologie im Bereich Predictive Maintenance	200
9.5	Fazit und Ausblick	201
	Literatur	201

T. Knabke (✉)

Universität Duisburg-Essen, Lehrstuhl für Wirtschaftsinformatik,
insbesondere Business Intelligence, Duisburg
E-Mail: tobias.knabke@uni-due.de

S. Olbrich

Universität Duisburg-Essen, Lehrstuhl für Wirtschaftsinformatik,
insbesondere Business Intelligence, Duisburg
E-Mail: sebastian.olbrich@uni-due.de

Zusammenfassung

In der unternehmerischen Praxis wachsen die Anforderungen an eine stärkere Verzahnung von operativen und analytischen Tätigkeiten. Dies erfordert ein Umdenken sowohl in der Organisation als auch in der IT-Strategie, da in den meisten Unternehmen derzeit operative und analytische Systeme voneinander getrennt sind. Eine Verschmelzung von operativen und analytischen Tätigkeiten ist mit den bisher eingesetzten IT-Systemen und den aktuellen Technologien nicht oder nur mit erheblichem Aufwand möglich. Durch die zunehmende Wirtschaftlichkeit von In-Memory (IM)-basierten Datenbanken und Applikationen verringert sich der zu erbringende Aufwand jedoch erheblich. Die IM-Technologie besitzt damit das Potential, auch im Unternehmensumfeld einen Zustand zu etablieren, der im privaten Alltag längst selbstverständlich ist – eine durchgängige Informationsversorgung.

Der einführende Abschnitt untersucht dieses Potential näher und umreißt das Spannungsfeld zwischen transaktionalen und analytischen Informationssystemen. Anschließend werden die technologischen Grundlagen der IM-Technologie skizziert. Dabei stehen die Funktionen der Datenanalyse im Vordergrund der Betrachtung. Der darauf folgende Abschnitt widmet sich den Auswirkungen, die der Einsatz von IM-Technologie auf die Bereiche Business Intelligence (BI) und Business Analytics hat. Der letzte Abschnitt beschreibt konkrete Anwendungsszenarien. Ausgehend von den aktuellen Entwicklungen wird ein Ausblick über zukünftige Integrationsszenarien und -trends gegeben.

„Tape is Dead, Disk is Tape, Flash is Disk, RAM Locality is King.“ Jim Gray 2006

9.1 Einleitung und Motivation

1994 artikulierte Microsoft-Gründer Bill Gates die Vision der *„Information at your fingertips“*. In vielen Bereichen des Alltags ist diese heute Wirklichkeit geworden. So bedarf es bspw. bei einer Kontostandsabfrage weder eines persönlichen Besuchs in der Bankfiliale noch eines physischen Kontoauszugs. Diese Informationen sind sekundenschnell online verfügbar. Ebenso hat individuell konfigurierbares Fernsehen Einzug in das tägliche Leben erhalten. Derartige Dienste sind nicht nur zuhause verfügbar, sondern können nahezu überall und jederzeit mit mobilen Endgeräten genutzt werden. Darüber hinaus bieten Internetsuchmaschinen Suchvorschläge und Trefferlisten in Echtzeit an – und dies bei einer ständigen, ortsunabhängigen Verfügbarkeit.

Eine zeitnahe Informationsversorgung ist im Unternehmensumfeld derzeit jedoch oftmals die Ausnahme. Besonders bei komplexen Anfragen oder großen Datenmengen, wie sie häufig bei Business Analytics und Business Intelligence (BI) vorliegen, sind Antwortzeiten von mehreren Minuten oder gar Stunden keine Seltenheit. Um die operativen Systeme mit solchen Anfragen nicht zu belasten und um optimierte Ablagestrukturen für Analysen nutzen zu können, sind operative und analytische Systeme derzeit technisch

voneinander getrennt. Durch die Nutzung von IM-Technologien könnten Antwortzeiten von unter einer Sekunde („*the speed of thought*“, Plattner und Zeier 2011) auch für komplexe Analyseszenarien zur Regel werden (Plattner 2009; Plattner und Zeier 2011). Der Einsatz von IM-Technologie in operativen Enterprise-Resource-Planning (ERP)-Anwendungen kann zu einer Aufweichung der traditionellen Trennung von transaktionsorientierten (Online-Transaction-Processing-Systemen, OLTP) und analyseorientierten (Online-Analytical-Processing-Systemen, OLAP) Systemen führen. Diese Trennung ist hauptsächlich auf die hard- und softwarebedingten Restriktionen im Zusammenhang mit Analysen großer Datenmengen zurückzuführen und war zu Zeiten der Etablierung von Data-Warehouse-(DWH)-Systemen ein unabdingbarer Kompromiss zwischen erwarteter Flexibilität und Analysegeschwindigkeit sowie zusätzlichem Aufwand für die Informationsbereitstellung und Versorgung (Chamoni 2011; Plattner 2009). Allerdings birgt eine solche Umstellung technische Herausforderungen, da transaktionale und analytische Anfragen grundsätzlich andere Anforderungen an die Datenorganisation stellen (French 1995). Besonders für Business Analytics und BI wird ein Technologiewechsel weitreichende Auswirkungen, z. B. auf die vorhandenen Architekturen und Datenmodelle, haben.

9.2 Grundlagen In-Memory-Datenbanken

Im betrieblichen Umfeld besteht die Anforderung an eine zunehmend schnellere Informationsversorgung. So entwickelt sich das Reporting von einer vergangenheitsbezogenen Betrachtung hin zu einer aktiveren Steuerung, auch von operativen Unternehmensdaten (vgl. Abschn. 4). Neben technischen Veränderungen betrifft dies unterschiedlichste Bereiche vom Organisationsmanagement bis hin zur Unternehmenskultur. Im folgenden Abschnitt wird der technologische Aspekt näher beleuchtet.

9.2.1 Aktuelle Entwicklungen im Hauptspeicherbereich

Die derzeitige Trennung von operativen und analytischen Systemen, stark steigende Datenmengen sowie die Forderung einer zeitnahen Informationsversorgung stellen hohe Anforderungen an Leistungsfähigkeit und Wirtschaftlichkeit der eingesetzten Technologie. Die Preisentwicklung für Speicher kommt dieser Forderung entgegen. In den letzten Jahrzehnten ist die Leistungsfähigkeit von Speicher exponentiell angestiegen; umgekehrt ausgedrückt sind die Preise für Speicher stark gefallen. Dies gilt insbesondere für Hauptspeicher (Arbeitsspeicher, random access memory, RAM). Neben klassischen festplattenbasierten Datenbanksystemen (disk-resident database, DRDB) rücken Hauptspeicherbasierte Datenbanken (in-memory database, IMDB oder main-memory database, MMDB) immer stärker in den Fokus als Massenspeicher von Unternehmensdaten. Mit ihrem Einsatz kann der Spagat zwischen hoher Leistungsfähigkeit auf der einen und Wirtschaftlich-

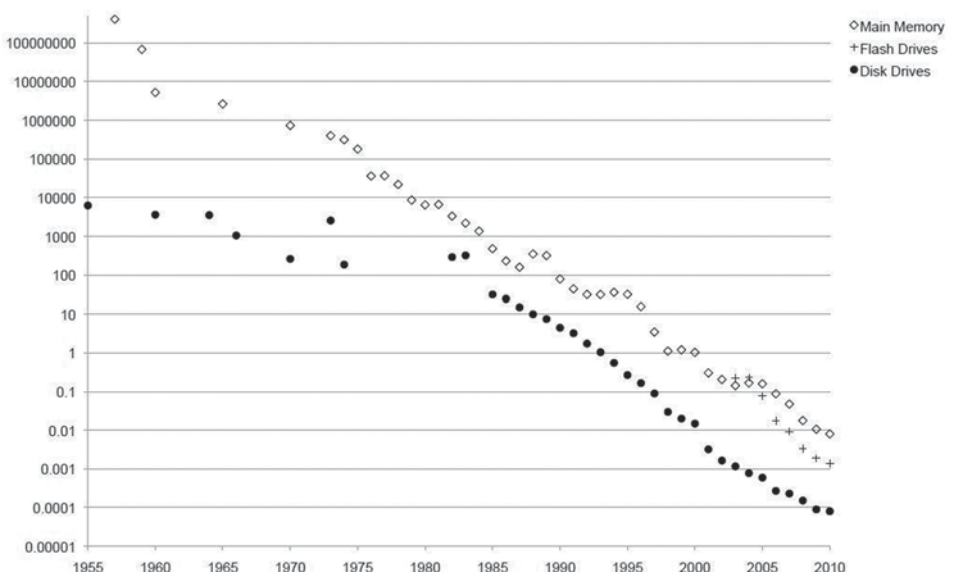


Abb. 9.1 Entwicklung von Speicherkosten in US-Dollar pro Megabyte. (Plattner und Zeier 2011, S. 15)

keit auf der anderen Seite gelingen, da der Preisverfall der letzten Jahre die IM-Technologie für Unternehmen realisierbar und wirtschaftlich gemacht hat. Diese Preisentwicklung von Haupt-, Festplatten- und Flash-Disk-Speicher ist in Abb. 9.1 dargestellt. Über die letzten Jahrzehnte ist das Verhältnis von Kosten zu Speichergröße exponentiell gesunken. So betrug der Preis für 1 MB Festplattenspeicher im Jahr 1970 mehr als 250 US\$, wohingegen der Preis in 2001 unter 0,01 US\$ gefallen ist. Eine Entwicklung, die sich bis heute fortsetzt und auch für Arbeitsspeicher zu beobachten ist (Plattner und Zeier 2011).

Doch nicht nur der Preisverfall von Hardware ist ein Grund für die verstärkte Nachfrage von Unternehmen. Wie Tab. 9.1 zeigt ist der Zugriff auf den Arbeitsspeicher um ein Vielfaches (im Bereich von vier Zehnerpotenzen) schneller als auf Festplattenspeicher (Garcia-Molina und Salem 1992; Plattner und Zeier 2011). Besonders ins Gewicht fällt diese Differenz beim Lesen von großen Datenmengen.

In einem IM- basierten Datenbankmanagementsystem (DBMS) werden Daten permanent im Hauptspeicher vorgehalten, der direkt vom Hauptprozessor zugreifbar ist. Die

Tab. 9.1 Zugriffszeiten für Festplatten und Hauptspeicher. (Plattner und Zeier 2011)

Aktivität	Zeit (ns)
Hauptspeicherzugriff	100
Sequentielles Lesen von 1 MB aus dem Hauptspeicher	250.000
Suche auf dem Festplattenspeicher/Positionierung Festplattenkopf	5.000.000
Sequentielles Lesen von 1 MB von der Festplatte	30.000.000

Vorteile eines solchen Systems lassen sich auch durch DRDB-Systeme mit einem sehr großen Cache nicht kompensieren, da Zugriff und Indexstrukturen für Festplattenzugriff ausgelegt sind (Garcia-Molina und Salem 1992). Konzepte für den Einsatz von IM-Technologien sowie Überlegungen zu notwendigen Änderungen bei Implementations- und Zugriffstechnologien gegenüber DRDB sind nicht neu, sondern bestehen bereits seit rund drei Jahrzehnten (DeWitt et al. 1984; Garcia-Molina und Salem 1992; Eich 1987a, b). Dabei kommt der Datenorganisation eine wesentliche Bedeutung zu.

9.2.2 Datenorganisation

Die geeignete Organisation und Ablagestruktur der Daten auf dem Speichermedium wird wesentlich durch den Verwendungszweck bestimmt (DeWitt et al. 1984; Stonebraker et al. 2005; Schmidt-Volkmar 2008; Plattner 2009; Schaffner et al. 2009; Plattner und Zeier 2011). Schreiboptimierte, transaktionale Systeme (OLTP) bevorzugen eine zeilenbasierte Datenablage. Solche klassischen Systeme der Datenverarbeitung, wie es bei ERP-Systemen der Fall ist, sind geprägt von vielen Schreiboperationen und Änderungen an bestehenden Datensätzen. Allerdings ist eine derartige Datenorganisation weniger gut dazu geeignet, Ergebnismengen von einzelnen Spalten wiederzugeben. Mengenbasierte Leseoperationen wiederum werden durch eine spaltenorientierte Datenhaltung besser unterstützt. Diese leseoptimierten Abfragen stehen bei OLAP-Systemen im Vordergrund und sind kennzeichnend für BI-Applikationen. Der Unterschied zwischen zeilen- und spaltenbasierter Datenorganisation ist exemplarisch in Abb. 9.2 dargestellt. Die Datenbank enthält drei Datensätze (Zeilen z) mit je drei Attributen (Spalten s). In zeilenorientierten Datenbanken wird jede Zeile in zusammenhängenden Blöcken gespeichert. Bei Spaltenorientierung hingegen werden die Attributwerte der einzelnen Spalten einer Datenbanktabelle nacheinander abgelegt (Abadi et al. 2009).

Bei Analyseanwendungen mit vorwiegend mengenbasierten Leseoperationen werden meist sehr wenige, etwa 10% aller Attribute einer Tabelle verwendet (Plattner 2009). Bei

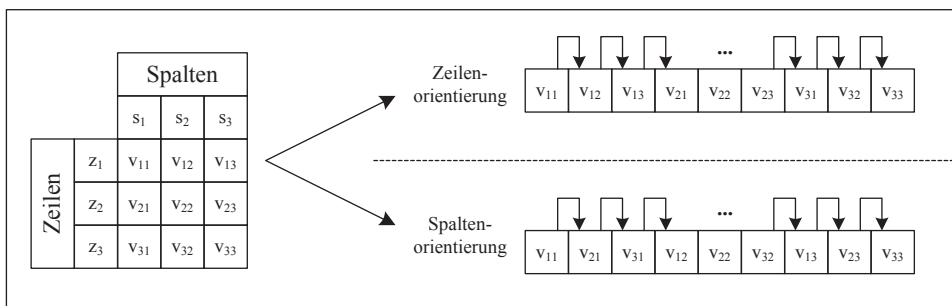


Abb. 9.2 Datenhaltungsalternativen. (In Anlehnung an Plattner und Zeier 2011)

Abb. 9.3 Zugriffsverhalten bei spaltenorientierter Datenhaltung. (vgl. Plattner 2009)

		Spalten		
		S ₁	S ₂	S ₃
Zeilen	Z ₁	V ₁₁	V ₁₂	V ₁₃
	Z ₂	V ₂₁	V ₂₂	V ₂₃
	Z ₃	V ₃₁	V ₃₂	V ₃₃

Select s₁ from db-table
where s₃ > 10

einem spaltenorientierten Design wird in diesem Fall nur auf die relevanten Datensätze und Spalten zugegriffen (vgl. Abb. 9.3) und nicht wie bei der zeilenorientierten Alternative zuerst auf den gesamten Datenbestand. Da eine Spalte nur die Werte einer Domäne bzw. eines Datentyps enthält, z. B. Jahreszahlen, und somit ähnliche Ausprägungen hat, ist eine bessere Komprimierung möglich. Darüber hinaus wird mit steigender Kompressionsrate die Leseperformance verbessert. Untersuchungen auf Systemen mit realen Unternehmensdaten ergaben Performancevorteile von bis zu Faktor 1000 gegenüber Zeilenorganisation (Plattner 2009). Zusätzlich wird durch die Komprimierung der Speicherverbrauch verringert. Plattner (2009) geht unter Berücksichtigung von Datenkompression, nicht mehr benötigten Aggregaten (materialisierte Views) und horizontaler Partitionierung von einem um Faktor 50 besseren Speicherverbrauch gegenüber DRDB aus. Zusammen mit der Parallelisierung und Zuweisung der Daten auf verteilte Hardwarekomponenten (Blades) lassen sich so auch die Daten von Großunternehmen vollständig in-memory halten (Plattner 2009).

Wie bereits angedeutet basiert die aktuell vorherrschende Trennung von OLTP- und OLAP-Systemen in der betrieblichen Praxis auf bestehenden Restriktionen (Integration und Harmonisierung, Performance, Kosten) (Wessel et al. 2013; Krüger et al. 2010). Um die Quellsysteme für analytische Auswirkungen nicht zu großer Last auszusetzen, werden Daten für dispositive Zwecke zusätzlich in ein dediziertes System, meist ein DWH, überführt. In diesen Systemen liegen die Informationen in der Auswertungsschicht in der Regel nicht so granular wie in den Quellsystemen vor. Darüber hinaus werden die Daten in vielen BI-Tools in einem denormalisierten Modell, etwa dem Star-Schema (Kimball 1996; Moody und Kortink 2000), gehalten, um den hohen Performanceanforderungen gerecht zu werden. Dadurch sind effiziente Leseoperationen auf großen Datenvolumina möglich. Nichtsdestotrotz stoßen diese Systeme, die gewöhnlich DRDBs als Speichermedium zur Datenhaltung nutzen, in puncto Performance zunehmend an ihre Grenzen. Seit einigen Jahren ist daher zu beobachten, dass Performanceengpässe in BI-Systemen durch den gezielten Einsatz von IM-Technologie für Analysen umgangen werden. Dabei setzen die großen Softwareanbieter hauptsächlich auf kostenpflichtige, zusätzliche Architekturbau- steine, die den Zugriff auf zeitkritische Auswertungen beschleunigen, indem sie die für das Reporting relevanten Daten im Hauptspeicher spaltenbasiert vorhalten.

Zwar mag eine Trennung von OLAP und OLTP den Bedürfnissen und Eigenschaften der unterstützten Prozesse und Datenstrukturen und sowie einer reduzierten Arbeitslast

der operativen Systeme gerecht werden (Stonebraker 2011), für Anwender aber gewinnen zeitnahe analytische Applikationen auf operativen Daten immer stärker an Bedeutung. Um dieser Forderung nachzukommen, finden sich zunehmend Technologien, die beide Ansätze auf einer Plattform verbinden und sowohl zeilen- als auch spaltenbasierte Datenorganisation bereitstellen. Da diese Architekturen OLTP und OLAP vereinen, werden sie auch als OLXP bezeichnet (Loos et al. 2011). Dieser Ansatz berücksichtigt die Kritik an der hohen Komplexität der Informationsversorgung und der redundanten Datenhaltung in einem separaten System (Chamoni 2011).

Um alle Möglichkeiten der IM-Technologie optimal nutzen zu können, bieten viele Hersteller sogenannte Appliances (d. h. eine Verknüpfung von Hard- und Software zu Services) an. Appliances umfassen nicht nur die Bereitstellung der reinen Technologie im Sinne der Hardware, sondern bieten gleichzeitig zahlreiche eingebaute, hardwarenahe Funktionen, die auf spezielle Anwendungsfälle ausgelegt sind. Diese verlagern einen Teil der Logik von der Anwendungs- zur Datenbankschicht, was u. a. in Planungsapplikationen zum Tragen kommt (Färber et al. 2010). Die bei Verteilungsaufgaben (disaggregation) notwendigen Berechnungen werden in diesem Ansatz nicht mehr in der Applikationsschicht durchgeführt und die Ergebnisse dann an die Datenbank geschickt. Stattdessen wird der zu verteilende Wert an die hardwarenahen Funktionen übergeben, die Berechnungen direkt auf der Datenbank ausgeführt und das Ergebnis dann zurückgeliefert. Über einen Performancegewinn hinaus reduziert dies auch die über das Netzwerk ausgetauschten Daten.

9.3 Konsequenzen für Business Intelligence und Business Analytics

Viele BI-Systeme basieren auf einem DWH-Ansatz, um Daten systematisch aus den operativen Vorsystemen zu extrahieren, harmonisieren und sind für das Reporting bereitzustellen. Mit zunehmendem Einsatz von IM-Technologie stellt sich die Frage, inwieweit dies Auswirkungen auf bewährte Architekturen und Ansätze hat.

9.3.1 Klassische Business Intelligence-Architekturen

Die bedeutendsten Anforderungen an ein DWH als Grundlage der dispositiven Informationsversorgung gehen auf Inmon (1996) zurück und sind Themenorientierung (subject-orientation), Integration bzw. Vereinheitlichung (integration), Zeitorientierung (time-variance) sowie Beständigkeit (non-volatility). Um eine adäquate Entscheidungsunterstützung gewährleisten zu können, liegt klassischen BI-Systemen eine mehrschichtige (DWH-) Architektur zugrunde. Der Aufbau einer solchen Referenzarchitektur ist in Abb. 9.4 dargestellt.

Im ersten Schritt werden die Daten aus den operativen Quellsystemen, bspw. ERP-Systemen, geladen und i.d.R. unverändert in die Extraktionsschicht übernommen. Wäh-

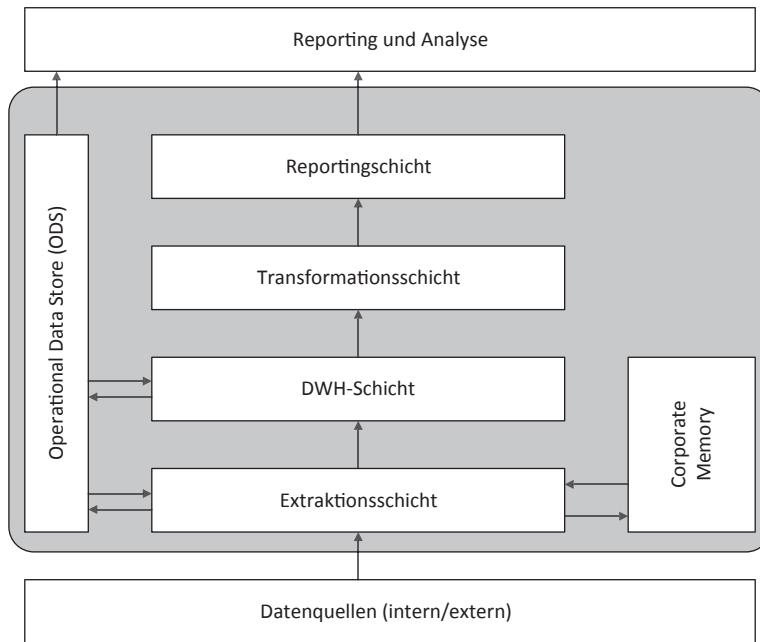


Abb. 9.4 Klassische DWH-Architektur

rend des Transformationsvorgangs in die DWH-Schicht finden dann Datenbereinigung, -harmonisierung und -konsolidierung statt. Diese Schicht stellt den Single Point of Truth (SPOT) dar, in dem die Daten applikationsunabhängig und konsolidiert vorliegen. Der SPOT bildet somit auch die Basis für funktionsspezifische Anwendungen (Hahne 2002, 2010). In den oberen Architekturkomponenten werden die Anforderungen der Fachbereiche und des Managements berücksichtigt, indem die unternehmensweit gültigen Informationen aus der DWH-Schicht mit bereichsspezifischer Logik angereichert werden. Die organisations- und anwendungsdomänenspezifische Ausrichtung in der Transformations- und Reportingschicht bezeichnet man als Data Mart. Um Performanceanforderungen für die Analyse gerecht zu werden, werden die Daten auf dem Weg in die Reportingschicht bei Bedarf aggregiert. Durch die redundante Datenhaltung und physische Speicherung in mehreren Schichten sind DWH-Systeme von mehreren Terabyte (TB) keine Seltenheit (Kemper et al. 2006; Hahne 2010). Das corporate memory enthält alle in das BI-System geladenen Daten und sichert die Unabhängigkeit von den operativen Quellsystemen im Falle von Anpassungsbedarf oder im Fehlerfall (Kimball 1996; Chamoni und Gluchowski 1998; Knabke und Olbrich 2011).

Traditionell dient BI der dispositiven Informationsversorgung. Diese strategische und taktische Entscheidungsunterstützung wird sukzessive auf operative Bereiche ausgedehnt. Damit erhält BI in die kurzfristige Prozessunterstützung Einzug (Baars et al. 2014). So lässt sich z. B. im Qualitätsmanagement der Produktionsindustrie der Ausschuss fehler-

hafter Werkstücke frühzeitig erkennen. Mithilfe von Sensordaten können Unregelmäßigkeiten im laufenden Produktionsprozess identifiziert und im Fehlerfall automatisch interveniert werden, um korrigierende Maßnahmen durchzuführen. Die Datengrundlage für ein operationales BI-System (OPBI) bilden granulare Datenspeicher, die konzeptionell unter dem Begriff Operational Data Stores (ODS) zusammengefasst werden (Winter 2000) und Analysen auf operativen Daten höchster Detaillierung ermöglichen (Inmon 1998).

Ein Nachteil der Nutzung von Hauptspeicher zur persistenten Speicherung von Daten liegt darin, dass er volatil ist und eine Unterbrechung der Stromversorgung zum Datenverlust führt. Daher sind zuverlässige und effiziente Datenwiederherstellungsstrategien bei dieser Speicherart unverzichtbar (Plattner und Zeier 2011).

Auf Basis der technologischen Fortschritte im Bereich Speichermedien, der stetig steigenden Datenmenge sowie einer zunehmend operativen Entscheidungsunterstützung ergeben sich Änderungen in der DWH-Architektur als zentraler Bestandteil von BI.

9.3.2 Zukünftige Business Intelligence-Architektur als semi-virtuelles Data-Warehouse

Insbesondere durch den Einsatz von IM-Datenbanken als Grundlage von transtransaktionalen und dispositiven IT-Systemen liegt der Gedanke nahe, Schichten nicht mehr zu persistieren, sondern lediglich logisch zu verwalten. Die klassischen DWH-Architekturen werden somit virtualisiert und die bisher hauptsächlich aus Performancegründen eingeführten materialisierten Schichten und Aggregate überflüssig. Abbildung 9.5 zeigt ein semi-virtuelles Data-Warehouse (SVDWH) auf Grundlage einer IMDB (Knabke und Olbrich 2011). Vergleichbar mit einem klassischen DWH basiert das semi-virtuelle DWH auf einer schichtenbasierten, skalierbaren Architektur. Allerdings werden die Daten nun komplett in-memory gehalten und nur die Extraktionsschicht separat gespeichert. Darüber liegende Schichten wie die DWH-Schicht sind rein logisch modelliert. Die Daten werden bei Analyseanfragen zur Laufzeit bereinigt, harmonisiert, berechnet und aggregiert, ohne sie dediziert abzulegen. Die Datenmodelle werden analog zur klassischen Architektur in einem Metadatenspeicher hinterlegt. Sind historische Daten verlässlich und langfristig bereits in einem IM-basierten Quellsystem vorhanden, kann auf die gesonderte Ablage in der Extraktionsschicht verzichtet werden (daher semi-virtuell). Darüber hinaus sollte das Quellsystem über geeignete Strukturen der Datenorganisation verfügen, wie etwa spaltenbasierte oder hybride Ansätze, d. h. Zeilenorganisation für OLTP und spaltenbasierte Speicherung für OLAP Anfragen. Die in Abbildung 9.5 vorgestellte SVDWH-Architektur berücksichtigt auch heterogene Applikationslandschaften, wie sie in vielen Organisationen zu finden sind (Knabke und Olbrich 2011).

Der ETL-Prozess in BI-Systemen ist ein kritischer und gewöhnlich zeitaufwändiger Prozess, der bei großen Organisationen nicht selten eine komplette Nacht in Anspruch nimmt. Da auf Analyse ausgelegte Systeme leseoptimiert sind, d. h. spaltenbasiert und komprimiert, bietet sich ein zusätzlicher, schreiboptimierter Speicher zur Datenübernah-

9.3.3 Konsequenzen für die Informationslandschaft

Der Einsatz von IM-Technologie als Basis eines SVDWH und somit von BI führt zu Veränderungen im Vergleich zur heutigen Informationslogistik (Knabke und Olbrich 2011).

Datenmodellierung Die redundante, physische Datenhaltung in einer skalierbaren Schichtenarchitektur (layered scalable architecture, LSA) entfällt größtenteils. Mit Ausnahme der Extraktionsschicht findet die Modellierung logisch statt. Auf Aggregate oder andere aus Performancegründen eingeführte Objekte und Schichten kann zugunsten weniger und nicht materialisierter Schichten verzichtet werden. Durch den Verzicht auf diese Verarbeitungsschichten werden Datenredundanzen und -inkonsistenzen vermieden und die Komplexität der Architektur reduziert (Winter et al. 2011). Eine mehrdimensionale Modellierung in der Auswertungsschicht wird weiterhin empfohlen, um eine leichte Anwendbarkeit durch Fachanwender sicherzustellen. Allerdings ist die logische Modellierung in der Auswertungsschicht ausreichend und muss nicht physisch abgelegt werden. Anpassungen von bestehenden Datenmodellen erfolgen dadurch schneller, insbesondere dann, wenn logische Modellierung mit Transformationen in Echtzeit verwendet werden. Eine logische Modellierung ohne zusätzliche Datenspeicherung ermöglicht ein Ad-hoc-Reporting unter Berücksichtigung von geänderten Anforderungen, ohne dass eine Neubeladung der Daten über alle Schichten des DWH notwendig ist.

Datenanalyse Auf eine Performanceoptimierung zur Sicherstellung der erwarteten Antwortzeiten kann weitestgehend verzichtet werden. So entfällt auch die Vorberechnung von Daten (Aggregate) und eine Vorgabe von Analysen. Dies erhöht die Individualität, Spontaneität und Flexibilität der Datenanalyse, weil potentielle Szenarien nicht ex ante antizipiert werden müssen, sondern diese flexibel zur Laufzeit berechnet werden. Die Analyseperformance steigt unter Verwendung von IMDB mit spaltenorientierter Datenhaltung aus mehreren Gründen signifikant: Die Daten befinden sich bereits im Hauptspeicher und müssen nicht erst von langsameren Medien wie Festplatten geladen werden. Darüber hinaus berücksichtigt eine spaltenorientierte Datenablage die Eigenschaften von Analyseabfragen, die vorwiegend nur eine kleine Menge der verfügbaren Attribute verwenden (vgl. Abschn. 2.2). Die Kombination mit besseren Komprimierungsmethoden erhöht den Performancevorteil zusätzlich. Logisch modellierte Echtzeitberechnungen von Aggregaten in Spaltendatenbanken bieten gegenüber vorberechneten, materialisierten Aggregaten in zeilenorientierter Datenhaltung große Geschwindigkeitsvorteile, wenn viele Aggregate benutzt werden (Plattner und Zeier 2011; Plattner 2009). Diese Verbesserungen tragen zur besseren Anpassungsfähigkeit der Systeme bei und steigern dadurch die Agilität der gesamten Infrastruktur.

Datenbereitstellung Die Datenbereitstellung in Echtzeit ist ein kritischer Erfolgsfaktor, wenn ein BI-System nicht nur zur strategischen Entscheidungsunterstützung, sondern auch zur operativen Prozessunterstützung eingesetzt wird. Im Gegensatz zu einem fest-

plattenbasierten Quellsystem werden die Daten im SVDWH-Ansatz nur in der Extraktionsschicht physisch gespeichert und im Übrigen direkt aus der Quelle gelesen (Knabke und Olbrich 2011). Neben einer Reduzierung des benötigten Speicherplatzes verkürzt dies die Zeit bis die Informationen für das Reporting und die Analyse zur Verfügung stehen. Transformationen und Berechnungen können on-the-fly durchgeführt werden, um eine bereinigte, harmonisierte und konsolidierte Datenbasis zu erhalten.

Anwendungsentwicklung Um das Potential von IM-Technologie nutzen zu können, ist vorhandenes Design und Coding der Applikationen zu überprüfen und an die neuen Gegebenheiten anzupassen (Loos et al. 2011). Das bisher akzeptierte Paradigma, Logik möglichst anwendungsnah vorzusehen, verschiebt sich in Teilen in Richtung Hardware.

Der Gedanke des virtuellen DWH konnte sich in der Vergangenheit aus verschiedenen Gründen bisher nicht durchsetzen. Durch den Einsatz von IM-Technologien in Kombination mit SVDWHs können viele Nachteile rein virtueller DWH behoben werden. Durch die so hinzugewonnene Flexibilität und Agilität werden semi-virtuelle Entscheidungsunterstützungssysteme in Zukunft eine wichtige Rolle für Business Intelligence und Business Analytics spielen. Dies trägt zur Verschmelzung von OLAP- und OLTP-Systemen bei, ohne Kernaspekte von DWHs wie Integration und Harmonisierung zu vernachlässigen.

9.4 Aktuelle Beispiele aus der betrieblichen Praxis

Die Ursprünge von BI im unternehmerischen Umfeld liegen in einem stabilen, harmonisierten und vor allem vergangenheitsbezogenem Reporting. Steigende Marktdynamik und Analysebedarfe erfordern jedoch eine aktive Steuerung und agile Ausrichtung auf Grundlage datengetriebener Konzepte. Dafür ist eine proaktive Entscheidungsunterstützung ein wesentlicher Bestandteil, die zunehmend auf der Grundlage von externen Informationsquellen beruht. Entsprechend ist eine stärkere Operationalisierung von BI zu beobachten, die mit einer steigenden Datenmenge sowie -komplexität und einem erhöhten Integrationsdrang externer Datenquellen einhergeht.

9.4.1 Steigende Datenmenge und zunehmende Integration unstrukturierter Daten erhöhen die Komplexität von Business Intelligence

Diese Herausforderung und der derzeit größte Trend im Bereich BI wird unter dem Begriff Big Data Analytics zusammengefasst (Chamoni 2011). Darunter versteht man mehr als nur die Analyse sehr großer Datenmengen (volume). Die Informationsversorgung zu bestimmten, fest definierten Zeitpunkten ist für heutige Analysebedarfe nicht mehr ausreichend. Die Herausforderung ist eine kontinuierliche Datenaufnahme, möglichst in Echtzeit (velocity). Nicht minder fordernd als eine zeitnahe Bereitstellung der Informationen ist die

Bewältigung der Datenvielfalt (variety). Besonders die Kombination aller bzw. mehrerer der 3 Vs bringt einen Mehrwert in der Verwendung von Big Data Analytics (Funke und Olbrich 2015). Neben strukturierten Daten, die bisher den BI-Bereich dominierten, spielen nun auch semi-strukturierte und unstrukturierte Daten z. B. aus externen Datenquellen wie Social-Media-Plattformen in den Auswertungen eine Rolle. Einer Studie des TDWI (Russum 2011) zufolge wird besonders die Integration von semi- sowie unstrukturierten Daten stark zunehmen, was die Komplexität der Datenintegration erhöht. Ein weiterer Fokus liegt auf der Integration in Echtzeit, der Datenqualität und dem Complex Event Processing. Beispiele für Big-Data-Anwendungsfälle sind das Aufspüren von Unregelmäßigkeiten im Finanz- und Versicherungsumfeld (Fraud-Detection) sowie eine intelligente Energieverbrauchssteuerung (Smart Metering). In vielen Unternehmen wird Big Data von IT-Abteilungen in Form von technologischen Pilotprojekten betrieben. Erfolgreiche Big Data-Initiativen erfordern jedoch stets auch einen fachlichen Bezug. Maßgeblich für den Erfolg sind Anwendungsfälle mit einem hohen Wertbeitrag, die ein konkretes fachliches, bisher nicht oder unzureichend durch Informationssysteme unterstützbares analytisches Problem lösen. Unter technischen Gesichtspunkten ist die Frage zu klären, welche Daten wie lange in-memory vorgehalten werden sollen und wann Daten auf günstigere Medien und Technologien ausgelagert werden.

9.4.2 Potentiale von In-Memory-Technologie am Beispiel von Handelsunternehmen

Handelsunternehmen weisen traditionell hohe Datenvolumina auf, was besondere Anforderungen an Handelsinformationssysteme und deren Performance darstellt. Große Handelsunternehmen stehen vor der Herausforderung, die Bedarfe in ihren Einzelhandelsketten, den Schnittstellen zum Großhandel sowie die Anforderungen der Zentralbereiche gleichermaßen abdecken zu müssen (Schütte 2011). Dazu kommen große Mengen Stamm- und Bewegungsdaten, die verstärkt für Analyseanwendungen zur Verfügung stehen müssen. Betrachtet man beispielsweise Kassensbondaten (point-of-sale, POS), so erreichen alleine diese Systeme schnell mehrere hundert Millionen Datensätze täglich. Diese Daten können zur Abverkaufs- und Bedarfsprognoserechnung herangezogen werden. Der Einsatz von IM-Technologien führt zu einer verbesserten Entscheidungsunterstützung, da bspw. Optimierungsläufe zur Forecast- und Replenishment-Berechnung aufgrund des großen Ressourcenbedarfs ohne diese Technologie häufig nur über Nacht durchgeführt werden können (Schütte 2011). Eine kurzfristige Anpassung und erneute Berechnung ist damit ausgeschlossen. Weiteres Potenzial bieten Auswertungsmöglichkeiten von POS-Daten zur Erfolgsbestimmung und Planung von Promotionskampagnen (Schütte 2011). Die Aufhebung der künstlichen Separierung von transaktionalen und analytischen Systemen, die durch die Verwendung von IMDB vorangetrieben wird, erleichtert die Aktionsplanung auf taktischer Ebene. Artikelplanung auf Basis von Abverkaufsdaten sowie die Definition der Artikel würden nicht mehr in getrennten Systemen, sondern zukünftig integriert erfolgen (Schütte 2011). Zudem bietet sich gerade im Einzelhandelsumfeld die

Integration von externen Kundeninformationen aus Kundenbindungsprogrammen an. Online-Händler können darüber hinaus wertvolle Informationen aus der Anbindung und Auswertung von Social-Media-Daten gewinnen.

9.4.3 Unterstützung von Geschäftsprozessen durch In-Memory-Technologie im Bereich Predictive Maintenance

In Analysesystemen werden häufig Modellierungsprinzipien wie das Star-Schema verwendet, um Performanceeinschränkungen aufgrund großer Datenvolumina und DRDB zu kompensieren. Im Gegensatz zur relationalen Modellierung weisen diese Ansätze in Kombination mit einer mehrschichtigen Architektur bei einigen Anwendungsfällen Nachteile auf. Im Bereich der Wartung, Reparatur und Überholung von Fahrzeugen oder Flugzeugen ist es eine essenzielle Anforderung, Preise für individuelle Serviceangebote zu kalkulieren und zu simulieren. Je nach Größe des Anbieters können die Materiallager über ein ganzes Land oder sogar weltweit verteilt sein, wie es in der Luftfahrtindustrie häufig der Fall ist. Der Dienstleister muss einen Überblick haben, welche Teile in welchem Zustand in welchem Lager verfügbar sind. Darüber hinaus sind zahlreiche Stammdateninformationen wie Alter, Kompatibilität oder der finanzielle Wert der Komponente notwendig. Um Aussagen über die Wertentwicklung eines Gerätepools treffen oder die dauerhafte Verfügbarkeit von Komponenten nachweisen zu können, werden die Informationen sowohl für den aktuellen Zeitpunkt als auch mit ihrem historischen Verlauf benötigt. Für die Preisgestaltung von Verträgen spielen die Zulässigkeit und der Wunsch des Kunden zum Einsatz von überholten und reparierten Komponenten eine große Rolle. Dies muss in Simulationen und Angebotsgestaltungen Berücksichtigung finden. Für die Analysen werden die Daten auf sehr granularer Basis ausgehend von Materialbewegungen benötigt. Eine besondere Herausforderung sind dabei n:m Beziehungen, z. B. von Komponente und Fahrzeug- bzw. Flugzeugtyp, da eine Komponente in mehrere Typen passt, ein Typ aber gleichzeitig aus mehreren Komponenten besteht. Diese Beziehung kann über Wartungslisten als Mappingtabelle abgebildet werden. Die aufgelöste Beziehung muss mit den Materialbewegungen verbunden werden, was das Datenvolumen in einem denormalisierten DWH-Ansatz mit mehreren physischen Schichten rapide ansteigen ließe. Die in der Praxis für Materialbewegungen üblichen Größenordnungen von hundert Millionen Datensätzen oder mehr müssten in diesem multidimensionalen Ansatz noch pro Komponenten-Typ-Kombination vervielfacht werden und ließen das Datenvolumen zusätzlich rapide ansteigen. Diese Vervielfachung von Sätzen wird durch den Einsatz referenzierender Datenstrukturen verhindert, da die Daten nur einmal physisch gespeichert werden und die notwendigen Berechnungen für die Analysen virtuell zur Laufzeit über Views durchgeführt werden. So wird eine notwendige Materialisierung aller möglichen Kombinationen ex ante vermieden. Somit ermöglicht der Einsatz von IM-Technologie oftmals erst die Geschäftsprozessunterstützung durch BI (Knabke et al. 2014). Zudem entfällt die komplette Neubeladung der Daten durch die relationale Modellierung im Gegensatz zur multidimensionalen Model-

lierung, wenn die Wartungslisten geändert oder erweitert werden. Dies ermöglicht flexible Analysen des Lagerbestands und Simulation der Wartungsintervalle.

9.5 Fazit und Ausblick

Wie anhand der voranstehenden Beispiele deutlich wurde, bringt der Einsatz von IM-Technologie einige wesentliche Vorteile für die betriebliche Praxis mit sich. Neben der beträchtlichen Erhöhung der Verarbeitungsgeschwindigkeit in der betrieblichen Informationsverarbeitung besteht das Potential, die aktuell vorherrschende Trennung von OLTP- und OLAP-Logik aufzuheben. Indem die Systemgrenzen von ERP und BI in den Organisationen durch den Einsatz von IMDB zunehmend verschwimmen, kann ein wesentlicher Meilenstein in Richtung einer durchgängigen Informationsversorgung erreicht werden.

Zur vollständigen Hebung dieses Potentials sind allerdings einige Bedingungen zu erfüllen: Zunächst einmal basieren die Zahlen für Performancegewinn meist auf Laborversuchen. Das Erreichen ähnlicher Größenordnungen im heterogenen betrieblichen Umfeld setzt genaue Kenntnisse über Technologie und die entsprechende Architektur der Anwendungen voraus. Ebenso ist ein erhöhter Einführungsaufwand durch die Umstellung der Datenmodelle zu erwarten. Durch die Positionierung der Produkte als Appliances (Verschmelzung von Hardware und Software) erhöht sich gleichzeitig die Abhängigkeit zu bestimmten Softwareanbietern.

Der zunehmenden Abhängigkeit der Kunden von der Produktausrichtung der großen ERP-Anbieter stehen weitreichende Integrationskonzepte gegenüber. Es ist schon jetzt zu beobachten, dass führende Anbieter von IM-Technologie ihre Produkte nicht mehr als reine Datenbanken vermarkten sondern auf die Funktionalitäten der Datenanalyse, des Datenmanagements und der Datenintegration hinweisen. Ziel ist der Aufbau eines konsolidierten Datenbestands über die gesamte Organisation hinweg, der jederzeit abrufbar ist. Der zukünftige Mehrwert wird voraussichtlich darin bestehen, diesen Datenbestand bzw. Teile davon auch für externe Geschäftspartner zugänglich zu machen und weitere Datenquellen außerhalb der eigenen Organisation in die Analysen einbeziehen zu können. Auf diese Weise können IM-Appliances als Katalysator für Verarbeitung von Big Data genutzt werden.

Literatur

- Abadi, D.J., Boncz, P.A., Harizopoulos, S.: Column-oriented database systems. Proc. VLDB Endow. **2**, 1664–1665 (2009)
- Baars, H., Felden, C., Gluchowski, P., Hilbert, A., Kemper, H.-G., Olbrich, S.: Shaping the next incarnation of business intelligence. Towards a flexibly governed network of information integration and analysis capabilities. Bus. Inf. Syst. Eng. **6**, 11–16 (2014)

- Calvanese, D., de Giacomo, G., Lenzerini, M., Nardi, D., Rosati, R.: Description logic framework for information integration. *Proceedings of the 6th International Conference on the principles of knowledge representation and reasoning (KR'98)*, S. 2–13 (1998)
- Chamoni, P.: BI-Strategie zum Ausgleich von Technologie-Push und Business-Pull. In: Lehner, W., Piller, G. (Hrsg.) *Innovative Unternehmensanwendungen mit In-Memory-Data-Management. Beiträge der Tagung IMDM 2011, 2.12.2011 in Mainz*, S. 13–22. Köllen Druck+Verlag GmbH, Bonn (2011)
- Chamoni, P., Gluchowski, P.: Analytische Informationssysteme – Einordnung und Überblick. In: Chamoni, P., Gluchowski, P. (Hrsg.) *Analytische Informationssysteme*, S. 3–25. Springer, Berlin (1998)
- DeWitt, D.J., Katz, R.H., Olken, F., Shapiro, L.D., Stonebraker, M.R., Wood, D.A.: Implementation techniques for main memory database systems. *Proceedings of the 1984 ACM SIGMOD International Conference on management of data*, S. 1–8. ACM, New York (1984)
- Eich, M.H.: A classification and comparison of main memory database recovery techniques. *Proceedings of the Third International Conference on data engineering*, S. 332–339. IEEE Computer Society, Washington, DC (1987a)
- Eich, M.H.: MARS: The design of a main memory database machine. *Proceedings of the 5th International Workshop on database machines*, S. 325–338 (1987b)
- Färber, F., Jäcksch, B., Lemke, C., Große, P., Lehner, W.: Hybride Datenbankarchitekturen am Beispiel der neuen SAP In-Memory-Technologie. *Datenbank-Spektrum* **10**, 81–92 (2010)
- French, C.D.: „One Size Fits All“ database architectures do not work for DSS, SIGMOD '95. *Proceedings of the 1995 ACM SIGMOD International Conference on management of data*, S. 449–450. ACM, New York (1995)
- Funke, K., Olbrich, S.: Increasing the value of big data projects – Investigation of industrial success stories. *Proceedings of the 48th Annual Hawaii International Conference on System Sciences*, (pp. Im Druck) (2015)
- Garcia-Molina, H., Salem, K.: Main memory database systems: an overview. *IEEE Trans. Knowl. Data Eng.* **4**, 509–516 (1992)
- Hahne, M.: *Logische Modellierung mehrdimensionaler Datenbanksysteme*. Deutscher Universitäts-Verlag, Wiesbaden (2002)
- Hahne, M.: Design des Enterprise Data Warehouse. Modellieren mehrschichtiger Architekturen. *BI-Spektrum* **5**, 8–12 (2010)
- Inmon, W.H.: *Building the data warehouse*. Wiley, New York (1996)
- Inmon, B.: The operational data store. *Information Management Magazine* (1998)
- Kemper, H.-G., Mehanna, W., Unger, C.: *Business Intelligence – Grundlagen und praktische Anwendungen. Eine Einführung in die IT-basierte Managementunterstützung*. Vieweg + Teubner, Wiesbaden (2006)
- Kimball, R.: *The data warehouse toolkit. Practical techniques for building dimensional data warehouses*. Wiley, New York (1996)
- Knabke, T., Olbrich, S.: Towards agile BI: applying in-memory technology to data warehouse architectures. In: Lehner, W., Piller, G. (Hrsg.) *Innovative Unternehmensanwendungen mit In-Memory-Data-Management, Beiträge der Tagung IMDM 2011, 2.12.2011 in Mainz*, S. 101–114. Köllen Druck+Verlag GmbH, Bonn (2011)
- Knabke, T., Olbrich, S., Fahim, S.: Impacts of in-memory technology on data warehouse architectures – a prototype implementation in the field of aircraft maintenance and service. In: Tremblay, M.C., et al. (Hrsg.) *Advancing the impact of design science: moving from theory to practice – 9th International Conference, DESRIST 2014, (Proceedings)*, S. 383–387, Miami (22–24 May 2014)
- Krüger, J., Grund, M., Tinnefeld, C., Eckart, B., Zeier, A., Plattner, H.: Hauptspeicherdatenbanken für Unternehmensanwendungen. *Datenmanagement für Unternehmensanwendungen im Kontext heutiger Anforderungen und Trends. Datenbank-Spektrum* **10**, 143–158 (2010)

- Loos, P., Lechtenböcker, J., Vossen, G., Zeier, A., Krüger, J., Müller, J., Lehner, W., Kossmann, D., Fabian, B., Günther, O., Winter, R.: In-memory databases in business information systems. *Bus. Inf. Syst. Eng.* **3**, 389–395 (2011)
- Moody, D.L., Kortink, M.A.R.: From enterprise models to dimensional models: a methodology for data warehouse and data mart design. In: Jeusfeld, M.A., et al. (Hrsg.) *Proceedings of the Second International Workshop on design and management of data warehouses, DMDW 2000*, Stockholm (CEUR-WS.org) (5, 6 June 2000)
- Plattner, H.: A common database approach for OLTP and OLAP using an in-memory column database: *Proceedings of the 35th SIGMOD International Conference on Management of Data*, Providence, Rhode Island (2009)
- Plattner, H., Zeier, A.: *In-Memory data management. An inflection point for enterprise application*. Springer, Berlin (2011)
- Russom, P.: Next generation data integration. TDWI best practices report, second quarter (2011)
- Schaffner, J., Bog, A., Krüger, J., Zeier, A.: A hybrid row-column OLTP database architecture for operational reporting: business intelligence for the real-time enterprise, S. 61–74. Springer, Berlin (2009)
- Schmidt-Volkmar, P.: *Betriebswirtschaftliche Analyse auf operationalen Daten*. Gabler, Wiesbaden (2008)
- Schütte, R.: Analyse des Einsatzpotenzials von In-Memory-Technologien in Handelsinformationssystemen. In: Lehner, W., Piller, G. (Hrsg.) *Innovative Unternehmensanwendungen mit In-Memory-Data-Management. Beiträge der Tagung IMDM 2011, 2.12.2011 in Mainz*, S. 1–12. Bonn (2011)
- Stonebraker, M.: Stonebraker on data warehouses. *Commun. ACM* **54**, 10–11 (2011)
- Stonebraker, M., Abadi, D.J., Batkin, A., Chen, X., Cherniack, M., Ferreira, M., Lau, E., Lin, A., Madden, S., O’Neil, E., O’Neil, P., Rasin, A., Tran, N., Zdonik, S.: C-store: a column-oriented DBMS. *Proceedings of the 31st VLDB Conference*, S. 553–564 (2005)
- Wessel, P., Köffer, S., Becker, J.: Auswirkungen von In-Memory-Datenmanagement auf Geschäftsprozesse im Business Intelligence: 11th International Conference on Wirtschaftsinformatik, S. 1781–1795, Leipzig, 27th February–1st March (2013)
- Winter, R.: Zur Positionierung und Weiterentwicklung des Data Warehousing in der betrieblichen Applikationsarchitektur: *Data-Warehousing-Strategie*, S. 127–139. Springer, Berlin (2000)
- Winter, R., Bischoff, S., Wortmann, F.: Revolution or evolution? Reflections on in-memory appliances from an enterprise information logistics perspective. *IMDM 2011, 2. Dezember 2011, Mainz*. In: Lehner, W., Piller, G. (Hrsg.) *Innovative Unternehmensanwendungen mit In-Memory-Data-Management. Beiträge der Tagung IMDM 2011, 2.12.2011 in Mainz*, S. 23–34. Bonn (2011)

Jens Lechtenbörger und Gottfried Vossen

Inhaltsverzeichnis

10.1 Einleitung und motivierende Beispiele. 206

10.2 NoSQL- und NewSQL-Datenbanken 207

 10.2.1 Grundlagen: Partitionierung, Replikation, CAP-Theorem, Eventual Consistency. 208

 10.2.2 NoSQL 211

 10.2.3 NewSQL 212

10.3 Big Data und Map-Reduce/Hadoop 213

 10.3.1 Technologie zur Handhabung von Big Data: Map-Reduce 214

 10.3.2 Nutzung von Big Data 219

10.4 Zusammenfassung und Ausblick 221

Literatur 222

Zusammenfassung

Traditionelle relationale Datenbanken haben in der jüngeren Vergangenheit Konkurrenz in Form von NoSQL- und NewSQL-Datenbanken sowie von parallelen Datenhaltungs- und -analysesystemen wie Hadoop erhalten. Die zugrundeliegenden Entwicklungen werden motiviert, technische Grundlagen erläutert sowie Besonderheiten der Ansätze vorgestellt.

J. Lechtenbörger (✉) · G. Vossen
Institut für Wirtschaftsinformatik, Westfälische Wilhelms-Universität, Münster, Deutschland
E-Mail: lechten@wi.uni-muenster.de

G. Vossen
E-Mail: vossen@uni-muenster.de

10.1 Einleitung und motivierende Beispiele

Nachdem relationale Datenbanken lange Zeit weitgehend konkurrenzlos die Grundlage betrieblicher Informationssysteme gebildet hatten, kam unter den Schlagworten *NoSQL* (Flexibilität, Open Source, schwache Konsistenz, „Web-scale“), später *Hadoop* (verteilte, leistungsfähige, ausfallsichere Datenanalyse, insbesondere für Big Data) sowie zuletzt *NewSQL* (relationale Datenbanken mit klassischen Konsistenzgarantien, aber hochgradiger Skalierbarkeit) wieder Bewegung in das Datenbankumfeld.

Zwei Beispiele mögen diese Entwicklungen motivieren. Zum einen betrachte man ein weltweit operierendes soziales Netzwerk, das als Web-Anwendung die Kommunikation zwischen Mitgliedern durch Nachrichten und „Likes“ ermöglicht. Derartige Anwendungen weisen unter anderem folgende Charakteristika auf: Sie generieren ein hohes Datenaufkommen in parallelen Abläufen, was eine verteilte Datenhaltung mit (zumindest teilweise) replizierten Daten sinnvoll erscheinen lässt. Mitglieder erwarten höchste Verfügbarkeit, haben aber geringe Erwartungen an Konsistenz: Ob verschiedene, geographisch weit entfernte Mitglieder vorübergehend unterschiedliche Like-Zähler oder eine unterschiedliche Sicht auf Kommentare wahrnehmen, spielt keine entscheidende Rolle. Zudem basiert die Kommunikation zwischen Client-Browser und Web-Server oft auf JSON, einem leichtgewichtigen Standard zur Darstellung semi-strukturierter Information in Attribut-Wert-Paaren. Angesichts dieser Charakteristika sind (objekt-) relationale Datenbankmanagementsysteme ((O)RDBMS) mit ACID-Transaktionen (Weikum and Vossen 2002) und starrem Schema nicht das optimale Werkzeug für eine verteilte und skalierbare Datenhaltung im Web-Umfeld, was unter dem Schlagwort *NoSQL* zur Entwicklung zahlreicher alternativer Datenbankformen geführt hat, wie in Abschn. 10.2 thematisiert wird.

Zum anderen gewinnt die Verarbeitung von Massendaten zunehmend an Bedeutung, beispielsweise bei der Auswertung von Datenströmen, die mittlerweile allgegenwärtig produziert und ausgewertet werden, etwa in Industrie, Handel und Logistik, für vielfältige Formen des wissenschaftlichen Rechnens (Klimaforschung, Astronomie, Kernphysik, Genetik), im Web-Umfeld (Klickstream-Analysen, Werbeeinblendungen) oder im Privatleben („measured self“, Home-Automation). Der Unterschied zwischen einem Datenbanksystem und einem Datenstromsystem ist in Abb. 10.1 für Anfragen dargestellt: Eine Datenbank-Anfrage kann ohne Vorankündigung an eine Datenbank geschickt werden; sie wird individuell verarbeitet und erzeugt entsprechende Ergebnisse (Abb. 10.1a), denn nach einem Laden sind die Daten, auf denen die Anfrage ausgewertet wird, dauerhaft vorhanden. In einem Datenstromsystem wird dagegen eine Anfrage, die zunächst im System registriert wurde, kontinuierlich am vorhandenen Datenstrom ausgewertet; die Verarbeitungskomponente produziert Ergebnisse anhand des Teils des Datenstroms, der im aktuellen Fenster betrachtet werden kann (Abb. 10.1b). Abbildung 10.1 ist insofern unvollständig, als ein Stromprozessor häufig ebenfalls auf lokalen (Puffer-) Speicher zugreifen kann oder sogar Teil eines klassischen Datenbanksystems ist.

Oftmals bestehen die Inhalte von Datenströmen aus Schlüssel-Wert-Paaren, wobei die Werte unstrukturiert sind und von entsprechender Anwendungslogik interpretiert und ana-

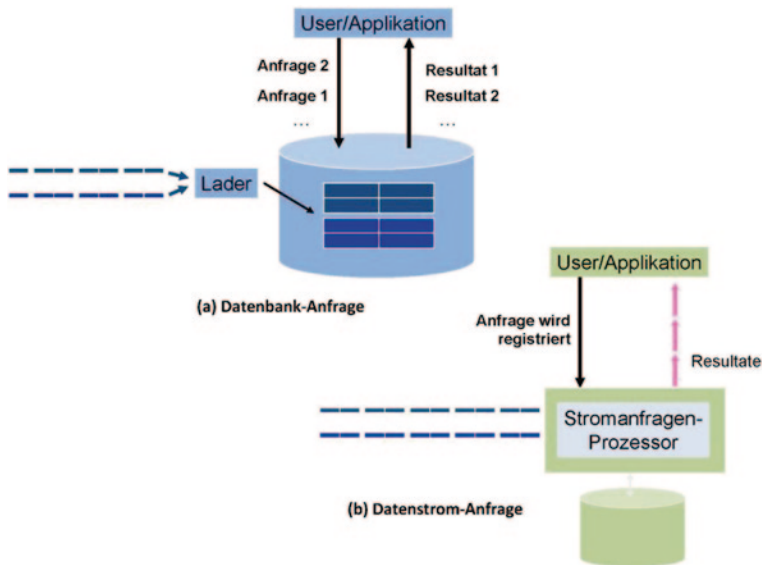


Abb. 10.1 Datenbank-Anfrage vs. Datenstrom-Anfrage

lysiert werden müssen, was sich in der Regel hochgradig parallelisieren lässt. Auch hier drängen (O)RDBMS sich nicht als Mittel der Wahl auf. Stattdessen ist mit der Rückbesinnung auf (verteilte) Dateisysteme und Map-Reduce ein neues Paradigma zum verteilten, skalierbaren und fehlertoleranten Umgang mit massiven Datenbeständen entwickelt worden, auf das in Abschn. 10.3 eingegangen wird.

10.2 NoSQL- und NewSQL-Datenbanken

Seit einigen Jahren gibt es vielfältige Datenbank-Aktivitäten und -Produkte unter dem Schlagwort „NoSQL“. Dennoch (oder gerade deswegen) gestaltet sich eine genaue Abgrenzung der zugehörigen Bestrebungen schwierig. Zunächst ist „NoSQL“ der Name einer seit Ende der 1990er Jahre weiterentwickelten relationalen Open-Source-Datenbank¹, die keinen Zugriff über die standardisierte Datenbanksprache SQL anbietet.

Nach heutigem Verständnis gehört diese Datenbank allerdings *nicht* der NoSQL-Bewegung an, denn das „No“ wird mehrheitlich nicht mit „nein“, sondern mit „not only“ interpretiert. Die Bewegung wurde insbesondere durch zwei Konferenzen im Jahre 2009 geprägt, das NOSQL Meeting² im Juni 2009 und die no:sql(east)³ im Oktober 2009.

¹ http://www.strozzi.it/cgi-bin/CSA/tw7/I/en_US/nosql/Home%20Page

² <http://nosql.eventbrite.com/>

³ <https://nosqleast.com/2009>

Die erste dieser Konferenzen hatte zum Ziel, Beschränkungen gängiger relationaler Datenbanken und deren Überwindung mit Hilfe verteilter, nicht-relationaler Open-Source-Datenbanken zu diskutieren. Ausgangspunkt dieser Fokussierung war die Tatsache, dass im Web-Umfeld große Datenbestände aus Performanz-, Flexibilitäts- und Skalierbarkeitsgründen häufig *nicht* mit relationalen Datenbanken verwaltet und ausgewertet wurden, sondern durch alternative Systeme, die auf SQL und starre Datenbankschemata oder ACID-Transaktionen verzichten. Einflussreich waren hier insbesondere die Kernideen von Bigtable (Chang et al. 2008) und Map-Reduce (Dean and Ghemawat 2008) (Datenhaltung und -auswertung bei Google) sowie von Dynamo (DeCandia et al. 2007) (Datenhaltung bei Amazon), die in einer steigenden Anzahl von Open-Source-Projekten aufgegriffen und in populären Web-Auftritten eingesetzt wurden. In Abschn. 10.2.1 werden zunächst Grundlagen der verteilten Datenhaltung und insbesondere den für viele NoSQL-Systeme charakteristischen Begriff der *Eventual Consistency* eingeführt, um in Abschn. 10.2.2 genauer auf diese Systeme einzugehen.

Parallel zu den NoSQL-Entwicklungen gab es allerdings bereits frühzeitig kritische Stimmen, die vor dem Verzicht auf SQL-Standardisierung und auf ACID-Transaktionen warnten. Entsprechend wurden und werden SQL-Datenbanken, die weiterhin auf ACID-Transaktionen setzen, aber die Skalierbarkeit von NoSQL-Systemen erreichen sollen, unter dem Schlagwort *NewSQL* beworben; derartige Systeme sind Gegenstand von Abschn. 10.2.3.

10.2.1 Grundlagen: Partitionierung, Replikation, CAP-Theorem, Eventual Consistency

Partitionierung und Replikation sind gängige Techniken zur Gewährleistung von Skalierbarkeit und Fehlertoleranz in verteilten Systemen (etwa im Cloud Computing und im Datenbankumfeld).

Verteilte Anwendungen verlassen sich hinsichtlich der notwendigen Rechenleistung typischerweise auf große Ansammlungen gebräuchlicher Hardware, bestehend aus konventionellen Prozessoren („Rechenknoten“), die zu Clustern zusammengefasst und über geeignete Netzwerktechnologie miteinander verbunden sind; derartige Cluster werden häufig innerhalb eines Datacenters oder über mehrere Datacenter hinweg *repliziert*. Replikation, also das gezielte Kopieren identischer Daten auf mehrere Server, als Form von Redundanz ist in diesem Zusammenhang der Schlüssel zur Zuverlässigkeit des betreffenden Systems und des fehlertoleranten Rechnens; analog werden Daten durch Replikation gegen Verlust geschützt. Das Ergebnis ist dann entweder ein verteiltes File-System (wie das Hadoop Distributed File System, kurz HDFS, s. Abschn. 10.3) oder eine global verteilte Datenbank (wie Google Spanner, s. Abschn. 10.2.3).

Neben Fehlertoleranz und (hoher) Verfügbarkeit ermöglicht Verteilung eine parallele Verarbeitung von Daten, speziell dann, wenn Rechenvorgänge unabhängig voneinander auf unterschiedlichen Teilen der Daten ausgeführt werden können. In einem solchen Fall

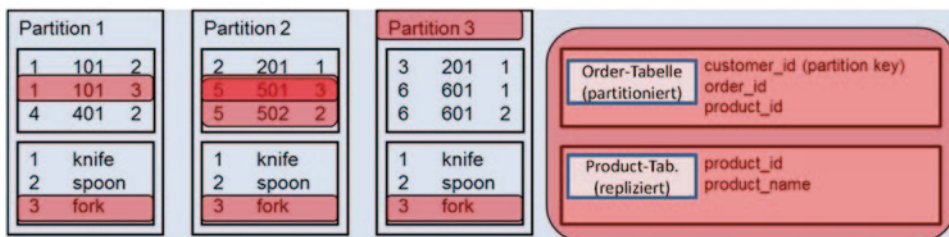


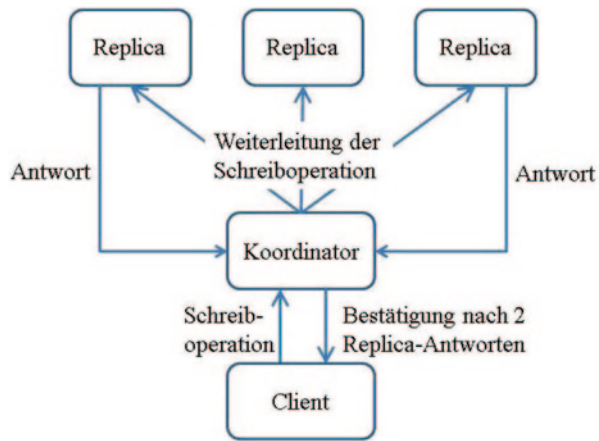
Abb. 10.2 Partitionierung vs. Replikation

werden Daten häufig über mehrere Cluster oder sogar Datacenter hinweg *partitioniert*; Abb. 10.2 illustriert den Unterschied zwischen Partitionierung und Replikation. In dem in dieser Abbildung gezeigten Beispiel sind diejenigen Daten aus einer relationalen Datenbank, die Kunden-Order (Bestellungen) betreffen, über drei Speicherorte partitioniert so, dass jedem Ort eindeutige Kundennummern zugeordnet sind. Die Product-Tabelle ist dagegen an denselben Orten repliziert (d. h. identisch kopiert). Anfragen und Updates können damit an eine oder gleichzeitig an mehrere Partitionen gerichtet werden, wobei Replikation und Partitionierung für Benutzer transparent sind.

Es sollte unmittelbar einleuchten, dass schreibende Zugriffe auf replizierten Daten besonderer Vorkehrungen bedürfen, um die Konsistenz der Daten sicherzustellen. Während klassische relationale Datenbanken mit ACID-Transaktionen die Serialisierbarkeit als Konsistenzbegriff verfolgen, gilt im Kontext verteilter Systeme die Ein-Kopien-Konsistenz (engl. *single copy consistency*), also die Illusion der Abwesenheit von Replikation, als starke Konsistenz, und es gibt eine ganze Reihe schwächerer Formen von Konsistenz wie *Eventual Consistency*, *Read-Your-Writes*, *Monotonic Reads*, *Monotonic Writes* oder kausale Konsistenz (Vogels 2009).

Zur Durchsetzung von starker Konsistenz auf replizierten Daten kommen oftmals Konsens-Protokolle wie Paxos (Lamport 1998) zum Einsatz, etwa bei Google in Bigtable (Chang et al. 2008) und Spanner (Corbett et al. 2013). *Eventual Consistency* beruht wie im Falle von Amazon Dynamo (DeCandia et al. 2007) typischerweise auf Quorum-Protokollen, unter denen Lese- und Schreiboperationen auf mehreren (möglicherweise allen) Kopien ausgeführt werden müssen, bevor sie als erfolgreich durchgeführt gelten: Bei Dynamo werden Daten auf N Knoten repliziert, und es gibt zwei konfigurierbare Parameter, R und W . R spezifiziert die minimale Anzahl von Knoten, die für eine erfolgreiche Lese-Operation erforderlich sind, während W diese Anzahl für Schreib-Operationen vorgibt. Im Falle von $R + W > N$ liegt ein klassisches Quorum-System vor, in dem sich Lese- und Schreiboperationen immer überlappen und wo daher starke Konsistenz garantiert werden kann. Da jedoch viele Knoten beteiligt sind, ist eine hohe Latenz zu erwarten. *Eventual Consistency* liegt vor, falls $R + W \leq N$ gilt; hier ist die Latenz reduziert, allerdings können Leseoperationen auf Knoten ausgeführt werden, die noch nicht alle Schreiboperationen nachvollzogen haben, was dann in veralteten Werten resultiert. In Abb. 10.3 wird (Bailis et al. 2014) folgend eine Schreiboperation für den Fall $N=3$ und $W=2$ illustriert: Die Schreib-operation wird erfolgreich abgeschlossen, wenn zwei der drei Kopien die

Abb. 10.3 Schreiboperation in Quorum-System mit $N=3$ und $W=2$



Schreib-operation erfolgreich durchgeführt haben. Im Falle $R=1$ könnten nachfolgende Leseoperationen dann den Wert der noch nicht aktualisierten Kopie liefern.

Schwächere Formen von Konsistenz (wie Eventual Consistency) sind vor allem aus zwei Gründen interessant. Erstens können sie Latenzen vermeiden, die aus Synchronisationsmechanismen resultieren. Zweitens besagt das *CAP-Prinzip* (Fox and Brewer 1999), das später in (Lynch and Gilbert 2002) als *CAP-Theorem* bewiesen wurde, dass man von den drei Eigenschaften Konsistenz (engl. *Consistency*), Verfügbarkeit (engl. *Availability*) und Partitionstoleranz (engl. *Partition tolerance*) höchstens zwei gleichzeitig sicherstellen kann. Partitionstoleranz liegt vor, wenn Anwendungen mit der Situation umgehen können, dass die zugrunde liegenden Server aufgrund von Netzwerkproblemen in verschiedene Partitionen zerfallen, zwischen denen zeitweise keine Kommunikation möglich ist. Für hochgradig (weltweit) verteilte Anwendungen lassen sich Netzwerkpartitionierungen in der Regel nicht ausschließen. Wenn dann noch hohe Verfügbarkeit garantiert werden soll, müssen laut CAP-Theorem also Kompromisse bezüglich der Konsistenz eingegangen werden.

Entsprechend verfolgen viele NoSQL-Datenbanken nicht mehr das Ziel von starker Konsistenz, sondern versprechen „irgendwann konsistent“ (engl. *eventually consistent*) zu werden: Wenn Partitionierungen behoben sind und hinreichend lange keine neuen Aktualisierungen vorgekommen sind, ist sichergestellt, dass irgendwann sämtliche Updates so durch das System propagiert wurden, dass alle Knoten wieder konsistent sind. NoSQL-Datenbanken greifen damit das *BASE-Prinzip* auf (anstelle von ACID), das im Kontext verteilter Internet-Anwendungen bereits lange bekannt ist (Fox et al. 1997): *Basically Available, Soft State, Eventually Consistent*.

Bemerkenswert ist allerdings, dass das CAP-Theorem *keine* Rechtfertigung liefert, auf starke Konsistenz generell zu verzichten: Wie Brewer in (2012) erläutert, ist der Verzicht auf Konsistenz nur während der Dauer von Partitionierungen notwendig. Solange keine Partitionierung auftritt, was dem Regelbetrieb entspricht, können auch hochverfügbare Systeme starke Konsistenz garantieren.

10.2.2 NoSQL

Wie bereits erwähnt subsumiert „NoSQL“ verschiedene Formen der Datenhaltung, die auf die eine oder andere Art von „klassischen“ SQL-DBMS abweichen, wobei in der Regel Open Source, horizontale Skalierbarkeit und hohe Verfügbarkeit bei Vernachlässigung von Konsistenz erwartet werden. Die Bandbreite reicht von *Key-Value Stores* (z. B. Amazon SimpleDB und Dynamo, LinkedIn Voldemort, Riak, Redis) und *Column Stores* (z. B. Google Bigtable (Chang et al. 2008), Apache Hbase oder Cassandra, Yahoo! PNUTS) über *Dokument-Datenbanken* (z. B. MongoDB oder Couchbase) bis zu *Graphen-Datenbanken* (z. B. Neo4J der Allegro) (Redmond and Wilson 2012). Nachfolgend werden einige Besonderheiten von Column Stores am Beispiel Bigtable sowie von Key-Value Stores am Beispiel Dynamo vorgestellt⁴.

Das Datenmodell von Bigtable (Chang et al. 2008) beruht auf Tabellen, denen Spalten zugewiesen werden, die ihrerseits aus verwaltungstechnischen Gründen in sogenannte *Spaltenfamilien* gruppiert werden. Während eine Tabelle typischerweise wenige Spaltenfamilien umfasst, kann die Anzahl der Spalten unbeschränkt sein. Eine Zeile einer Tab. (also ein Datensatz) darf Werte für eine beliebige Auswahl von Spalten annehmen. Werte gehören einem einheitlichen Datentyp „String“ an und können im Zeitverlauf versioniert werden. Konzeptionell kann eine Tabelle somit als Abbildung von Tripeln der Form (Zeilen-ID, Spalten-ID, Zeitpunkt) auf beliebige Zeichenketten verstanden werden. Lese- und Schreibzugriffe auf einzelne Zeilen sind atomar, und es werden Transaktionen auf einzelnen Zeilen unterstützt (jedoch nicht zeilenübergreifend).

Technisch setzt Bigtable auf dem Google File System (GFS) auf und läuft in Server-Clustern, wobei einzelne Tabellen dynamisch in Zeilenbereiche, sogenannte *Tablets*, partitioniert und auf unterschiedliche Tablet-Server verteilt werden. Die Daten werden mit Hilfe des Paxos-Algorithmus konsistent repliziert, und neue Tablet-Server können zur Laufzeit hinzugefügt werden, wodurch Skalierbarkeit und Hochverfügbarkeit sichergestellt werden.

Der bekannteste Bigtable-Klon im Open-Source-Umfeld ist das in Java programmierte HBase⁵, das aufbauend auf dem Hadoop Distributed File System (HDFS) Bigtable-ähnliche Funktionalität bereitstellt und beispielsweise bei Adobe und Twitter eingesetzt wird.

Im Gegensatz zu dieser tabellenorientierten Speicherung verwaltet Amazon Dynamo (DeCandia et al. 2007) als Key-Value Store Schlüssel-Wert-Paare. Zu einem Schlüsselwert kann ein beliebiger Wert als Blob (Binary Large Object) per PUT-Operation gespeichert und per GET-Operation erfragt werden. Es gibt weder Operationen, die mehrere Schlüssel betreffen (also auch keine Transaktionen) noch ein relationales Schema. Die Besonderheit von Dynamo ist, dass PUT-Operationen unabhängig von Netzwerk- und Serverausfällen jederzeit möglich sein sollen (always writable; z. B. sollen Kunden jederzeit neue Artikel in ihre Warenkörbe legen können, auch wenn der Server abstürzt, mit dem ein Browser bisher kommuniziert hat). Dies bedeutet insbesondere, dass im Falle von

⁴ <http://nosql-database.org/>

⁵ <http://wiki.apache.org/hadoop/Hbase>

Netzwerkpartitionierungen inkonsistente Wert-Versionen zu einem Schlüssel entstehen können. Entsprechend bietet Dynamo lediglich Eventual Consistency als Konsistenzgarantie. Technisch nutzt Dynamo Partitionierung und Replikation der Daten basierend auf verteilten Hash-Tabellen (Distributed Hash Tables, DHT), wobei Vektorkuhren eingesetzt werden, um unterschiedliche Wert-Versionen zu einem Schlüssel erkennen und verwalten zu können.

Viel zitierte Open-Source-Projekte, die ebenfalls auf verteilten Hash-Tabellen und Eventual Consistency aufbauen, sind Project Voldemort⁶, das bei LinkedIn eingesetzt wird, das ursprünglich bei Facebook entwickelte Apache Cassandra⁷ und Basho Riak⁸.

10.2.3 NewSQL

Die NoSQL-Entwicklungen, insbesondere der Fokus auf Eventual Consistency, wurde bereits frühzeitig von kritischen Kommentaren begleitet. Beispielsweise hat Michael Stonebraker schon im Jahre 2010 in einem Blogeintrag der Communications of the ACM⁹ Gründe angegeben, warum Unternehmen NoSQL-Systeme entweder nicht kennen würden oder nicht interessiert seien; seine Kern-Statements lauteten damals bereits (1) „No ACID Equals No Interest“, da für viele Unternehmen OLTP das Kerngeschäft ausmacht und dieses nicht ohne harte Konsistenzgarantien auskommt, (2) „A Low-Level Query Language is Death“, da Anwendungsprogrammierer nicht auf Block- oder Record-Ebene arbeiten wollen, und (3) „NoSQL Means No Standards“, was für viele Unternehmen schon deshalb von Bedeutung ist, weil sie viele Datenbanken betreiben und einheitliche Schnittstellen zwischen diesen benötigen.

In einem weiteren Blogeintrag¹⁰ benutzte Stonebraker 2011 den Begriff „New SQL“ für die wieder aufkommenden SQL-Datenbanken mit ACID-Konsistenz, die versprachen, den NoSQL-Datenbanken gleichwertige Skalierbarkeitseigenschaften aufzuweisen (etwa Clustrix, NimbusDB/NuoDB, VoltDB). Derartige Datenbanken zielen also auf den klassischen Datenbankmarkt, allerdings mit besonderem Fokus auf extremer Skalierbarkeit, etwa im Big-Data-Umfeld (vgl. Abschn. 10.3).

Eine etwas andere Art von Datenbank, die ebenfalls der Kategorie NewSQL zugerechnet wird, ist die F1-Datenbank (Shute et al. 2013) von Google. Selbst in Unternehmen wie Google, die ihr Geld auf völlig andere Weise als „Brick-and-Mortar“-Firmen verdienen, gibt es viele Anwendungen, die z. B. strenge Konsistenz in Gegenwart von weit verteilter Replikation erfordern, die sich ein dem Relationenmodell zumindest ähnliches Datenmo-

⁶ <http://project-voldemort.com/>

⁷ <http://cassandra.apache.org/>

⁸ <http://www.basho.com/riak/>

⁹ <http://cacm.acm.org/blogs/blog-cacm/99512-why-enterprises-are-uninterested-in-nosql>

¹⁰ <http://cacm.acm.org/blogs/blog-cacm/109710-new-sql-an-alternative-to-nosql-and-old-sql-for-new-oltp-apps/fulltext>

dell wünschen und die ferner auf Versionierung setzen. Bei Google trifft dies z. B. auf das Anzeigengeschäft zu; das Unternehmen hat darauf mit der Entwicklung von Spanner (Corbett et al. 2013) reagiert, einer hoch skalierbaren, global verteilten Mehrversionen-Datenbank mit synchroner, anwendungsbezogen konfigurierbarer Replikation. Spanner garantiert Hochverfügbarkeit durch Replikation über Data Center oder sogar über Kontinente hinweg. Erster Kunde für Spanner ist Googles F1-Datenbank (Shute et al. 2013), das Backend von Google AdWords mit SQL-Unterstützung, verteilten SQL-Anfragemöglichkeiten, optimistischer Concurrency Control sowie Möglichkeiten zur Datenbank-Reorganisation und dynamischen Schema-Veränderungen.

Es stellt sich heraus, dass selbst Web-Applikationen wie das AdWords-System (welches wirtschaftliche Daten verwaltet und daher harten Anforderungen an Daten-Integrität und Daten-Konsistenz genügen muss) nicht ohne klassische ACID-Transaktionsgarantien auskommen, während Eventual Consistency von NoSQL-Systemen an dieser Stelle nicht ausreichend ist. F1-Transaktionen bestehen aus vielen Lese- und optional einer Schreib-Operation, welche die Transaktion beendet. Es werden Snapshot-, pessimistische und optimistische Transaktionen unterschieden und vom Scheduler unterschiedlich behandelt, wobei letzterer Typ die Voreinstellung ist. Optimistische Transaktionen können lange laufen, bestimmte Client-Fehler ausgleichen sowie Server-Failovers mitmachen.

Da F1 die synchrone Replikation von Spanner mit dem Ziel höchster Verfügbarkeit über Rechenzentrumsgrenzen hinweg einsetzt (laut (Shute et al. 2013) mit Kopien in fünf Rechenzentren, je zwei an der US-Ost- und -Westküste, eines im Landesinneren), sind die Commit-Latenzen von schreibenden Transaktionen mit 50–150 ms deutlich höher als bei den von Stonebraker genannten NewSQL-Datenbanken, was bei Google durch spezielle Schema-Cluster-Techniken und Programmierstrategien für Anwendungen ausgeglichen wird.

Die New SQL-Entwicklung und insbesondere Google zeigen, dass durch eine angemessene Systemarchitektur sowie eine geschickte Kombination von Datenbank-Technologien auch global verfügbare Informationssysteme gebaut werden können, die höchsten Anforderungen an Ausfallsicherheit, Skalierbarkeit, Konsistenz und Usability genügen können. Es ist davon auszugehen, dass andere Hersteller dieser Entwicklungsrichtung folgen werden, so dass nicht unbedingt von einem „Back to the Roots“ bei relationalen Datenbanken gesprochen werden kann, aber von einer klaren und zeitgemäßen Evolution, die mit ACID-Transaktionen, Skalierbarkeit und Hochverfügbarkeit bei globaler Verteilung die Stärken zweier paralleler Entwicklungslinien zusammenführt.

10.3 Big Data und Map-Reduce/Hadoop

Neben der fortschreitenden Digitalisierung und der Tatsache, dass digitale Objekte (z. B. Email-Anhänge oder Digitalbilder sowie -videos) im Laufe der Zeit immer größer geworden sind, hat die technologische Entwicklung immer schnelleren Datentransport sowie umfassendere Datenspeicherung ermöglicht und die Erzeugung von Daten sowohl

automatisch, also durch Maschinen, als auch durch User dramatisch zunehmen lassen. In einer neueren Statistik¹¹ hat die Firma Intel festgestellt, dass in einer einzigen Internet-Minute rund 650 TB an IP-Daten über das Netz transportiert werden. Das Ergebnis dieser Entwicklung ist so überwältigend, dass die Bezeichnung „Big Data“ angemessen erscheint; dieser Abschnitt beschreibt die Herausforderungen, Techniken und Anwendungen von Big Data und wirft dabei einen Blick auf mögliche zukünftige BI-Architekturen.

Das *Volumen* (engl. Volume) stellt bereits ein erstes Charakteristikum von Big Data dar. Daten werden laut (Cavage and Pacheco 2014) als „big“ betrachtet, wenn sie groß genug sind, um signifikant von paralleler Verarbeitung in einer Armada von Rechnern zu profitieren, wobei die Verwaltung der Berechnung selbst eine beachtliche Herausforderung darstellt. Weitere Charakteristika von Big Data, die zusammen mit dem Volumen als die „fünf Vs“ bezeichnet werden, sind die (hohe) *Geschwindigkeit* (engl. Velocity), mit der Daten produziert werden und verarbeitet werden müssen, die (hohe) *Vielfalt* (engl. Variety), die Daten annehmen können, die (nicht immer gegebene) Präzision, Genauigkeit oder *Vertrauenswürdigkeit* (engl. Veracity), in der Daten vorliegen, und der *Wert* (engl. Value), den Daten (hoffentlich) darstellen. (Die ersten drei dieser Charakteristika werden dem Analysten Doug Laney¹² von Gartner zugeschrieben.) Geschwindigkeit bezieht sich auf die Tatsache, dass Daten heute oft in Form der bereits erwähnten Datenströme auftreten, welche der die Daten analysierenden Einrichtung keine Chance lassen, die Daten zu Analysezwecken dauerhaft zu speichern; stattdessen muss die Verarbeitung unmittelbar erfolgen. Vielfalt bedeutet, dass die Daten in unterschiedlichsten Formen und Formaten auftreten, darunter als unstrukturierter Text, als semi-strukturierte Daten (z. B. XML-Dokumente) oder als strukturierte Daten (z. B. relationale Tabelle). Diese Eigenschaften von Big Data sind in Abb. 10.4 zusammengefasst. Wert ist eine Eigenschaft, die Daten erst im Nachhinein, also durch angemessene Analyse, Verdichtung oder sonstige Verarbeitung erhalten.

10.3.1 Technologie zur Handhabung von Big Data: Map-Reduce

Zum Umgang mit Big Data sind in den letzten Jahren zahlreiche Techniken, Methoden und Technologien entwickelt worden, über die als Nächstes ein kurzer Überblick gegeben wird. Wenn Daten in so hohen Quantitäten auftreten, dass lokal vorhandene Speicher- und Verarbeitungskapazitäten nicht mehr ausreichen, kann es nicht verwundern, dass „traditionelle“ Technologie, die auf einer zentralen Datenbank fußt, nicht mehr ausreicht. Um zu verstehen, was stattdessen benötigt wird bzw. angemessen ist, sollen zunächst Anforderungen an die Verarbeitung von Big Data zusammengestellt und dann Technologien betrachten werden, die diesen genügen. Diese Anforderungen lassen sich wie folgt zusammenfassen:

¹¹ <http://www.intel.com/content/www/us/en/communications/internet-minute-infographic.html>

¹² <http://blogs.gartner.com/doug-laney/>

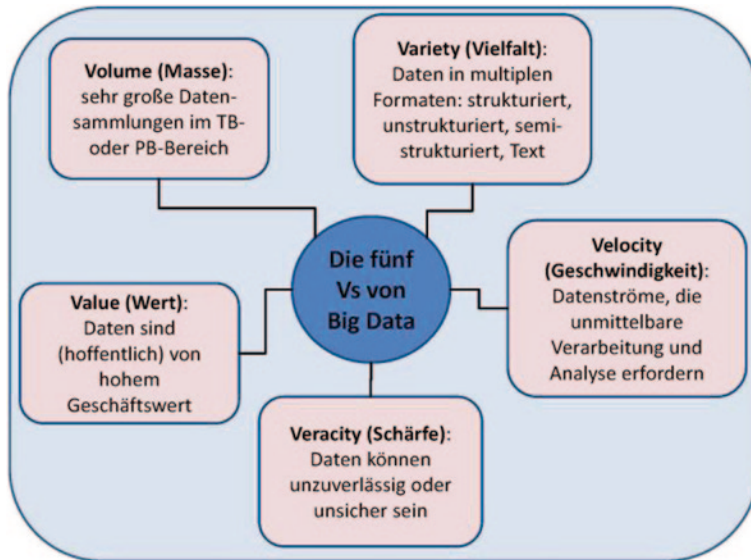


Abb. 10.4 Die definierenden „Fünf Vs“ von Big Data

- Angemessene, hohe Verarbeitungsleistung für komplexe Rechnungen;
- skalierbare, verteilte und fehlertolerante Datenverarbeitung mit temporärem oder dauerhaftem Speicher;
- parallele Programmier- und Verarbeitungsparadigmen, die mit großen Datenmengen angemessen umgehen können;
- angemessene Implementierungen und Ausführungsumgebungen für diese Programmiermodelle und Paradigmen.

Bzgl. Hardware-Lösungen zur Verarbeitung von Big Data sei der Leser auf (Saecker and Markl 2012) verwiesen. Ebenfalls relevant in diesem Zusammenhang ist ein Wiederaufkommen von Hauptspeicher- bzw. In-Memory-Datenbanktechnologie, eine Entwicklung, die ursprünglich aus den 1980er Jahren stammt (Eich 1989) und die mittlerweile – dank enormer technologischer Fortschritte in den letzten 30 Jahren – in kommerziellen Produkten verfügbar ist (Loos et al. 2011; Plattner and Zeier 2011). Bei Datenbanken sind ferner NoSQL- und NewSQL-Systeme entstanden, die wir im vorigen Abschnitt skizziert haben.

Während Replikation – wie in Abschn. 10.2.1 beschrieben – als Maßnahme zum Erhalt von Datenverfügbarkeit gilt, erweist sich Partitionierung als Schlüssel zur algorithmischen Behandlung zahlreicher Big-Data-Probleme. Partitionierung folgt im Wesentlichen dem „alten“ Prinzip des *Teile und Herrsche* (engl. *Divide and Conquer*), das bereits von Caesar verwendet wurde und in Algorithmen der Informatik eine lange Tradition hat. Wenn Daten in mehrere unabhängige Teile zerlegt werden können (wie im Partitionierungsbeispiel von Abb. 10.2 oben), kann eine Verarbeitung dieser Teile parallel erfolgen, z. B. in verschiedenen Kernen einer Multicore-CPU, in unterschiedlichen CPUs eines Clusters oder sogar

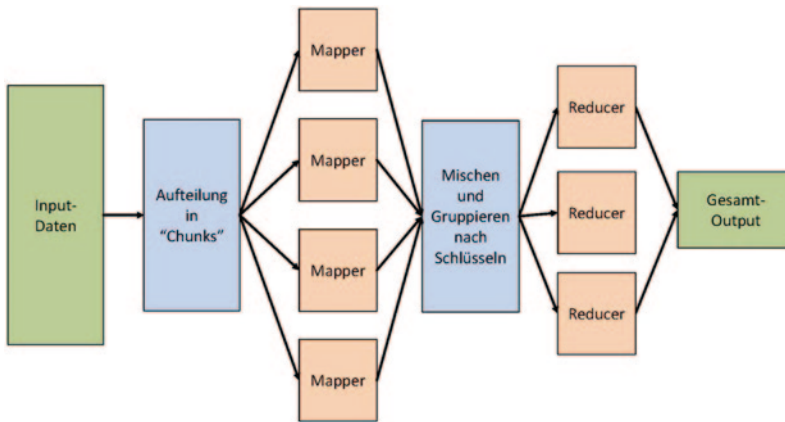


Abb. 10.5 Prinzip einer Map-Reduce-Berechnung

in verschiedenen Data Centern; die dort erzielten Einzelresultate müssen dann zu einem Gesamtergebnis zusammengesetzt werden. Dies ist bereits die grundlegende Idee von Googles Map-Reduce-Ansatz (Dean and Ghemawat 2008) (US Patent 7,650,331 von Januar 2010), der aus der funktionalen Programmierung bekannte Funktionen höherer Ordnung zur Spezifikation von Berechnungen auf sehr großen Datenmengen heranzieht.

Map-Reduce kombiniert zwei Funktionen, *map* und *reduce*, die auf Schlüssel-Werte-Paare angewendet werden und die im Grundsatz dem Group-By von SQL gefolgt von einer Aggregation ähneln. Eine Map-Reduce-Berechnung arbeitet grundsätzlich wie in Abb. 10.5 gezeigt: Eingabedaten werden in sog. „Chunks“ (Partitionen) bereitgestellt und stammen typischerweise aus einem verteilten Dateisystem. Die Chunks werden zunächst von Map-Tasks auf *Mapper* genannten Komponenten verarbeitet. Mapper konvertieren Chunks in eine Folge von Schlüssel-Wert-Paaren; wie genau dies passiert, hängt von der spezifischen Aufgabe ab und wird durch den Code bestimmt, welchen der Benutzer für die Map-Funktion bereitstellen muss. Als nächstes werden die von den Mappern erzeugten Zwischenergebnisse durch einen Master-Controller gesammelt und anhand ihrer Schlüssel gruppiert. Die Schlüssel und die ihnen zugeordneten Gruppen werden dann so an Reduce-Tasks weitergeleitet, dass alle Schlüssel-Wert-Paare mit demselben Schlüssel derselben Reducer-Komponente zugewiesen werden. Die Reducer bearbeiten jeweils einen Schlüssel und kombinieren alle mit diesem assoziierten Werte in der Weise, wie die Anwendung es erfordert und der zugehörige Code spezifiziert.

Am Beispiel aus (White 2012) lassen sich Wetterdaten analysieren, die als langer String von Wetterstationen geliefert werden; interessant ist hier die Höchsttemperatur pro Jahr. Input-Daten können in diesem Fall wie in Abb. 10.6 gezeigt aussehen. Die betreffende Wetterstation sendet (als Datenstrom!) regelmäßig lange Zeichenreihen, die entsprechend interpretiert werden müssen; jeder solche String enthält u. a. die Stations-ID, das Datum der Messung, Längen- sowie Breitengrad des Orts der Station und die aktuelle Temperatur.

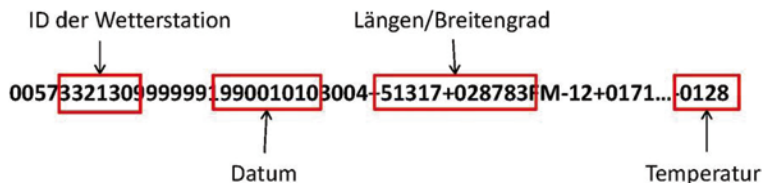


Abb. 10.6 Exemplarische Wetterdaten

Angenommen, der folgende Input liege vor; die für die Bestimmung der maximalen Jahrestemperatur relevanten Teile sind kenntlich gemacht (und Temperaturwerte auf ganze Zahlen gerundet):

```
006701199099999991990051507004+51317+028783FM-12+0171...+0000
004301199099999991990051512004+51317+028783FM-12+0171...+0022
004301199099999991990051518004+51317+028783FM-12+0171...-0011
004301199099999991989032412004+51317+028783FM-12+0171...+0111
004301199099999991989032418004+51317+028783FM-12+0171...+0078
```

Falls dies als ein Chunk an einen Mapper geleitet wird, extrahiert dieser das Jahr (als Schlüssel) und die Temperatur (als Wert) wie gewünscht und gibt aus: (1990, 0), (1990, 22), (1990, -11), (1989, 111), (1989, 78). Gruppierung anhand der Schlüsselwerte liefert (1989, [111, 78]), (1990, [0, 22, -11]), woraus ein Reducer nun die Maxima bestimmen kann als (1989, 111) und (1990, 22).

Es sollte einleuchten, dass eine Aufgabe wie diese, die in der Realität sehr große Datenmengen von Wetterstationen umfassen wird, die aber alle unabhängig voneinander verarbeitet werden können, ein perfekter Kandidat für eine Map-Reduce-Verarbeitung ist. Andere Aufgaben dieser Art sind z. B. das Zählen des Vorkommens bestimmter Wörter in einer Textsammlung (was z. B. für Indexerzeugung und -wartung bei einer Suchmaschine relevant ist), Matrix-Vektor-Multiplikation (relevant für eine Berechnung des PageRank zur Anordnung von Suchergebnissen) oder Operationen der Relationenalgebra in Datenbanken (wie Joins oder Aggregat-Operationen, die z. B. im Kontext von Anfrage-Optimierung relevant sind) (Rajaraman et al. 2013).

Damit eine solche Map-Reduce-Berechnung bzw. -Ausführung funktioniert, sind verschiedene Frage zu klären: Wie wird der Code für eine bestimmte Map-Reduce-Task erstellt? Wie last sich ein gegebenes Problem datenseitig so in Chunks zerlegen, dass diese parallel verarbeitet werden können? Wie lassen sich Tasks angemessen an Rechenknoten (die einen Mapper oder einen Reducer ausführen) zuweisen? Wie werden die Teile einer Berechnung, die auf verschiedenen Knoten stattfindet, synchronisiert? Wie last sich ein solches Szenario robust gegen Fehler bzw. Ausfälle machen?

Die erste dieser Fragen kann nur durch einen Benutzer beantwortet werden, der die benötigten Map- bzw. Reduce-Funktionen – typischerweise in einer Hochsprache wie Java – programmiert. Für das oben gezeigte Wetter-Beispiel findet sich der entsprechende

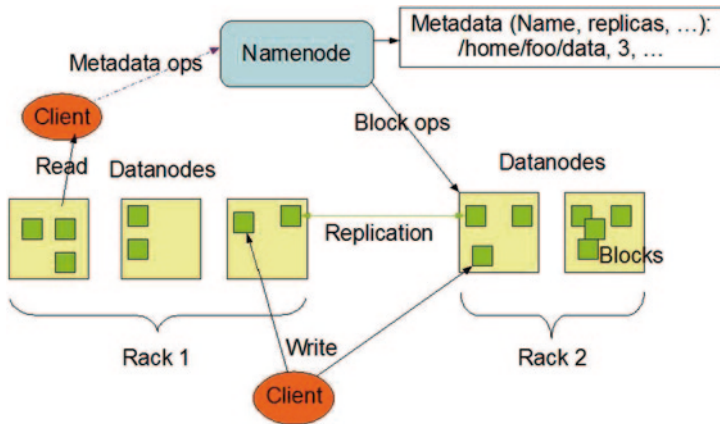


Abb. 10.7 Architektur von HDFS

Code beispielsweise in (White 2012). Die weiteren Fragen wurden in den letzten Jahren auf unterschiedliche Weisen beantwortet; die bekannteste Antwort ist die Software-Bibliothek Hadoop¹³, die bereits in Version 2 verfügbar ist. Hadoop (White 2012) unterstützt skalierbare Berechnungen, die auf verteilten Rechner-Clustern laufen. Die Kernkomponenten der ersten Hadoop-Version sind die *Map-Reduce Engine* sowie das *Hadoop Distributed File System* (HDFS). Erstere ist für Ausführung und Kontrolle von Map-Reduce-Jobs zuständig; HDFS ist ein verteiltes Dateisystem, in welchem große Datenmengen abgelegt, gelesen und ausgegeben werden können. Nutzerdaten werden in Blöcke unterteilt, die über die lokalen Speicher von Cluster-Knoten repliziert werden. HDFS basiert auf einer Master-Slave-Architektur, bei welcher ein *Namenode* als Master den Datei-Namensraum einschließlich der Datei-Block-Zuordnung und der Lokalisierung von Blöcken verwaltet und *Datanodes* als Slaves die eigentlichen Blöcke verwalten. Dies ist in Abb. 10.7 gezeigt, welche dem *HDFS Architecture Guide*¹⁴ entnommen ist. Neben diesen Hauptkomponenten gibt es zahlreiche Erweiterungen von Hadoop um spezifische Funktionalität, die zusammen das sog. *Hadoop Ecosystem* ergeben. Inzwischen wurden auch verschiedene Alternativen zu Hadoop vorgeschlagen (z. B. Disco, Skynet, Twister oder FileMap) sowie dessen Weiterentwicklung zu *Hadoop NextGen Map-Reduce* (YARN) bzw. Hadoop 2.0. Wer sich für Anwendungen bzw. Anwender von Hadoop interessiert, sei auf die Apache Website¹⁵ verwiesen.

Das Map-Reduce-Paradigma und seine Implementierung Hadoop hat in den letzten Jahren nicht nur zahlreiche Entwicklungen angestoßen, die zu zahlreichen neuen kommerziell verfügbaren Produkten geführt haben, sondern auch eine Reihe von Forschungsaktivitäten (Afrati et al. 2012, 2013; Fedak 2013; Rajaraman et al. 2013; Sauer and Härder

¹³ <http://hadoop.apache.org/>

¹⁴ http://hadoop.apache.org/docs/stable/hdfs_design.html

¹⁵ <http://wiki.apache.org/hadoop/PoweredBy>



Abb. 10.8 Mögliche Big Data-Adoptionsstrategie

2013; Shim 2013). Eine typische Erweiterung des grundlegenden Map-Reduce-Paradigmas ist PACT, ein in (Battre' et al. 2010) beschriebenes Programmiermodell, welches das Map-Reduce-Modell verallgemeinert durch Hinzunahme weiterer Funktionen sowie durch die Möglichkeit, Verhaltensgarantien für Funktionen zu spezifizieren.

10.3.2 Nutzung von Big Data

Als Nächstes soll ein Blick auf die Frage geworfen werden, wie ein Unternehmen Nutzen aus Big Data ziehen kann. Was ist zu tun bzw. zu verändern, wenn das Unternehmen bereits ein Data Warehouse für Datenanalysezwecke aufgesetzt hat? Dazu lässt sich hier eine Modifikation der „klassischen“ Data Warehouse-Architektur so skizzieren, dass Big Data-Anforderungen Rechnung getragen wird.

Wie bei vielen anderen Einführungsentscheidungen bzgl. IT auch, die man im Laufe der letzten 30 Jahre zu fällen hatte, macht es Sinn, eine Entscheidung, ob ein Big Data-Projekt aufgesetzt oder ob in Big Data-Technologie investiert werden soll, wohlfundiert zu treffen. Dabei kann zunächst ein Blick auf Anwendungsbereiche¹⁶ helfen, die möglicherweise zum eigenen vergleichbar sind, ferner z. B. eine SWOT-Analyse, mit welcher sich Stärken, Schwächen, Chancen und Risiken einer bestimmten Technologie oder eines Projektvorhabens analysieren lassen. Alternativ kann zur Entscheidungsfindung eine Kontextanalyse herangezogen werden, die nach Zielen und der Eigenschaft *Value* fragt, welcher von dem Projekt zu erwarten ist. Beide Werkzeuge sind populär und haben sich etwa im Zusammenhang mit Geschäftsprozessmodellierung bewährt (Schönthaler et al. 2012). Umfassender als spezifische Analysen ist die Entwicklung einer Big Data-Strategie, die z. B. wie in Abb. 10.8 gezeigt aussehen kann.

Die Strategie startet mit einer Phase der Informationsbeschaffung und Planung, was z. B. eine SWOT- oder eine Kontextanalyse (oder beides) umfassen kann und die Definition des Business Case enthalten sollte. Falls eine Entscheidung zugunsten eines Big Data-Projekts oder der Adoption von Big Data-Technologie gefällt wird, müssen zunächst relevante Datenquellen identifiziert und selektiert werden; ein Unternehmen wird typischerweise eine Reihe interner Quellen wie vorhandene Datenbanken, aber und vor allem externe Quellen insbesondere aus dem Web nutzen wollen, da gerade letztere relevante Daten (etwa aus öffentlichen Blogs) liefern können. Die dritte Phase besteht aus einer Detailplanung und umfasst die Wahl der zu verwendenden Technologie, etwa die

¹⁶ <http://www.bigdata-madesimple.com/17-important-case-studies-on-big-data/>

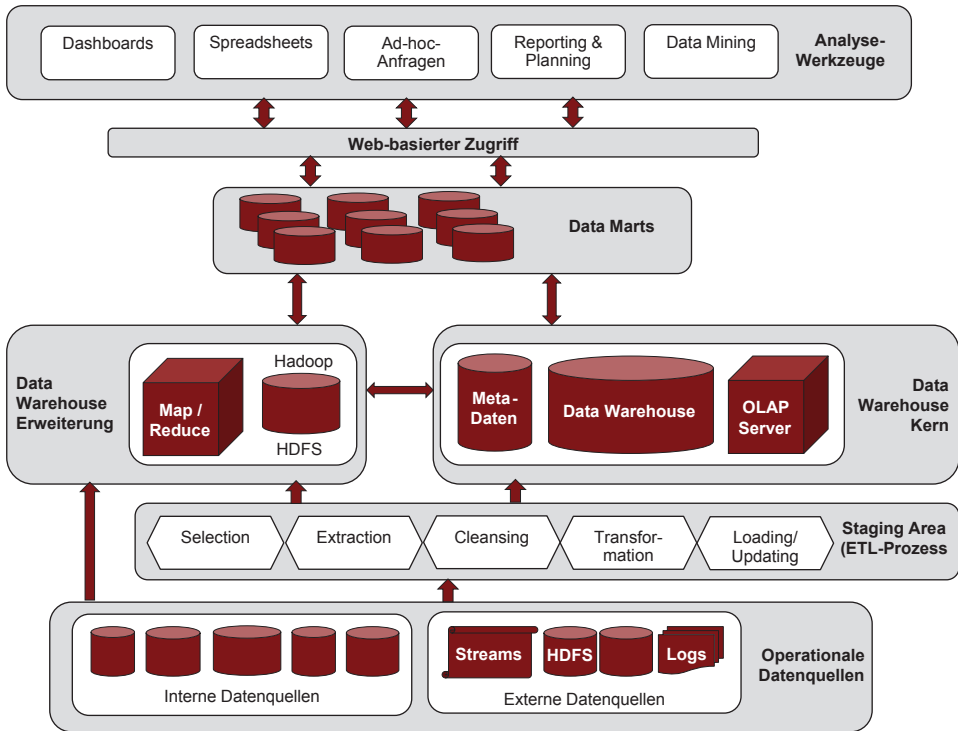


Abb. 10.9 Für Big Data-Nutzung erweiterte Data Warehouse-Architektur

Entscheidung für eine bestimmte Hadoop-Implementierung. Sodann kann eine Realisierung stattfinden; schließlich ist das System bzw. das Projekt operational und benötigt ggfs. regelmäßige oder gelegentliche Wartung.

Wir gehen hier nicht auf weitere Einzelheiten einer Strategieentwicklung ein, bemerken jedoch, dass es auch für Unternehmen, deren Kernkompetenz nicht in der IT liegt, sinnvoll sein kann, von den modernen Möglichkeiten zur Datenanalyse wie Data Warehousing mit OLAP und Data Mining bzw. allgemeiner Business Intelligence Gebrauch zu machen, die in anderen Kapiteln dargestellt werden.

Die grundsätzliche Architektur eines Data Warehouse ist in der rechten Hälfte von Abb. 10.9 gezeigt, die gleichzeitig andeutet, wie sich eine traditionelle Data Warehouse-Architektur für Big Data-Anwendungen erweitern lässt. Neu sind z. B. die weiter gefasste Selektion externer Datenquellen sowie die Erweiterung um eine Map-Reduce-Engine wie z. B. Hadoop (linke Hälfte der Abbildung), was auch bedeutet, dass hier Datei- und Datenbanksysteme neben- und miteinander existieren. Es werden natürlich zusätzliche Kommunikationswege zwischen den alten und den neuen Komponenten benötigt; das Ergebnis kann dann wie in der Abbildung gezeigt aussehen.

Es sei erwähnt, dass Business Intelligence- sowie Analyse-Anwendungen nicht notwendigerweise das Vorhandensein eines Data Warehouse erfordern. Viele der heute verfügbaren Werkzeuge können als Aufsatz oder Erweiterung eines operationalen Systems

oder einer Datenbank betrieben werden, das bzw. die im Unternehmen bereits im Einsatz ist. In einem solchen Fall ist ein expliziter Architekturentwurf nicht erforderlich, allerdings zeigt die Erfahrung, dass insbesondere gut dokumentierte Strategie- sowie Architekturüberlegungen ein Unternehmen vor einem Scheitern entsprechender Projekte schützen kann.

10.4 Zusammenfassung und Ausblick

In diesem Kapitel haben wir mit NoSQL, NewSQL, Map-Reduce und Hadoop neuere Entwicklungen im Datenbankumfeld skizziert. NoSQL-Datenbanken bilden ein Sammelurium unterschiedlichster Ansätze, denen Flexibilitäts- und Skalierbarkeitsanforderungen gemeinsam sind, die von traditionellen, relationalen Datenbanken nicht zufriedenstellend erfüllt werden konnten. Insbesondere wurde und wird in vielen Projekten Eventual Consistency als abgeschwächter Konsistenzbegriff verfolgt, was Latenzvorteile mit sich bringt, aber den Umgang mit ggf. inkonsistenten Daten der Anwendung überlässt, woraus – neben dem Verzicht auf einen etablierten Sprachstandard wie SQL – eine erhöhte Komplexität in der Anwendungsentwicklung resultiert. Gewissermaßen in Verteidigung der Werte Standardisierung und Konsistenz werden demgegenüber NewSQL-Datenbanken entwickelt, die sich wie klassische relationale Systeme nutzen lassen, allerdings auch unter höchsten Anforderungen an Skalierbarkeit. Parallel zu diesen Entwicklungen ist Big Data zum Thema der Jahre 2012 und 2013 (und darüber hinaus) geworden. Daten fallen heute in vorher nicht gekanntem Überfluss an, womit Daten-Verarbeitungstechnologie allerdings inzwischen effizient umgehen kann, was wirtschaftliche Chancen durch eine angemessene Auswertung dieser Daten verspricht. In diesem Kontext haben wir Map-Reduce als Paradigma der parallelisierten, fehlertoleranten Datenverarbeitung und Hadoop als mögliche Implementierungsvariante für dieses Paradigma vorgestellt.

Abschließend wollen wir die Dynamik verdeutlichen, die den konzeptionellen und technischen Entwicklungen rund um die Verarbeitung massiver Datenbestände nach wie vor innewohnt. Im Hinblick auf NoSQL-Systeme wurde kürzlich in (Bailis et al. 2014) mit dem Konzept der *Probabilistically Bounded Staleness* (PBS) ein Werkzeug entwickelt, das es erlaubt, den Grad von Inkonsistenz unter Eventual Consistency in Quorum-Systemen wie Dynamo, Cassandra, Riak und Voldemort präzise zu fassen und zu analysieren. Dies erlaubt es, den Trade-Off zwischen Latenz und Inkonsistenz über Daumenregeln hinaus zu verstehen und in den Entwurf einzubeziehen.

Mit dem auch von NewSQL-Systemen verfolgten Ziel höchster Skalierbarkeit ohne Konsistenzverluste setzt (Bailis et al. 2015) auf das Konzept der Invarianten-Konfluenz, das formalisiert, wann durch Invarianten beschriebene Konsistenzbedingungen in verteilten Systemen lokal, also ohne latenztreibende Koordination garantiert werden können. Die Evaluation des zugehörigen Systems zeigt am etablierten TPC-C-Benchmark eine lineare Skalierbarkeit mit 12,7 Mio. Transaktionen pro Sekunde auf 200 Servern, was sämtliche bisher publizierten Resultate um den Faktor 25 übertrifft.

Schließlich sei darauf hingewiesen, dass es sich bei Map-Reduce um ein Programmierparadigma auf (verglichen mit deklarativen Anfragesprachen wie SQL oder logikbasierten Ansätzen) niedrigstem Abstraktionsniveau handelt. Entsprechend schwierig gestalten sich Programmierung und Optimierung von Analyseaufgaben. Zahlreiche Projekte streben demgegenüber an, ein relationalen Datenbanken vergleichbares Abstraktionsniveau für die Verwaltung und Analyse von Big Data zu erreichen. Stellvertretend seien hier die Open-Source-Projekte AsterixDB (Alsubaiee et al. 2014) und Stratosphere (Alexandrov et al. 2014) (weiterentwickelt als Apache Flink¹⁷) genannt.

Literatur

- Afrati, F., Das Sarma, A., Salihoglu, S., Ullman, J.D.: Vision paper: towards an understanding of the limits of map-reduce computation. CoRR abs/1204.1754 (2012)
- Afrati, F., Das Sarma, A., Salihoglu, S., Ullman, J.D.: Upper and lower bounds on the cost of a map-reduce computation. PVLDB 6(4), 277–288 (2013)
- Alexandrov, A., et al.: The Stratosphere platform for big data analytics. In: The VLDB Journal, Springer-Verlag, Berlin Heidelberg 2014, http://stratosphere.eu/assets/papers/2014-VLDBJ_Stratosphere_Overview.pdf, Zugriffen: 1. Aug. 2015
- Alsubaiee, S., et al.: Storage management in AsterixDB. PVLDB 7(10), 841–852 (2014)
- Bailis, P., Venkataraman, S., Franklin, M.J., Hellerstein, J.M., Stoica, I.: Quantifying eventual consistency with PBS. VLDB J. 23(2), 279–302 (2014)
- Bailis, P., Fekete, A., Franklin, M.J., Ghodsi, A., Hellerstein, J.M., Stoica, I.: Coordination avoidance in database systems, to appear in Proc. VLDB (pre-print at arXiv:1402.2237) (2015)
- Battre, D., et al.: Nephel/PACTs: a programming model and execution framework for web-scale analytical processing. In: Proceedings of ACM Symposium on Cloud Computing (SoCC), S. 119–130. (2010)
- Brewer, E.A.: CAP twelve years later: how the “rules” have changed. IEEE Comput. 45(2), 23–29 (2012)
- Cavage, M., Pacheco, D.: Bringing arbitrary compute to authoritative data. Commun. ACM 57(8), 40–48 (2014)
- Chang, F., et al.: Bigtable: a distributed storage system for structured data. ACM Trans. Comput. Syst. 26(2), 1–26 (2008)
- Corbett, J.C., et al.: Spanner: Google’s globally distributed database. ACM Trans. Comput. Syst. 31(3), 8:1–8:22 (2013)
- Dean, J., Ghemawat, S.: Map-reduce: simplified data processing on large clusters. In: Proceedings of the 6th symposium on Operating System Design and Implementation (OSDI), 2004, and Comm. ACM. 51(1), 107–113 (2008)
- DeCandia, G., et al.: Dynamo: Amazon’s highly available key-value store. In: Proceedings of the 21st ACM symposium on operating systems principles, S. 205–220. (2007)
- Eich, M.H.: Main memory database research directions. In: Proceedings of the international workshop on Database Machines, S. 251–268. (1989)
- Fedak, G.: Special issue of map-reduce and its applications. Concurr. Comput. Pract. Exper. 25(1), 1 (2013)

¹⁷ <https://flink.incubator.apache.org/>

- Fox, A., Brewer, E.A.: Harvets, yield, and scalable tolerant systems. In: Proceedings of the 7th hot topics in Operating Systems, S. 174–178. (1999)
- Fox, A., Gribble, St.D., Chawathe, Y., Brewer, E.A., Gauthier, P.: Cluster-based scalable network services. In: Proceedings of the 16th ACM symposium on Operating Systems Principles, S. 78–91. (1997)
- Lamport, L.: The part-time parliament. *ACM Trans. Comput. Syst.* **16**(2), 133–169 (1998)
- Loos, P., et al.: In-memory databases in business information systems. *Bus. Inf. Syst. Eng.* **6**, 389–395 (2011)
- Lynch, N., Gilbert, S.: Brewer’s conjecture and the feasibility of consistent, available, partition-tolerant web services. *ACM SIGACT News* **33**(2), 51–59 (2002)
- Plattner, H., Zeier, A.: In-memory data management – an inflection point for enterprise applications. Springer, Berlin (2011)
- Rajaraman, A., Lescovec, J., Ullman, J.D.: Mining of massive datasets, 2. Aufl. verfügbar unter <http://infolab.stanford.edu/~ullman/mmds.html> (2013)
- Redmond, E., Wilson, J.R.: Seven databases in seven weeks: a guide to modern databases and the NoSQL movement. Pragmatic Programmers (2012)
- Saecker, S., Markl, V.: Big data analytics on modern hardware architectures: a technology survey. European Business Intelligence Summer School (eBISS), Springer LNBPI, Lecture Notes in Business Information Processing, 138, S. 125–149. (2012)
- Sauer, C., Härder, T.: Compilation of query languages into map-reduce. *Datenbank-Spektrum* **13**(1), 5–15 (2013)
- Schönthaler, F., Vossen, G., Oberweis, A., Karle T.: Business Processes for Business Communities. Springer-Verlag, Berlin (2012)
- Shim, K.: Map-reduce algorithms for big data analysis. Springer LNCS. **7813**, 44–48 (2013)
- Shute, J., et al.: F1: A distributed SQL database that scales. *Proc. VLDB Endow.* **6**(11), 1068–1079 (2013)
- Vogels, W.: Eventually consistent. *Commun. ACM.* **52**(1), 40–44 (2009)
- Weikum, G., Vossen, G.: Transactional Information Systems – Theory, Algorithms, and the Practice of Concurrency Control and Recovery. Morgan Kaufmann, San Francisco (2002)
- White, T.: Hadoop: The Definitive Guide, 3. Aufl. O’Reilly Media, Sebastopol (2012)

Peter Gluchowski

Inhaltsverzeichnis

11.1 Motivation..... 226

11.2 Technologische und organisatorische Entwicklungen bei Analytischen Informationssystemen..... 226

 11.2.1 Advanced/Predictive Analytics..... 227

 11.2.2 Datenmanagement von BI-Systemen..... 228

 11.2.3 Agile BI..... 232

 11.2.4 Self Service BI..... 233

 11.2.5 Mobile BI..... 236

11.3 Zusammenfassung..... 237

Literatur..... 237

Zusammenfassung

Die Technologien und Konzepte, die sich im Bereich der Analytischen Informationssysteme ausmachen lassen, sind stetigen Entwicklungen und Veränderungen unterworfen. Nicht zuletzt aufgrund der hohen kommerziellen Relevanz des Themas lässt sich die zu beobachtende Volatilität sogar als besonders ausgeprägt bezeichnen. Vor diesem Hintergrund greift der vorliegende Beitrag einige augenfällige Entwicklungstendenzen auf und beschreibt die zentralen Neuerungen.

11.1 Motivation

Das Begriffsgebilde Business Intelligence (BI) hat sich in den letzten Jahren zunächst in der Praxis und später auch in der wissenschaftlichen Diskussion als Synonym für IT-Lösungen zur Unternehmensplanung, -steuerung und -kontrolle fest etabliert (Gluchowski et al. 2008, S. 89). Als zentrale Aufgabe der BI in den Unternehmen gilt es, den Fach- und Führungskräften durch das Angebot erforderlicher Inhalte und Funktionalitäten wertvolle technologische Unterstützung zu liefern (Kemper et al. 2010, S. 8). Nach dem heute vorherrschenden, weiten Begriffsverständnis zählen aus technologischer Perspektive zu Business Intelligence alle Systemkomponenten, die dabei helfen, das entscheidungsrelevante Datenmaterial zu sammeln und aufzubereiten, dauerhaft und nutzungsorientiert zu speichern, aufgabenadäquat zu analysieren und in geeigneter Form anzuzeigen (Gluchowski 2001, S. 7; Gluchowski und Kemper 2006, S. 14).

Rückblickend wurde dabei häufig auf etablierte Technologien gesetzt, die seit mehr als einer Dekade verfügbar sind und sich als stabil sowie funktional ausgereift erweisen. Allerdings ist gerade in jüngster Vergangenheit eine Tendenz dahingehend auszumachen, dass neue und innovative Ansätze das vorhandene funktionale, organisatorische und technologische Spektrum im BI-Bereich wenn nicht gar ersetzen, dann aber zumindest ergänzen. Der vorliegende Beitrag widmet sich im folgenden Kapitel diesen aufkommenden Trends einschließlich ihrer Auswirkungen auf die Fachbereiche im Unternehmen. Eine Zusammenfassung greift abschließend die zentralen Ergebnisse der Betrachtung auf.

11.2 Technologische und organisatorische Entwicklungen bei Analytischen Informationssystemen

Ein Überblick über die derzeit intensiv diskutierten und in der Praxis beachteten aktuellen BI-Themen lässt sich durch eine Betrachtung der kürzlich durchgeführten einschlägigen Tagungen zu diesem Themenkontext erlangen. Als führende anbieterunabhängige europäische Konferenz zu den Themenkreisen Business Intelligence und Data Warehousing gilt die Sommertagung des TDWI Germany e. V. Nicht nur an den schweizerischen Markt richtet sich auch die inzwischen alljährlich durchgeführte und von der Universität St. Gallen mitausgerichtete Tagung Data Warehousing in Zürich (das jeweilige Vortragsprogramm dieser Events findet sich unter www.tdwi.eu). Für die vorliegende Untersuchung wurden die Tagungsprogramme jeweils für die Jahre 2012, 2013 und 2014 untersucht und die am häufigsten präsentierten Themen als aktuell identifiziert.

Als meist behandelte Themen präsentieren sich einerseits **Advanced/Predictive Analytics** mit 29 Beiträgen und das **Datenmanagement von BI-Systemen** mit 26 Beiträgen. Eine hohe Anzahl (22) weisen ebenfalls die Vorträge zu **Big Data** auf. Gleichauf mit jeweils 12 Präsentationen liegen die Themenkreise **Neue Datenbanktechnologien** und **Self Service BI**. Knapp dahinter folgt mit 11 angenommenen Einreichungen **Agile BI**. Als

vergleichsweise neues Thema, das verstärkt im Jahre 2013 aufgekommen ist, kann **Mobile BI** mit fünf Beiträgen verstanden werden.

Da die Themenfelder Big Data (Analytics) und Neue Datenbank-Technologien im vorliegenden Sammelband mit spezifischen Beiträgen adressiert werden, konzentriert sich der Text im weiteren Vorgehen auf die verbleibenden fünf der identifizierten Themenfelder.

11.2.1 Advanced/Predictive Analytics

Neben der Bereitstellung einer aufbereiteten, abgestimmten und problemrelevanten Datenbasis stellt vor allem die Auswahl und Anwendung geeigneter Analyseverfahren eine besondere Herausforderung dar. Neben den traditionellen Zugriffsverfahren auf die Daten, wie sie Standard- und Ad-Hoc-Berichtslösungen sowie On-Line-Analytical-Processing-(OLAP-) Systeme für die freie Navigation in multidimensionalen Datenstrukturen bieten, treten heute verstärkt anspruchsvolle Auswertungsverfahren, die ihre Wurzeln sowohl in der Mathematik und Statistik als auch in der Informatikdisziplin der Künstlichen Intelligenz haben. Während diese Verfahren in der Vergangenheit häufig unter dem Oberbegriff Data Mining diskutiert worden sind, findet heute hierfür oftmals das Wortgebilde Advanced Analytics Verwendung. Sind die Analysen primär darauf ausgerichtet, zukünftige Entwicklungen vorherzusagen, dann erweist sich die Begrifflichkeit Predictive Analytics als gebräuchlich (Felden et al. 2012, S. 522).

Die Einsatzfelder für Predictive Analytics sind breit gestreut und reichen von der Ermittlung der Reaktionswahrscheinlichkeit einzelner Kunden und Kundengruppen auf spezifische Marketingaktivitäten über die Prognose des zukünftigen Zahlungsverhaltens eines neuen Geschäftspartners bis hin zur Vorhersage der voraussichtlichen Performance eines Unternehmens.

Vielfach sind die Predictive-Analytics-Anwendungen eng mit den operativen Abläufen verwoben und unterstützen Mitarbeiter beispielsweise im direkten Kundenkontakt dabei, die richtigen Entscheidungen zu treffen. In diesen Fällen, die unter dem Schlagwort **Operational BI** thematisiert werden, steht die zeitnahe Durchführung der Analysen einschließlich der maschinellen Ableitung einer Handlungsempfehlung im Vordergrund. Ein Anwendungsfall hierfür ist etwa im Call Center gegeben, wo eine schnelle Einschätzung der Bonität eines Neukunden unmittelbar während eines Telefongesprächs erfolgen muss, um dem Kunden eine geeignete Zahlungsmethode anbieten zu können. Ein weiteres Einsatzgebiet für Operational BI findet sich bei den Empfehlungssystemen, die dem Kunden bei Zugriff auf das Web-Angebot alternative oder ergänzende Produkte oder Produktbündel am Bildschirm anzeigen.

Durch die Auswertung nicht nur strukturierter Daten, wie sie beispielsweise in relationalen Datenbanksystemen vorliegen, sondern auch semi- und unstrukturierter Daten, die sich beispielsweise in Log-Dateien von Web-Servern, E-Mails oder bei Sensoren einstellen, ergeben sich vollkommen neue Analysechancen. Zu denken wäre hier etwa an

die Auswertung von Maschinendaten im Produktionssektor, von Daten von Antriebsaggregaten in Automobilen und Flugzeugen, aber auch von Wetter-, Bewegungs- und sogar Gesundheitsdaten.

Zusätzliche Erkenntnisse über Wirkungsketten im eigenen geschäftlichen Bereich, welche die komplexe Realität strukturieren und womöglich gar prognostizierenden Charakter tragen, sind seit jeher interessant für die Fachbereiche. In die Zukunft gerichtete Aussagen beispielsweise über die Auswirkungen des Markteintritts neuer Anbieter können helfen, frühzeitig auf dynamisch sich ändernde Wettbewerbsgegebenheiten zu reagieren und Gegenmaßnahmen zu ergreifen.

11.2.2 Datenmanagement von BI-Systemen

Genau genommen erweist sich das Datenmanagement in Unternehmen nicht als neuer Trend, sondern eher als thematischer Dauerbrenner, der allerdings vor dem Hintergrund einer Explosion des verfügbaren und damit nutzbaren Datenvolumens zunehmend an Bedeutung gewinnt. Längst werden Daten nicht mehr als notwendiges Übel der Informationsverarbeitung verstanden, sondern als wichtige Ressource, die nicht nur zur effektiven Steuerung des operativen Geschäfts dient, sondern großes Potenzial zur Erlangung von Wettbewerbsvorteilen birgt.

Aus diesen Gründen haben zahlreiche Unternehmen inzwischen exakte organisatorische Zuständigkeiten bzw. Verantwortlichkeiten für einzelne Datenbereiche, wie Lieferanten, Kunden- oder Artikeldaten, definiert (Otto 2007). Ein derartiger **Data Owner** hat als Manager die Aufgabe, vorgegebene Standards und Richtlinien für die Daten (**Data Governance**) zu erfüllen und dabei vor allem ein Mindestmaß an Datenqualität zu gewährleisten. Ihm zur Seite stehen die **Data Stewards** für die inhaltliche und technische Umsetzung der zugehörigen Vorgaben.

Als eher neuer Ansatz im Bereich des Datenmanagements gilt das **Information Lifecycle Management**. Hierbei geht es darum, die Daten entsprechend ihrer Nutzungshäufigkeit und ihres Wertes für das Unternehmen auf möglichst geeigneten Speichermedien aufzubewahren. Schließlich erweisen sich langsamere Speichermedien als tendenziell kostengünstiger in Anschaffung und Unterhalt, so dass es durchaus sinnvoll ist, selten genutzte Detaildaten lange vergangener Perioden auf eher kostengünstigen Medien (z. B. Bandlaufwerken) zu verwalten, aktuelle Daten dagegen für einen schnellen Zugriff vorzuhalten.

Der Erkenntnis folgend, dass die eigenen Kernkompetenzen in der Regel nicht in der Verwaltung großer Datenmengen liegt, gehen Unternehmen – in Deutschland noch eher zögerlich, in anderen Ländern verstärkt – auch dazu über, die BI-Daten in eine Cloud zu verlagern. Hierbei werden von kommerziellen Anbietern IT-Ressourcen bedarfsgerecht vermietet, auf die dann über das Internet zugegriffen werden kann. Als Vorteil ist zu werten, dass die **Cloud-BI**-Nutzer weitgehend von den Verwaltungsaufgaben des Datenmanagements befreit werden, wie das Einspielen von Software-Updates oder die Datensi-

cherung. Dadurch lassen sich IT-Personalkosten in erheblicher Größenordnung einsparen. Als kritisch ist demgegenüber der Know-How-Verlust einzustufen. Zudem lassen sich die Anforderungen an den Datenschutz auf diese Weise nicht immer gewährleisten.

Die Vorgabe verbindlicher Standards und Richtlinien für die Daten (Data Governance), die sich heute auch darüber hinaus auf den gesamten BI-Sektor im Unternehmen bezieht (BI Governance) führt in Verbindung mit einer exakten Definition organisatorischer Zuständigkeiten dazu, dass einerseits den einzelnen Mitarbeitern – auch im Fachbereich – ein klarer Handlungsrahmen vorgegeben wird und andererseits die Leitungsebene damit ihrer Aufsichtspflicht nachkommt.

Als essenziell zur Erfüllung der Aufgaben einer erfolgreichen Data und BI Governance gilt die Existenz eines angemessenen Metadatenmanagements.

Exkurs Metadatenmanagement

Business-Intelligence-Systeme erweisen sich als komplexe technologische Gebilde, bei denen das Zusammenspiel der einzelnen Komponenten als zentraler Erfolgsfaktor zu werten ist. Häufig finden sich heute Systemlösungen, die sukzessive über die Jahre erweitert worden sind und sich aus zahlreichen Bestandteilen zusammensetzen. Die Wartung, Pflege und Weiterentwicklung dieser Konstrukte gestaltet sich demzufolge als zunehmend schwieriger. Nur durch eine umfassende und aussagekräftige Beschreibung der einzelnen Systemobjekte und -prozesse sowie deren Zusammenspiel lässt sich die Beherrschung der Systeme gewährleisten.

Aus Gründen der Übersichtlichkeit kann eine derartige Beschreibung nicht auf der Ebene einzelner Datenobjekte oder Verarbeitungsschritte erfolgen; vielmehr muss eine höhere Abstraktionsebene gewählt werden. Als geeignetes Medium zur Beschreibung der einzelnen Systembestandteile einer BI-Lösung erweisen sich die zugehörigen BI-Metadaten (Melchert 2004, S. 7).

Definition

Allerdings greift in einem BI-Kontext die aus dem Datenbankumfeld bekannte Kurzdefinition von Metadaten (griechisch meta=hinter, nach, über) als „Daten über Daten“ zu kurz, zumal das Metadatenverständnis hier zudem auch Angaben über die einzelnen Datenverarbeitungsvorgänge und organisatorischen Zuordnungen umfasst (Melchert et al. 2002, S. 3). Vielmehr beinhalten BI-Metadaten dann beschreibende Informationen über Inhalt, Typ, Struktur, Kontext und Bedeutung von Daten, aber auch prozess- und organisationsbezogene Informationen über die Verarbeitung, Verwaltung und Nutzung dieser Daten (Bauer und Günzel 2013, S. 80; Gluchowski et al. 2008, S. 131; Devlin 1997; Auth 2004, S. 33).

Somit beziehen sich Metadaten auf die in einem BI-System enthaltenen Objektdaten sowie den zugehörigen Verarbeitungskontext und abstrahieren von einzelnen Sachverhalten der Realwelt (Melchert 2004, S. 8). Ein noch weiter gefasstes Begriffsverständnis versteht unter Metadaten alle Informationen, die für den Entwurf, die Konstruktion oder Benutzung eines BI-Systems benötigt werden (Bauer und Günzel 2013, S. 340).

Folgerichtig kann der Verwaltung der Metadaten in einem BI-System eine zentrale Rolle zugesprochen werden, da hier alle Informationen über Strukturen und Abläufe der Systems vollständig und widerspruchsfrei vorgehalten werden.

Nutzenpotenziale für BI

Ein angemessener Umgang mit den Metadaten in Form eines geeigneten Metadatenmanagements verspricht unterschiedliche Nutzenpotenziale für die verschiedenen Stakeholder eines BI-Systems wie Entwickler, Entscheider und Anwender. Vor allem lassen sich hier Effizienzsteigerungen in BI-Entwicklung und -Betrieb sowie Effektivitätssteigerungen bei der BI-Nutzung anführen (Melchert 2004, S. 7 f.).

Ein Metadatenmanagement verfolgt somit das Ziel, einerseits den Aufwand für den Aufbau und den Betrieb eines komplexen BI-Systems zu reduzieren und andererseits den Nutzen bei der Verwendung des Systems zu steigern. Eine Aufwandsreduktion lässt sich sowohl durch eine verbesserte Transparenz über das Zusammenspiel einzelner Systemkomponenten erreichen, als auch durch die Nutzung von Metadaten zur Automatisierung der Systemadministration und zur Unterstützung bei der Systemgestaltung, z. B. durch Abhängigkeitsanalysen. Zudem kann gewährleistet werden, dass Änderungen am System durch Orientierung an den Metadaten in einheitlicher Art und Weise durchgeführt werden. Die Erhöhung des Verwendungsnutzens stellt sich durch eine verbesserte Nachvollziehbarkeit der Herkunft einzelner Datenobjekte sowie durch eine eindeutige terminologische Begriffsverwendung ein, die eine präzise und eindeutige Interpretation begünstigt, sich allerdings in der Praxis häufig als große Herausforderung erweist (Eckerson 2004, S. 4).

Ausprägungen

Metadaten finden sich zu allen Komponenten einer BI-Architektur, dazu gehören Staging Areas, Core Data Warehouses, Data Marts, Reporting- und Analysewerkzeuge, ETL-Komponenten und im weitesten Sinne auch Quelldatensysteme (Bauer und Günzel 2013, S. 80).

Grob lassen sich technische von fachlichen Metadaten abgrenzen. Technische Metadaten fallen beim Aufbau (built time) und beim Betrieb (run time) der BI-Lösung quasi als Abfallprodukt an (Lehner 2003, S. 46) und lassen sich oftmals ohne Medienbruch aus den jeweiligen Systemen auslesen und weiterverarbeiten. Sie weisen hilfreiche und teilweise wichtige Informationen für Systementwickler und Administratoren auf. Typische Ausprägungen technischer Metadaten sind konzeptionelle und logische Datenmodelle, physische Speicherinformationen (Indizes, Partitionierung), Datenqualitätsinformationen sowie beispielsweise Erstellungszeitpunkte, Abarbeitungsrhythmen und Verarbeitungsvolumina zu Extraktions-, Transformations- und Ladeprozessen.

Dagegen beschreiben fachliche Metadaten relevante Datenobjekte und Verarbeitungsprozesse aus einer geschäftsorientierten (business) Perspektive und müssen oftmals manuell erfasst werden. Fachliche Metadaten richten sich in erster Linie an den Endanwender und beinhalten z. B. Definitionen, Synonyme und Homonyme, Kommentare sowie Verantwortlichkeiten und Eigentümerschaften.

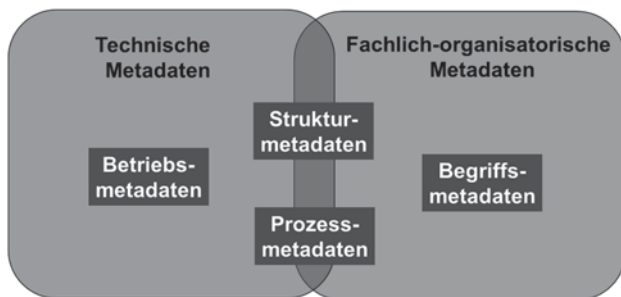


Abb. 11.1 Ausprägungen von Metadaten

Obwohl die Wichtigkeit des Themas erkannt wird, fristet das Metadatenmanagement in den Unternehmen oftmals ein Schattendasein. Während technische Metadaten häufig von den eingesetzten Systemen automatisch zur Verfügung gestellt werden, sind es meist die fachlichen Metadaten, die unzureichend gepflegt sind (Gluchowski et al. 2010, S. 678).

Eine etwas andere Einteilung unterscheidet zwischen Betriebs-, Struktur-, Prozess- und Begriffsmetadaten (in Anlehnung an (Bauer und Günzel 2013, S. 80), vgl. Abb. 11.1). Hierbei sind unter Betriebsmetadaten alle Arten von aufgezeichneten Systeminformationen zu fassen, die während des laufenden Betriebs der Systeme anfallen (Logging). Im Einzelnen handelt es sich um Laufzeiten einzelner Komponenten, Berichte oder Abfragen sowie Angaben über Größe und Art verarbeiteter oder gespeicherter Datenmengen (Speicherplatz, Tabellenzeilen, Anzahl und Durchsatz paralleler Datenbankverbindungen, Benutzerstatistiken, Anzahl erfolgreicher und nicht erfolgreicher ETL-Prozesse sowie der zugehörige Durchsatz).

Demgegenüber beinhalten Strukturmetadaten strukturelle Angaben über die Architektur des gesamten BI-Systems. Im Vordergrund stehen hier die logischen und physischen Datenschemata, einschließlich der Beschreibung der zugehörigen Datenfelder und Datentypen. Weiterhin gehören dazu Integritätsbedingungen, Indizes, logische und physisch materialisierte Sichten, Dimensionen und Fakten sowie deren Partitionierung und physische Aggregation (Bauer und Günzel 2013, S. 80). Darüber hinaus lassen sich auch personelle bzw. rollenorientierte Zuordnungen bezüglich Ersteller, Besitzer, Betreiber und Nutzer vornehmen, aber auch statistische Zugriffshäufigkeiten.

Zu den Prozessmetadaten zählen alle Angaben zum Transport und zur Verarbeitung von Daten, die sich im implementierten System in den zugehörigen ETL-Jobs wieder finden. Als Metadaten ergeben sich hier Angaben zum Ersteller, Erstellungs- und Änderungszeitpunkten sowie zur Reihenfolge der abzuarbeitenden ETL-Pakete (scheduling). Zudem lassen sich auch zentrale Transformationsschritte durch Metadaten beschreiben, auf einer fachlichen Ebene in Form sogenannter Business Rules.

Begriffsmetadaten werden ausschließlich vom Fachbereich gepflegt und beinhalten folgerichtig fachliche Angaben zur Erläuterung oder Interpretation von Datenobjekten. Hierzu gehören Definitionen, Einsatzbereiche und Berechnungsvorschriften sowie Querbeziehungen zu verwandten Begriffen. Häufig wird in diesem Zusammenhang der Aufbau

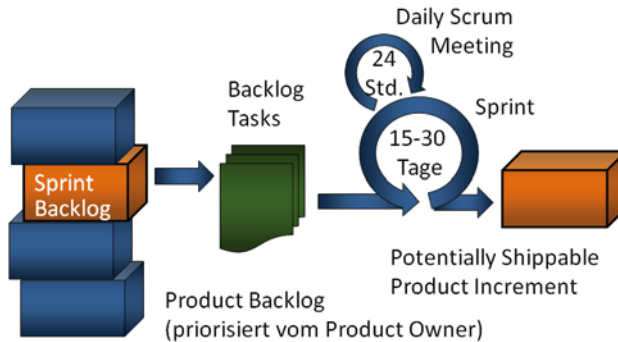


Abb. 11.2 Scrum als Agile Entwicklungsmethodik

eines unternehmensweiten Glossars für die fachliche Beschreibung der relevanten Datenobjekte gefordert (Bauer und Günzel 2013, S. 343).

11.2.3 Agile BI

Vor allem bei größeren Unternehmen erweisen sich die langen Umsetzungszeiten vom Entstehen des Bedarfs bis zur Implementierung und zur Freigabe oftmals als großes Problem im BI-Bereich. Klassisch lineare Entwicklungsprozesse, wie beispielsweise durch das Wasserfallmodell vorgegeben, sind zu schwerfällig, um mit dynamisch sich ändernden Anforderungen umgehen zu können und zeitnah Lösungen zu liefern. Besonders die Bedarfe hinsichtlich der Lösungen für die Informationsversorgung und die Analyse unterliegen zudem häufig einer hohen Volatilität.

Als Antwort hierauf gelten die agilen Entwicklungsmethoden, die vor allem die BI-Ablauforganisation, aber auch die BI-Aufbauorganisation und die BI-Architektur betreffen. Häufig werden hierbei agile Vorgehensmodelle und besonders häufig **Scrum** (Komus 2013, S. 86) eingesetzt (vgl. Abb. 11.2), die versprechen, innerhalb weniger Wochen neue Versionen der BI-Lösung anbieten zu können.

Als zentraler Schritt bei der Scrum-Entwicklungsmethodik gilt der Sprint, der nach einem festgelegten Zeitintervall im Bereich weniger Wochen ein vorab definiertes Ergebnis mit funktionellem Mehrwert liefern soll, das dem Anwender zur Verfügung gestellt werden kann (**Potentially Shippable Product Increment**). Tägliche Scrum Meetings informieren die Teammitglieder über den Entwicklungsfortschritt und dienen der Absprache anstehender Arbeitspakete. Der iterative Durchlauf durch aufeinander folgende Sprints erfolgt so lange, bis alle relevanten Anforderungen erfüllt sind. Ein **Product Backlog** sammelt die artikulierten Anforderungen in Form sogenannter User Stories. Vor Beginn eines Sprints werden die hier umzusetzenden Anforderungen bzw. User Stories (**Sprint Backlog**) im Entwicklungsteam diskutiert, um dann diese eher fachlichen Beschreibungen in technische Aufgaben (**Backlog Tasks**) zu übersetzen.

Um einen reibungslosen Ablauf beim Einsatz der Methodik gewährleisten zu können, sind aufbauorganisatorisch einige Rollen mit fest vorgegebenen Zuständigkeiten und Verantwortlichkeiten vorgegeben, wie die des Product Owners und die des Scrum Masters. Der **Product Owner** hat die inhaltliche und wirtschaftliche Verantwortung für das zu entwickelnde Produkt. Dabei gehört es beispielsweise zu seinen Aufgaben, die Anforderungen im Product Backlog zu priorisieren und hierbei eine enge Verbindung zu den Fachanwendern zu halten. Der **Scrum Master** ist für die korrekte Anwendung der Scrum Methodik zuständig und beseitigt alle auftretenden Hindernisse, die beispielsweise einem täglichen Treffen entgegenstehen oder aus mangelnder Kommunikation zwischen den Mitgliedern des Entwicklungsteams resultieren.

Auch für die BI-Architektur ergeben sich aus der Anwendung agiler Entwicklungsmethoden Konsequenzen, zumal eine allzu starre und monolithische Architektur, bei der sich Änderungen nur mit erheblichem Aufwand realisieren lassen, kurzen Entwicklungszyklen mit inkrementellen Erweiterungen entgegensteht. Vielmehr bedarf es flexibler und gut anpassbarer BI-Komponenten mit gutem Zusammenspiel.

Für die Fachbereiche erweist sich die Agile BI deshalb als interessant, weil sie verspricht, (Teil-) Ergebnisse bereits nach kurzer Entwicklungszeit produktiv nutzen zu können. Insbesondere der zunehmenden Dynamik und dem wachsenden Zeitdruck soll so wirksam begegnet werden. Als spezielle Ausprägung erweist sich hierbei das Pair Programming (Krawatzek et al. 2013, S. 60), bei dem Mitarbeiter aus der Fachabteilung gemeinsam mit IT-Entwicklern vor einem Bildschirm Anwendungen in kurzen Zyklen erweitern.

11.2.4 Self Service BI

Als häufig zu beobachtendes Phänomen in den Fachabteilungen von Unternehmen erweist sich der Umstand, dass trotz sehr leistungsfähiger und individuell zugeschnittener BI-Applikationen die Mitarbeiter eine Nutzung von Tabellenkalkulationsprogrammen weiterhin vorziehen. Die Gründe hierfür sind sicherlich vielfältig. Einerseits könnte eine Ursache darin liegen, dass sich die BI-Lösungen oftmals immer noch als zu starr und zu wenig flexibel erweisen, wenn es darum geht, neue Anforderungen in einer komplexen BI-Landschaft rasch umzusetzen (siehe die Ausführungen zu Agile BI) oder innovative Analysen testweise im eigenen Hause auszuprobieren, um deren Wertbeitrag zu überprüfen.

Hier setzt das Konzept der Self Service BI an, das verspricht, mit neuartigen Produkten den Fachanwender in die Lage zu versetzen, nicht nur eigenständig und flexibel den Datenbestand nach neuen Verknüpfungen zu untersuchen, sondern auch unabhängig von der Unternehmens-IT seine Systeme betreiben und intuitiv bedienen zu können (Imhoff und White 2011, S. 5; vgl. Abb. 11.3).

Häufig sind dazu über die herkömmlichen Funktionalitäten für die Navigation und Investigation von Daten hinaus auch weiterführende Analysefunktionen aus der Statistik

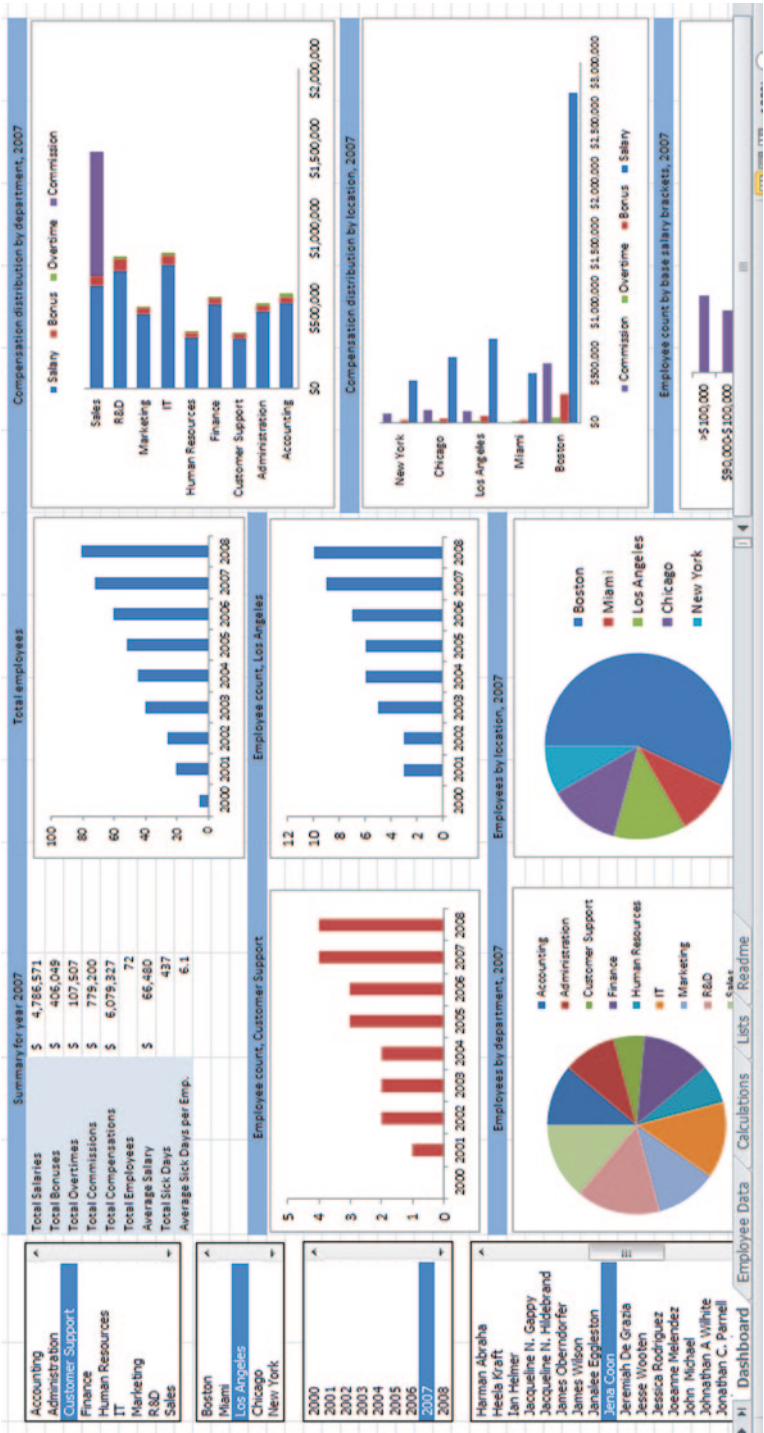


Abb. 11.3 Self-Service-BI-Anwendung

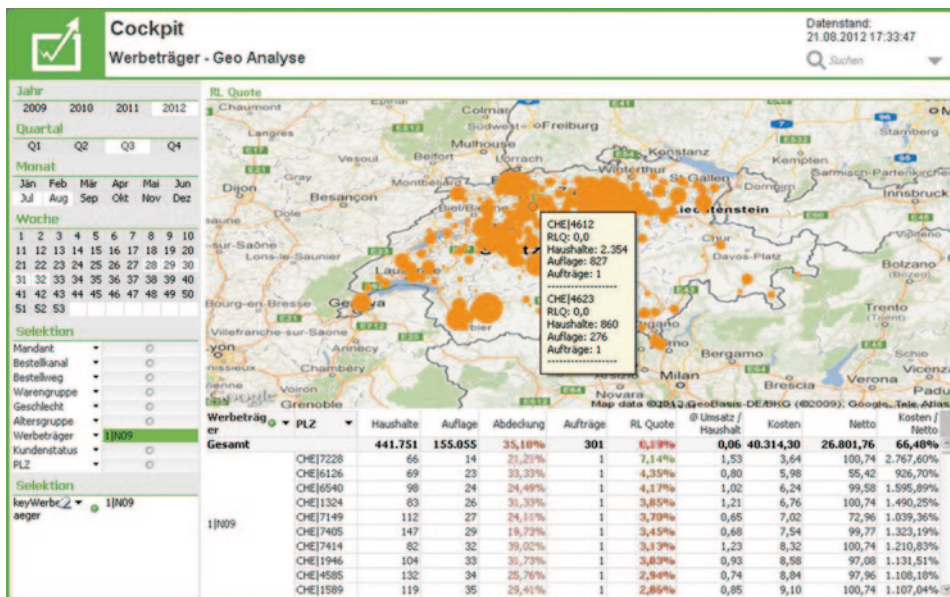


Abb. 11.4 BI-Anwendungen mit geografischen Informationen

und aus dem Forschungsbereich Human Computer Interaction zur Visualisierung und Untersuchung hochdimensionierter Datenstrukturen implementiert (**Visual BI**).

Als sehr eingängig erweisen sich zudem Kartendarstellungen, die es erlauben, quantitative Unternehmensgrößen in einer geografischen Darstellung zu visualisieren, z. B. um Verkaufszahlen in einzelnen Vertriebsregionen darzustellen (vgl. Abb. 11.4).

Um hierzu auch dringende Informationsbedarfe befriedigen zu können, wird oftmals auf das Konzept des **Sandboxing** zurückgegriffen, bei dem den Fachanwendern Datenmaterial in einem abgegrenzten Datenbankbereich zur freien Verfügung gestellt wird. Erst wenn sich dieses Datenmaterial nach einer eingehenden Erprobung durch den Fachanwender und nach der testweisen Erstellung spezifischer Analysen und Berichte als werthaltig für das Unternehmen erweist, werden die Daten und zugehörigen Auswertungen in den Regelbetrieb übernommen und auch anderen BI-Usern zur Verfügung gestellt.

Self Service BI in Verbindung mit dem Sandboxing adressiert unmittelbar die zunehmende Mündigkeit der Fachanwender im Hinblick auf Systeme und Anwendungen, die durch die Option zur eigenständigen Durchführung neuer Auswertungen und Analysen dem steigenden Zeitdruck und der wachsenden Dynamik wirksam begegnen können. Durch die fortschrittlichen Darstellungsformen der Visual BI gelingt es zudem, auch hochkomplexe Sachverhalte zu strukturieren und auszuwerten (Eckerson und Hammond 2011, S. 7). Je nach Anspruchshaltung der Anwender bieten die Werkzeuge einen leichten Einstieg für gelegentliche Nutzer, aber auch fortgeschrittene Funktionalitäten mit methodischer Tiefe.



Abb. 11.5 BI-Anwendungen auf mobilen Endgeräten

11.2.5 Mobile BI

Die generelle Tendenz zur Nutzung mobiler Endgeräte wie Smartphones und Tablets macht auch vor dem Unternehmenskontext keinen Halt. Schließlich sind auch außerhalb des eigenen Unternehmensgeländes heute vielfältige Möglichkeiten gegeben, einen Zugang zum Internet aufzubauen und diesen für geschäftliche Zwecke zu nutzen. Zu denken ist hier an die vielfältigen öffentlich zugänglichen Netzknoten (Hotspots), auf die im kabellosen Modus über ein Wireless LAN (WLAN) zugegriffen werden kann.

Bisweilen gehen die Unternehmen dazu über, den Mitarbeitern zu erlauben, auch private Endgeräte in das Unternehmensnetzwerk einzubringen (bring your own device) und damit die Unternehmensapplikationen zu nutzen (Disterer und Kleiner 2013, S. 92–100). Die Möglichkeit, Business Intelligence Applikationen mobil zu verwenden, wird heute unter dem Schlagwort Mobile BI diskutiert (Stodder 2012, S. 5). Dabei erweist es sich als wichtig zu verstehen, dass es nicht nur darum geht, die vorhandenen Applikationen in eine mobil nutzbare Umgebung zu transferieren, sondern gleichsam die verfügbaren Bedienkonzepte zur Navigation und Selektion (wie z. B. das „Wischen“ über den Bildschirm) zu adaptieren und in geeigneter Weise einzusetzen (vgl. Abb. 11.5).

Verstärkt gelingt es auf diese Weise, auch wenig IT-affinen Mitarbeitern den Zugang zu BI-Informationen und -Analysen zu öffnen, zumal sich die verfügbaren Oberflächen als weitgehend selbsterklärend und intuitiv nutzbar erweisen.

Als großes Problem erweist sich nach wie vor in diesem Kontext der Aufbau sicherer Verbindungen zwischen den mobilen Endgeräten und dem Unternehmensnetzwerk. Schließlich handelt es sich hier um höchst sensible Daten, die zum Teil über öffentlich zugängliche Netzwerke transportiert werden müssen. Auch für den Fall des Verlustes bzw. eines Diebstahls der Endgeräte muss Vorsorge getroffen werden, damit die möglicherweise lokal gesicherten Daten nicht durch Dritte geschäftsschädigend genutzt werden können.

Für die Fachbereiche ergibt sich aus Mobile BI die Option zur Übertragung des allseits gelebten Trends Mobilität auf den Sektor der Geschäftsapplikationen. Informationen und Analysefunktionalitäten an jedem Ort verfügbar zu haben, kann sich als entscheidender Wettbewerbsfaktor herausstellen.

11.3 Zusammenfassung

Der vorliegende Beitrag verfolgt das Ziel, die aktuellen Trends im Business-Intelligence-Sektor zu erörtern. Dazu wurden fünf zentrale BI-Trends unter den Oberbegriffen Advanced/Predictive Analytics, Datenmanagement, Agile BI, Self Service BI sowie Mobile BI beschrieben.

Die Ausführungen konnten belegen, dass sich sowohl auf organisatorischer als auch auf technischer Ebene vielfältige Neuerungen einstellen, mit denen sich die Aufgaben in den Fachbereichen besser als in der Vergangenheit bewältigen lassen. Sicherlich sind die Betrachtungen lediglich als Momentaufnahme zu verstehen, zumal stetig neue Innovationen im BI-Umfeld hervorgebracht werden, die das verfügbare instrumentelle Spektrum erweitern.

Literatur

- Auth, G.: Prozessorientierte Organisation des Metadatenmanagements für Data-Warehouse-Systeme. Norderstedt (2004)
- Bauer, A., Günzel, H.: Data-Warehouse-Systeme: Architektur, Entwicklung, Anwendung, 4. Aufl. dpunkt, Heidelberg (2013)
- Devlin, B.: Data warehouse – from architecture to implementation. Addison Wesley, Amsterdam (1997)
- Disterer, G., Kleiner, C.: BYOD – Bring your own device. In: Gluchowski, P., Reinheimer, S. (Hrsg.) Agilität in der IT, HMD, 50. Jg., Bd. 290, S. 92–100. dpunkt, Heidelberg (2013)
- Eckerson, W.: Mastering metadata. Bus. Intell. J. 9(4), Fall S. 34–37 (2004)
- Eckerson, W., Hammond, M.: Visual reporting and analysis, TDWI best practice report, Renton 1. Quartal (2011)
- Felden, C., Koschtial, C., Buder, J.: Predictive Analytics in der Strategischen Anlagenwirtschaft. In: Mertens, P., Rässler, S. (Hrsg.) Prognoserechnung, 7. Aufl., S. 519–537. Physica, Berlin (2012)
- Gluchowski, P.: Business intelligence. Konzepte, Technologien und Einsatzbereiche. HMD Theorie und Praxis der Wirtschaftsinformatik, 38 Jg., Heft 222, S. 5–15. dpunkt, Heidelberg (2001)
- Gluchowski, P., Kemper, H.-G.: Quo Vadis Business Intelligence? Aktuelle Konzepte und Entwicklungstrends. In: BI-Spektrum, 1. Jg. H. 1, S. 12–19. Sigs Datacom, Troisdorf (2006)

- Gluchowski, P., Gabriel, R., Dittmar, C.: Management Support Systeme und Business Intelligence, Computergestützte Informationssysteme für Führungskräfte und Entscheidungsträger, 2. Aufl. Springer, Berlin (2008)
- Gluchowski, P., Kurze, C., Wunderlich, M.: Modellierung und Dokumentation von BI-Systemen. In: Controlling – Zeitschrift für erfolgsorientierte Unternehmensplanung, 22. Jg., Heft 12, S. 672–678. (2010)
- Imhoff, C., White, C.: Self-service business intelligence, empowering users to generate insight, TDWI best practice report, Renton 3. Quartal (2011)
- Kemper, H.-G., Baars, H., Mehanna, W.: Business Intelligence – Grundlagen und praktische Anwendungen, 3. Aufl. Vieweg, Wiesbaden (2010)
- Komus, A.: Agile Methoden in der Praxis – Studie zur Anwendung und Zufriedenheit. In: Gluchowski, P., Reinheimer, S. (Hrsg.) Agilität in der IT, HMD, 50. Jg., Bd. 290, S. 84–91. dpunkt, Heidelberg (2013)
- Krawatzek, R., Zimmer, M., Trahasch, S.: Agile Business Intelligence – Definition, Maßnahmen und Herausforderungen. In: Gluchowski, P., Reinheimer, S. (Hrsg.) Agilität in der IT, HMD, 50. Jg., Bd. 290, S. 56–63. dpunkt, Heidelberg (2013)
- Lehner, W.: Datenbanktechnologie für Data-Warehouse-Systeme: Konzepte und Methoden. dpunkt, Heidelberg (2003)
- Melchert, F.: Metadatenmanagement im Data Warehousing, Ergebnisse einer empirischen Studie, Universität St. Gallen (2004)
- Melchert, F., Auth, G., Herrmann, C.: Integriertes Metadatenmanagement für das Data Warehousing. Grundlagen, Nutzenpotenziale, Architektur, Universität St. Gallen (2002)
- Naisbitt, J.: Megatrends. 10 Perspektiven, die unser Leben verändern werden, Bayreuth (1984)
- Otto, B.: Datenmanagement als strategischer Erfolgsfaktor, Vortrag im Rahmen des SAP NetWeaver Master Data Management Infotag 2007, Regensburg (2007)
- Stodder, D.: Mobile business intelligence and analytics, extending insight to mobile workforce, TDWI best practice report, Renton 1. Quartal (2012)

Teil III

Betriebswirtschaftliche Anwendung und spezielle Aspekte

Planung und Informationstechnologie – Vielfältige Optionen erschweren die Auswahl

12

Karsten Oehler

Inhaltsverzeichnis

12.1	Einleitung	242
12.2	Werkzeug-Kategorien	242
12.2.1	ERP (Enterprise Resource Planning)	243
12.2.2	Tabellenkalkulation	243
12.2.3	Generische OLAP Werkzeuge (Generisches BI)	244
12.2.4	Dedizierte Planungsanwendungen (BI Anwendungen)	244
12.3	Die Beurteilung im Einzelnen	245
12.3.1	Zentrale Administration	245
12.3.2	Workflow-Unterstützung	246
12.3.3	Planung auf verdichteten Strukturen	247
12.3.4	Validierung	247
12.3.5	Automatisierte Top-Down-Verteilungen	248
12.3.6	Simulationsrechnungen	248
12.3.7	Vorbelegung	249
12.3.8	Aggregation	250
12.3.9	Kommentierungen	250
12.3.10	Eingabe über das Internet	251
12.3.11	Detailplanungen	251
12.3.12	Abstimmung des Leistungsnetzes	252
12.3.13	Übernahme von Stamm- und Bewegungsdaten aus ERP-Systemen	252
12.3.14	Verknüpfung von Teilplanungen	253
12.3.15	Abbildung abhängiger Planungsgrößen	254
12.3.16	Nachvollziehbarkeit	255

K. Oehler (✉)
CPM Solutions, pmOne, Alzenau, Deutschland
E-Mail: karsten.oehler@pmone.com

12.4 Zusammenfassende Bewertung	255
12.5 Ausblick	258
Literatur	258

Zusammenfassung

Software-Systeme aus den unterschiedlichsten Kategorien können zur Unterstützung von Planungsaufgaben eingesetzt werden. Häufig werden allerdings suboptimale Entscheidungen aus mangelnder Kenntnis der Eignung der jeweiligen Systeme getroffen. Die wichtigsten Anforderungen sollen in diesem Beitrag diskutiert und die Leistungspotenziale der Softwarewerkzeuge gegenübergestellt werden.

12.1 Einleitung

Unternehmensplanung ist ohne Informationstechnologie kaum effizient durchführbar. Hier besteht ein weitreichender Konsens, sodass von einer detaillierten Auseinandersetzung mit dem Thema abgesehen werden kann. Kontrovers ist allerdings die Art und Weise, wie eine solche Unterstützung erfolgen sollte. Zahlreiche Werkzeuge bieten dem Planenden mehr oder weniger umfangreiche Unterstützung bei der Durchführung seiner Aufgaben. (Oehler 2006, S. 296–334) Hier gibt es deutliche Unterschiede was den optimalen Einsatz angeht und somit besteht Klärungsbedarf,

Allerdings ist Planung keine homogene Aufgabe. Eine Vielzahl von Funktionen bildet das Unternehmensplanungssystem. Diese Funktionen stellen unterschiedliche Anforderungen an die IT-Unterstützung. Die umfassende Befriedigung der hieraus entstehenden Bedürfnisse an Unterstützung führen bei vielen Unternehmen zu einem hybriden Systemansatz.

In diesem Beitrag sollen die unterschiedlichen Aspekte der IT-Unterstützung diskutiert werden. Hierbei werden die unterschiedlichen Werkzeuge hinsichtlich ihrer Eignung untersucht. Diese Analyse wird mit Hilfe eines Funktionskataloges durchgeführt.

Die Software-Kategorien sind nicht homogen und eine Einzelanalyse würde den Rahmen dieses Beitrags sprengen. Trotzdem können diese Ausführungen bei der Auswahlentscheidung hilfreich sein. Die Projektentscheider sollten bei der Auswahl versuchen, die für sie wichtigen Kriterien auszuwählen und die konkret ausgewählten Lösungsvorschläge anhand dieser beurteilen.

12.2 Werkzeug-Kategorien

Unter Planung, Budgetierung und Forecasting werden unterschiedlichste Prozesse subsumiert. Entsprechend reichhaltig sind die Möglichkeiten einer Unterstützung durch IT-Systeme. Dies zeigt sich auch in den unterschiedlichen Ansätzen, die von einer eher starren,

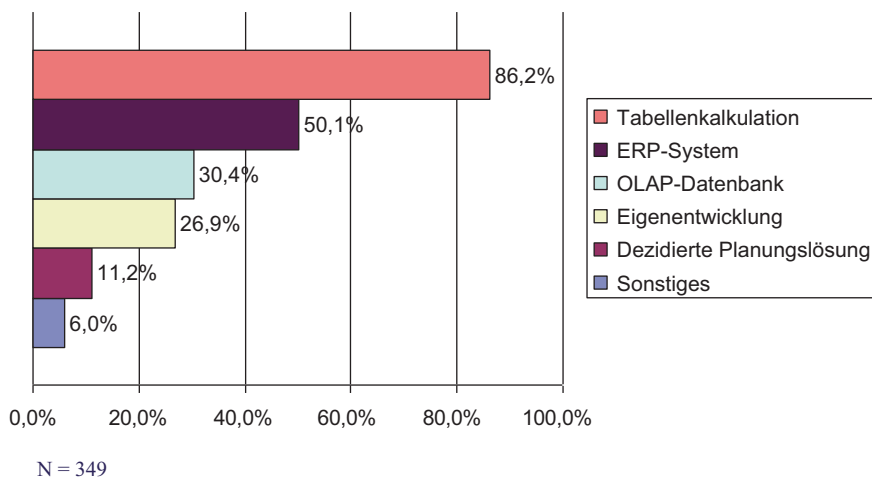


Abb. 12.1 Werkzeugeinsatz in der Planung. (Seufert und Oehler 2009)

zentralistisch orientierten Planungsunterstützung bis zur völlig flexiblen Dezentralisierung durch Office-Werkzeuge reichen. Die Verbreitung der unterschiedlichen Werkzeuge zeigt Abb. 12.1. Grundlage ist eine Studie, die vom Institut für Business Intelligence erstellt worden ist.

12.2.1 ERP (Enterprise Resource Planning)

ERP-Systeme sind primär für die Geschäftsprozessunterstützung und die Abrechnung entwickelt worden. Bekannte Systeme sind SAP ERP, Oracle eBusiness Suite oder Microsoft Dynamics. Integraler Bestandteil eines solchen Systems ist die operative Planung. Eine gemeinsame Nutzung betrieblicher Objekte wie Kostenstellen, Kontenrahmen usw. gleichermaßen für Planung und Ist-Abrechnung ist nahe liegend. In den etablierten ERP-Systemen sind zahlreiche Planungsfunktionalitäten enthalten. Absatzplanung, Budgetierung auf Kontenwerten oder Produktionsplanung sind einige Beispiele.

ERP-Systeme unterstützen in der Regel eine fest definierte Anzahl von Methoden und Modellen. So bieten diverse Anwender wie beispielsweise SAP spezielle Module für die Prozesskostenrechnung in Form einer Planungsrechnung an.

12.2.2 Tabellenkalkulation

Spitzenreiter bezüglich der Verbreitung ist die Tabellenkalkulation. Die meisten Planungsapplikationen werden auf dieser Basis aufgebaut. Selbst Großunternehmen nutzen solche Werkzeuge für den Planungsprozess.

Die Vorteile scheinen auf der Hand zu liegen: Die Tabellenkalkulation ist schnell einsetzbar, da geringe bzw. keine Anschaffungskosten entstehen, da sie meist schon unternehmensweit vorhanden ist. Außerdem können die meisten Benutzer bereits mit dem System umgehen. Diese Systeme bieten eine (scheinbar) hohe Flexibilität (RaEi00, S. 109). Allerdings ist diese Flexibilität trügerisch, da die Fehlerhäufigkeit mit zunehmender Komplexität deutlich steigt.

12.2.3 Generische OLAP Werkzeuge (Generisches BI)

OLAP steht für Online Analytical Processing und ist als Antwort zu OLTP (Online Transaction Processing) zu sehen. Eine beliebte Metapher zur Verdeutlichung von OLAP ist der Würfel, der bekanntermaßen drei Dimensionen besitzt. Unter einer Dimension ist eine interessierende Auswertungsrichtung zu verstehen. Am Beispiel einer Vertriebsauswertung können dies die Dimensionen Zeit, Produkte, Kundengruppen sowie Positionen wie Umsatz oder Deckungsbeitrag sein. Wichtige Systeme in diesem Bereich sind INFOR ALEA, Hyperion Essbase und IBM Cognos TM1.

12.2.4 Dedizierte Planungsanwendungen (BI Anwendungen)

Es gibt seit langem einen Markt für Planungs- und Budgetierungssysteme. Diese zeichnen sich durch eine hohe funktionale Ausrichtung auf die spezifischen Anforderungen der Planungsprozesse aus. Es ist eine starke Modularisierung mit begrenzten Möglichkeiten zur Integration in eine umfassende BI-Architektur festzustellen.

Typische Planungswerkzeuge bieten weit reichende Unterstützung beim Aufbau von Planungsmodellen. Durch die flexible Verknüpfung von Planungselementen eignen sich solche Systeme für die Simulation. So kann jede Basisänderung in ihrer vollständigen Auswirkung auf die Erfolgs- und Finanzsicht einer Unternehmung analysiert werden. In der Regel wird ein unmittelbarer Vergleich zwischen Simulationsstand und Ausgangsbasis unterstützt.

Bezüglich des Vordefinitionsgrades gibt es erhebliche Unterschiede. Während einige Anbieter Module wie eine integrierte Erfolgs- und Finanzlogik fest implementieren, unterstützen andere Werkzeuge eher den Baukastenansatz. Die Vor- und Nachteile sind im Hinblick auf die jeweiligen Umfeldparameter zu beurteilen. Eine hohe Dynamik lässt beispielsweise den Einsatz eines Baukastens als vorteilhafter erscheinen. Tabelle 12.1 listet einige verbreitete Planungsanwendungen auf.

Tab. 12.1 Anbieter von Planungssystemen

Produkt	Anbieter
ABC for OLAP	CUBUS
CoPlanner	CoPlanner
Corporate planner	Corporate planning
Cognos TM1	IBM (vormals cognos)
FMS	SAS
MPC	INFOR (vormals geac)
Planning	Oracle (vormals hyperion)
Professional planner	Winterheller software
BI IP, BPC	SAP
Tagetik	Tagetik/pmOne (für DACH)
TN planning	Thinking networks
4Plan MD	Software4You

12.3 Die Beurteilung im Einzelnen

Eine pauschale Beurteilung eines Planungssystems hilft kaum weiter. Spezifische Planungsfunktionen müssen beschrieben werden, um die Stärken und Schwächen der Systeme deutlich zu machen. Daher sollen im Folgenden typische Planungsfunktionen erläutert werden.

Die Analyse bezieht die bereits erwähnte Studie, Erkenntnisse aus der Literatur sowie die langjährige Erfahrung des Autors in Beratung und Umsetzung aller vorgestellten Lösungen ein. Da allerdings die Klassen bezüglich ihrer Leistungsfähigkeit nicht homogen sind, lastet dieser Einschätzung eine gewisse Subjektivität an.

12.3.1 Zentrale Administration

Eine zentrale Administration umfasst Funktionen wie beispielsweise Berichtserstellung, Anlegen von Benutzern, Definition von Abhängigkeiten etc. Die zentrale Administration ist aus Aufwandsgründen vorteilhaft. Die grundsätzliche Durchführbarkeit von Anpassungen durch die Zentrale während des Planungsprozesses ist kritisch. Werden Planungsobjekte physisch an die Planenden verteilt, wie es bei der Tabellenkalkulation der Fall ist, ist eine Strukturaktualisierung sehr aufwändig. Allerdings schränkt eine zentrale Administration auch die Flexibilität ein. Denn nicht alle Aspekte eines Planungsprozesses sind zentral bekannt. So ist das Wissen um die Planungsobjekte in den jeweiligen Funktionsbereichen meistens tiefer ausgeprägt. Die Anpassungsfähigkeit leidet darunter, dass entsprechende Abhängigkeiten zur Umsetzung erst an die zentrale Administration weitergeleitet werden müssen.

ERP-Systeme werden zentral verwaltet. Allerdings kann in der Regel das Einrichten der meisten Objekte an beliebige Benutzerrollen delegiert werden. Änderungen in den Strukturen sind in der Regel sofort bei der Einzelplanung wirksam.

Auch OLAP-Systeme werden üblicherweise zentral administriert. Einige Systeme unterstützen ein sogenanntes Multi-Administratorenkonzept. Rechte zur Verwaltung der Planungsobjekte können damit delegiert werden. Auch hier erfolgt eine unmittelbare Aktualisierung der Eingabemasken.

Das gleiche trifft auch für einige Planungssysteme zu. Auch hier sind Werkzeuge mit einem Multiadministratorenkonzept verfügbar. Allerdings basieren einige Werkzeuge auf der physischen Verteilung von Planungsscheiben. Sind diese Teilplanungen verteilt, ist eine Stammdatenaktualisierung etwas aufwändiger, kann aber systemseitig unterstützt werden.

Lediglich die Tabellenkalkulation schwächelt hier deutlich: Durch die physische Verteilung der Teilplanungen muss nach Anpassung der Strukturen eine erneute Verteilung stattfinden. Insofern sind Änderungen während des Planungsprozesses nur sehr eingeschränkt durchzuführen.

12.3.2 Workflow-Unterstützung

Eine Workflow-Unterstützung dokumentiert die jeweiligen Planungszustände im Prozess und leitet Planungsergebnisse an die jeweilig verantwortlichen Personen weiter. Eine wichtige Funktion ist beispielsweise die Benachrichtigung des Vorgesetzten bei der Fertigstellung einer Planscheibe. Parallel dazu wird in der Regel das Planungsergebnis nach der Fertigstellung der jeweiligen Planungsscheibe eingefroren. Das Genehmigungsverfahren muss auf unveränderlichen Informationen basieren und dies erfordert das Sperren der Vorschläge nach Abgabe der Zahlen. Hilfreich bei größeren Planungsprojekten ist auch ein Monitor, der den Status der jeweiligen Einheiten darstellt.

ERP-Systeme verfügen häufig über eine ausgereifte Workflow-Funktionalität. Hierzu gehören auch Vertretungsregeln und Eskalationen. Hier ist es relativ einfach, auch den Planungs- und Budgetierungsprozess einzuhängen.

In OLAP-Werkzeugen muss eine solche Funktionalität meistens hinzuentwickelt werden. Dies kann bei komplexen Prozessen recht aufwändig werden.

Die Tabellenkalkulation wird meistens in Kombination mit einer Groupware-Komponente wie IBM Lotus Notes oder Microsoft Outlook eingesetzt. Diese Werkzeuge ermöglichen die Einrichtung eines Workflows. Auch ein „Einfrieren“ der Daten nach der Abgabe erübrigt sich, da die Ergebnisse verschickt werden und damit für den Absender nicht mehr veränderbar sind. Allerdings ist das Handling unterschiedlicher Versionen kompliziert. Auch das Verschicken der Vorlagen und Resultate per Groupware ist aufwändig und muss häufig durch selbst entwickelte Makros unterstützt werden.

Spezifische Planungslösungen haben häufig eine eingebaute Workflow-Unterstützung, die auf die spezifischen Belange eines Budgetierungsprozesses ausgerichtet ist. Fester Bestandteil dieser Lösungen ist auch ein Status-Monitor.

12.3.3 Planung auf verdichteten Strukturen

Die Planung auf verschiedenen Verdichtungsstufen erleichtert die vertikale Planabstimmung, da auf jeder Ebene Planinformationen eingegeben werden können. Typische Aggregationsebenen sind Bereiche, Kostenartengruppen oder Gesamtjahr. Es erfolgt hierbei kein automatisches Herunterbrechen der Werte auf die detaillierte Basisebene. Eine Vergleichsrechnung ist auch nur auf dieser Ebene (oder einer höheren Ebene) sinnvoll.

ERP-Systeme verlangen in der Regel die Eingabe auf der niedrigsten Stufe. Das macht die Planung häufig sehr schwerfällig, weil dann die Erfordernisse der Abrechnung gemeinsam mit den Erfordernissen einer Planung zusammengebracht werden müssen. In Folge sind die Planungseingaben detaillierter als eigentlich wünschenswert.

Die meisten OLAP-Systeme erfordern auch die Eingabe auf der niedrigsten Ebene. Es gibt allerdings „Faulenzer“-Funktionen, die die Eingabe auf verdichteter Ebene zulassen. Im Hintergrund werden dann die Daten nach vordefinierten Schlüsseln verteilt. Das kann die Planung durchaus erleichtern. Es gibt aber auch Konzepte über sogenannte „Schattenhierarchien“, welche die Eingabe auf verdichteter Ebene zu ermöglichen. Dies erfordert allerdings einen erhöhten Einrichtungsaufwand.

Planungsanwendungen lassen zum Teil die Eingabe auf Knoten zu. Falls nicht, können ähnliche Hilfskonstruktionen wie bei OLAP zum Einsatz kommen.

Die Tabellenkalkulation lässt sich flexibel aufbauen. Allerdings müssen Bottom Up- und Top Down-Sicht in Auswertungen auch zusammengebracht werden, was den Konfigurationsaufwand deutlich erhöht.

12.3.4 Validierung

Die Validierung ermöglicht die unmittelbare Überprüfung eingegebener Werte nach vorab definierten Kriterien. Häufige Kriterien sind unter anderem die vollständige Eingabe aller Positionen oder die Einhaltung bestimmte Grenzen für Kostenpositionen. Eine Validierung kann auch auf verteilte, erfasste Sachverhalte wie Kosten beim Empfänger = Umsatz beim Sender angewendet werden. Hier muss dann festgelegt werden, auf welcher Seite die Validierung durchgeführt wird.

ERP-Systeme besitzen in der Regel keine explizite Validierung für Planzahlen. Zum Teil lassen sich solche Regeln über Scriptsprachen programmieren.

In OLAP muss ein Validierungsmechanismus auch erst aufgesetzt werden. Die meisten OLAP-Systeme verfügen über leistungsfähige Regel-Maschinen, die die Formulierung

von Validierungen erleichtern. Dabei werden üblicherweise ein oder mehrere Dimensionselemente definiert, welche das Ergebnis von Prüfungen aufnehmen.

In der Tabellenkalkulation lassen sich Validierungen relativ leicht implementieren. Eine konditionale Formatierung weist auf Regelverletzungen hin. Ein Nachteil ist dabei jedoch, dass hierbei nicht auf zentral hinterlegte Informationen zugegriffen werden kann. Damit ist eine Validierung in Bezug auf die Abstimmung zwischen Leistendem und Empfänger kaum zu implementieren. Auch kann das Versenden von in Bezug auf die Validierung mangelhaften Planungsblättern trotz Validierung nicht verhindert werden.

Spezifische Planungslösungen unterstützen in der Regel die Validierung in den skizzierten wünschenswerten Ausprägungen.

12.3.5 Automatisierte Top-Down-Verteilungen

Automatisierte Top-Down-Verteilungen erleichtern die Planvorbereitung. Hierbei wird auf aggregierter Ebene ein Wert eingegeben, der nach bestimmten Regeln auf die untergeordneten Elemente verteilt wird. So können beispielsweise Istdaten als Basis herangezogen werden, die dann pauschal um einen bestimmten Prozentsatz angepasst werden. Eine wichtige Erweiterung ist das Sperren von Elementen. Wird beispielsweise ein übergeordnetes Element gesperrt, muss die Anpassung eines untergeordneten Elements dazu führen, dass die Schwesterelemente die Veränderung aufnehmen. Eine weitere Eigenschaft ist die simultane Verteilung über beliebig viele Versionen. So kann beispielsweise ein Jahreswert für eine Produktgruppe gleichzeitig auf Monate und Einzelprodukte heruntergebrochen werden.

ERP-Systeme sind durch die Bottom-Up-orientierte Planungsvorgehensweise nicht in der Lage, eine solche automatisierte Verteilung zu unterstützen.

Einige OLAP-Systeme wie beispielsweise TM1 unterstützen die Top-Down-Verteilung über beliebige Dimensionen. Auch ein Sperren von Elementen für die Verteilung wird unterstützt.

Eine Top-Down-Verteilung sucht man bei Tabellenkalkulationen vergeblich. Dies liegt auch daran, dass Dimensionen und Hierarchien keine spezifischen Objekte in der Tabellenkalkulation darstellen. Gelegentlich wird über Makros eine solche Funktionalität implementiert. Hier treten allerdings Bedenken in Bezug auf die Anpassbarkeit auf.

Bei spezifischen Planungswerkzeugen gehört diese Funktionalität in der Regel zum Standardumfang.

12.3.6 Simulationsrechnungen

Simulationsrechnungen unterstützen üblicherweise zwei wesentliche Verfahren: „What If“ und „How to Achieve“ (Zielwertsuche). Bei der What-If-Analyse werden Basisparameter geändert und die Auswirkung auf die Zielgröße(n) betrachtet. Bei der Zielwertsuche

wird der Zielwert variiert und betrachtet, welche Veränderungen bei den Basisparametern notwendig sind. Man muss außerdem zwischen der reinen Wertesimulation („was passiert, wenn sich der Kurs um x% ändert?“) und der Struktursimulation („wie verändert sich das Bereichsergebnis, wenn das Produkt X einem anderen Bereich zugeordnet werden würde?“) unterscheiden. Letztes ist schwieriger zu implementieren, weil Stammdatenänderungen auf Szenario-Ebene möglich sein müssen.

Zwar ist es in einigen ERP-Werkzeugen möglich, unterschiedliche Szenarien zu definieren und hier in einem gewissen Umfang Alternativrechnungen aufzubauen. Eine Simulationsrechnung ist allerdings zu schwerfällig, da Iststrukturen in einem hohen Detaillierungsgrad die Ausgangsbasis darstellen.

Die Simulation auf OLAP-Basis ist möglich. Allerdings beschränkt sich diese auf die „What if“-Analyse. Die Implementierung eines Goal-Seeking ist schwierig. Es gibt Ansätze, ein Interpolationsverfahren umzusetzen, dies ist bei komplexen Modellen allerdings eher langsam. Strukturveränderungen können teilweise durch das Kopieren eines ganzen Würfels in eine Szenario-Variante umgesetzt werden. Allerdings sind Vergleichsanalysen zwischen Ausgangs- und Simulationsvariante nur aufwändig zu implementieren.

Die Tabellenkalkulation ist für Simulationsrechnungen recht gut geeignet. Auch komplexe Szenarien lassen sich durch die flexible Modellierung leicht abbilden. Allerdings sind nachträgliche Auswertungen wie beispielsweise ein Szenario-Vergleich schwer abzubilden.

Einige Planungswerkzeuge unterstützen beide Arten der Simulation in der Regel umfassend. Allerdings werden Struktursimulationen nur von wenigen Werkzeugen unterstützt.

12.3.7 Vorbelegung

In den seltensten Fällen beginnt eine Planung bei Null im Sinne eines Zero-Based-Budgetings. Istdaten oder Vorscheurechnungen bilden meistens die Basis für die Planung. Insofern ist es wichtig, dass relativ leicht Planungsvorschläge erzeugt werden können. Dabei ist es mit einem einfachen Kopieren jedoch nicht getan. So kann die Vorscheurechnung je Position auf unterschiedlichen Verfahren basieren. Beispiele sind Trendextrapolationen, Ist/Budget des aktuellen Jahres usw.

Zum Teil verfügen ERP-Systeme über Vorscheaufunktionen, mit denen Planungsszenarien vorbelegt werden können.

In OLAP lassen sich Kopierfunktionen in der Regel relativ einfach implementieren. Hierbei können auch Justierungen vorgenommen werden.

Die Übernahme von Ausgangsdaten in eine Planungsversion kann in der Tabellenkalkulation aufwändig werden. „Cut and Paste“ wird häufig verwendet, ist aber fehleranfällig.

Planungslösungen stellen zur Fortschreibung in der Regel standardisierte Methoden bereit. Zum Teil existieren Module zur detaillierten Forecast-Ableitung.

12.3.8 Aggregation

Die automatische Zusammenführung von dezentral geplanten Daten ist eng mit der Art der Datenspeicherung verbunden. Wird direkt in eine zentrale OLAP-artige Datenbasis eingegeben, kann die Aggregation in der Regel automatisiert durchgeführt werden. Eine Frage betrifft allerdings den Zeitpunkt der Aggregation. Erfolgt dies unmittelbar nach Eingabe von Basisinformationen oder ist ein spezieller Prozess zu starten?

Außer der Tabellenkalkulation unterstützen eigentlich alle Systeme die direkte Verdichtung zufriedenstellend. Einige OLAP- und Planungssysteme unterstützen die Aggregation in Echtzeit.

Excel bietet die Möglichkeit, Daten aus gleichartigen Dateien zu verdichten. Allerdings ist das Verfahren fehleranfällig. Ist eine Datei beispielsweise leicht angepasst worden (z. B. durch eine neue Spalte), wird zwar aggregiert, aber unter Umständen mit einer völlig falschen Zuordnung. Viele solcher Probleme bleiben häufig länger unentdeckt.

12.3.9 Kommentierungen

Kommentierungen ermöglichen es, neben den Zahlen auch Erläuterungen hinzuzufügen. Dies kann insbesondere die Kommunikation zwischen den Bereichen verbessern. Es hilft aber auch, später beim Soll-Ist-Vergleich Abweichungen besser zu verstehen. Es ist notwendig, im Planungsprozess auf verschiedenen Ebenen Kommentare hinterlegen zu können. Daneben erscheint es auch wichtig, erläuternde Zusatzinformationen in Form von angehängten Dateien mitgeben zu können. Für die Kommentierungen gelten die gleichen Anforderungen an die Nachvollziehbarkeit wie für die eigentlichen Zahlen (s. u.): Autor und Änderungszeitpunkt sollten unveränderlich hinterlegt werden können. Es ist auch sinnvoll, mehrere Kommentare pro Zelle zuzulassen. Hiermit kann auch eine Diskussion abgebildet werden.

Die Bearbeitung von Texten ist eine Schwachstelle von ERP-Systemen. Zwar ist die Integration zum Content-Management schon weit fortgeschritten. Allerdings sind Planung und Budgetierung hiervon scheinbar noch entkoppelt.

OLAP-Systeme unterstützen die Kommentierung. Ein Kommentar wird in der Regel einer Zelle zugeordnet. Ein Vorteil von leistungsfähigen Regel-Sprachen ist es, dass die Anzahl der Kommentare auf unteren Dimensionselementen gezählt werden können. So kann man leicht erkennen, unter welchen Gruppenelementen Kommentierungen hinterlegt sind.

Auch Planungslösungen bieten in der Regel die Möglichkeit, umfassend zu kommentieren. Häufig werden dann auch spezielle Kommentarberichte angeboten, die alle Kommentare eines frei definierten Planungsausschnitts auflisten.

In einer Tabellenkalkulation können ebenfalls Kommentare hinterlegt werden. Zellen mit Kommentaren werden entsprechend markiert. Allerdings lassen sich die Kommentare nur schwer zu einer zentralen Datenbasis zusammenführen.

12.3.10 Eingabe über das Internet

Die Möglichkeit, über das Internet Planungsdaten einzugeben, ist insbesondere für internationale Unternehmen mit einer hohen Dezentralisierung von Bedeutung. In der Regel kann ein Planer bei web-basierten Anwendungen ohne größeren Installationsaufwand an das Planungssystem angebunden werden. Aber hier gibt es deutliche Qualitätsunterschiede: Wie genau können die Planungsmasken gestaltet werden? Kann die Validierung oder auch andere Funktionen (z. B. die Top Down Verteilung) genutzt werden? Eine flexible Internet-Eingabemöglichkeit erhöht die Akzeptanz gegenüber einer starren Eingabemaske.

Außer für die Tabellenkalkulation gibt es mittlerweile für die meisten Werkzeuge eine entsprechende Eingabeunterstützung über das Internet. Zwar lässt sich ein Tabellenkalkulationsblatt über einen Browser aufrufen. Der Zugriffsschutz ist dabei jedoch nur aufwändig sicherzustellen.

OLAP-Werkzeuge können häufig über unabhängige eingabefähige Berichtswerkzeuge beplant und analysiert werden. Allerdings ist der Eingabekomfort häufig eher begrenzt.

Planungssysteme unterscheiden sich in Bezug auf die Internetunterstützung zum Teil gravierend. Während einige Systeme lediglich die Eingabe von Zahlen in sehr einfachen Formularen ermöglichen, gibt es Systeme, bei denen der Anwender sich seine eigenen Planungsmasken auf der Basis einer Vorlage selbständig anpassen kann. Besonders interessant sind Werkzeuge, die die gleichen Planungsdesigns für Web- und Excel-Eingabe verwenden können.

12.3.11 Detailplanungen

Häufig orientiert man sich an zentral definierten Größen wie Konten- oder Kostenpositionsrahmen. In den dezentralen Bereichen wird jedoch häufig eher Maßnahmen-bezogen geplant. Einige Systeme bieten die Möglichkeit, unterhalb des zentral definierten Planungsrasters bereichsindividuell und dezentral freie Positionen anzulegen. Die Werteingaben dieser Positionen werden automatisch in das zentrale Planungsraster übergeleitet. Dieses Verfahren wird häufig als „Line Item Detail“ bezeichnet. Die Möglichkeiten einer solchen lokalen Modellerweiterung kann variieren: Im einfachsten Fall werden lediglich Detailpositionen eingeben und auf Planungsrasterebene verdichtet. Allerdings besteht häufig auch die Anforderung, komplett eigenständige Modelle lokal zu hinterlegen und deren Ergebnisse mit der zentralen Anwendung zu verlinken.

Detailplanungen sind in ERP in der Regel nicht möglich und können auch nicht verlinkt werden. Zum Teil werden Upload-Funktionen angeboten, mit denen lokale Planungsdaten in das System überspielt werden können.

In OLAP-Systemen ist es häufig schwierig, lokale Detailplanungen zu integrieren. Häufig lässt sich zumindest das Line Item Detail über eine zusätzliche Dimension implementieren. Diese Dimension enthält einen Zähler, so dass die lokalen Planer ihre Planungspositionen im Detail eingeben können.

Detailplanungen sind in Excel äußerst leicht zu integrieren. Allerdings verbirgt sich hier auch ein gewisses Risiko. Die Leichtigkeit der Änderung verführt häufig dazu, viele Ad hoc-Anpassungen durchzuführen. Dadurch leidet die Qualität.

Viele Planungslösungen bieten die Möglichkeit an, unterhalb von Planungspositionen Detailelemente zu hinterlegen. In der Regel muss ein entsprechendes Element als „Line Item Detail“ frei geschaltet werden. Dann kann der jeweilige Planer die Planungsgranularität frei festlegen.

12.3.12 Abstimmung des Leistungsnetzes

Ein häufiger Zeitfresser bei der Planung ist die Koordination zwischen leistendem und empfangendem Bereich. Abweichungen in den Planungen treten häufig erst bei der Verdichtung in der Zentrale auf. Eine dezentrale Abstimmung kann hier bei der Beschleunigung helfen. Hier gibt es unterschiedliche Verfahren: Gelegentlich wird die Übereinstimmung dadurch erzwungen, dass eine der Parteien den Betrag bzw. die Menge einträgt und diese automatisiert in die Planung des Partners übernommen wird. In einem anderen Fall planen die Partner den jeweiligen Bedarf/das Angebot. Der Einsatz der Validierung erscheint hier sinnvoll: Erst wenn die Differenz einen bestimmten Betrag nicht überschreitet, dürfen die Planungen freigegeben werden.

Mit der Tabellenkalkulation ist es schwierig, einen entsprechenden Prozess zu etablieren. Dadurch, dass die jeweiligen Planungsdaten verteilt vorliegen, ist eine Echtzeitabstimmung nicht möglich. Hinzu kommt die Vielzahl der möglichen Kontraktionsbeziehungen.

ERP-Systeme bieten teilweise komfortable Möglichkeiten der Bestimmung des Leistungsnetzes.

Eine Leistungsabstimmung ist in OLAP durch die zentrale Datenhaltung möglich. Allerdings muss ein entsprechender Prozess erst implementiert werden. Hierbei wird die Organisationsdimension gespiegelt, so dass Planungspositionen gleichzeitig mit liefernder oder empfangender Kostenstelle (-träger) eingegeben werden können.

Einige Planungssysteme bieten vorbereitete Prozessvorlagen zur Leistungsabstimmung an.

12.3.13 Übernahme von Stamm- und Bewegungsdaten aus ERP-Systemen

Der Aufbau von Planungslösungen ist häufig recht aufwändig. Die Möglichkeit, Stamm- und Bewegungsdaten aus ERP-Systemen zu übernehmen kann den Aufwand deutlich reduzieren. Hierbei ist allerdings auch zu berücksichtigen, dass die Strukturen für Planungen überarbeitet und auch erweitert werden müssen. Eine typische Situation ist die Anlage von

neu zu planenden Projekten oder Artikeln. Muss ein solches Objekt erst im ERP-System angelegt werden, kann dies den Planungsablauf verzögern.

Für die Planung mit ERP-Systemen ist dieser Punkt im engeren Sinne nicht relevant. Es ist allerdings zu betrachten, wie Informationen aus anderen Teilsystemen übernommen werden können. So erfordert beispielsweise die analytische Kostenplanung den Zugriff auf vielfältige Kostentreiber (z. B. Fertigungsstunden oder Rüstzeiten). Die Möglichkeit der freien definierbaren Übergabe ist nicht in allen ERP-Systemen gegeben.

OLAP-Systeme verfügen häufig über leistungsfähige Importwerkzeuge. Damit lassen sich planungsrelevante Informationen aus nahezu beliebigen Vorsystemen übernehmen. Die Herausforderung liegt meistens in der komplexen Struktur der Vorsysteme. Für die verbreiteten OLAP-Systeme stellen professionelle ETL-Anbieter Schnittstellen bereit. Für Planungssysteme, die auf OLAP-Datenbanken basieren, gilt das ausgeführte. Planungssysteme mit einer eigenen Datenhaltung stellen auch eigene ETL-Funktionen zur Verfügung, die aber häufig in der Funktionalität eingeschränkt sind.

Die größten Probleme ergeben sich bei der Informationsübernahme in die Tabellenkalkulation. Da die Tabellenkalkulation keine Dimensionen kennt, ist die Verarbeitung recht aufwändig. Man definiert in der Regel Transferbereiche. Das sind eigene Tabellenblätter, die die Importdaten aufnehmen. Über sogenannte Lookup-Funktionen werden dann den jeweiligen Planungsobjekten die entsprechenden Inhalte zugeordnet. Dies ist fehleranfällig und auch bei größeren Datenmengen zeitkritisch.

12.3.14 Verknüpfung von Teilplanungen

Ein umfassendes Planungssystem besteht aus diversen Teilplanungen. Diese Teilplanungen sind abhängig voneinander und müssen abgestimmt werden. Eine wichtige Frage hierbei ist der Zeitpunkt der Abstimmung. Stehen die Ergebnisse aus anderen Teilplanungen unmittelbar nach deren Aktualisierung zur Verfügung?

Einen hohen Aufwand in die Verknüpfung von Teilplanungen hat die SAP in ihr ERP gesteckt. Hieraus ist ein umfassend integriertes Planungssystem entstanden, welches sich allerdings durch eine hohe Komplexität auszeichnet. Da die Übergaben zwischen den Planungssystemen in der Regel nicht online erfolgt, ist die Einhaltung der Planungsreihenfolge äußerst wichtig. Die Schwerfälligkeit dieses Ansatzes mag auch ein Grund dafür gewesen sein, Planungsfunktionalität schon vor einiger Zeit in das BW auszulagern.

OLAP-Modelle lassen sich in der Regel leicht verknüpfen. Hier kann häufig auch zwischen einer dynamischen und der zeitpunktbezogenen Aktualisierung gewählt werden. Bei einer dynamischen Verknüpfung sind allerdings Auswirkungen auf die Performance zu berücksichtigen. Es ist nicht von der Hand zu weisen, dass Planungsmodelle auf der Basis von OLAP ohne Vordefinitionen sehr komplex werden können.

Noch problematischer ist das Arbeiten mit der Tabellenkalkulation. Auf die mangelnde Nachvollziehbarkeit ist schon hingewiesen worden. Häufig befinden sich Teilplanungen in unterschiedlichen Dateien. Zwar kann man üblicherweise auch auf Datei-externe

Tabellenblätter referenzieren. Wenn hierbei allerdings Aktualisierungsketten entstehen, kann dies zu Fehlern führen.

In typischen Planungssystemen sind Teilplanungen bereits definiert. Klassiker ist hier die integrierte Erfolgs- und Finanzplanung. Durch Bereinigung der Erfolgs- und Bilanzplanung kann der Zahlungsstrom hinreichend genau bestimmt werden. Sind solche Standardrechnungen bereits definiert, reduziert sich der Integrationsaufwand deutlich.

12.3.15 Abbildung abhängiger Planungsgrößen

Die Unterstützung abhängiger Planungsgrößen durch Regeln kann die Planung erheblich vereinfachen. Man fokussiert in dem Fall auf die Treiber und lässt Kosten und Erlöse hieraus automatisch ableiten. So könnte beispielsweise der Umsatz aus der aktuellen Pipeline abgeleitet werden. Wichtig ist dabei, von wem und wann solche Abhängigkeiten definiert werden. Bei einer zentralen Modellierung fällt es häufig schwer, alle spezifischen Abhängigkeiten berücksichtigen zu können. Dies würde auch zu einer eher schwerfälligen Modellierung führen.

In ERP-Systemen sind wenige Abhängigkeiten fest vordefiniert. In der analytischen Kostenplanung können beispielsweise Mengen und spezifische Verbräuche und Preise hinterlegt werden. Hieraus ermitteln das System die entsprechenden Kosten. Individuell definierte Planungsfunktionen können jedoch nicht ohne aufwändiges Customizing hinterlegt werden.

Die Regelsprachen von OLAP-Systemen ermöglichen den Aufbau beliebiger Abhängigkeiten. Allerdings sei hier auf die erhebliche Komplexität hingewiesen, die durch eine umfängliche Nutzung von Abhängigkeiten entstehen kann. Damit steigt die Fehleranfälligkeit. Beliebte Fehler, die häufig auftreten, sind beispielsweise die Addition von Preisen auf Produktgruppenebene oder die Verdichtung von Anlagenwerten von Monaten auf Jahresebene.

Planungssysteme bieten den Vorteil, dass vielfältige Abhängigkeiten bereits vordefiniert sind. So ist die Ermittlung von Cash Flow-Positionen in einigen Systemen bereits vordefiniert. Bei der Definition freier Abhängigkeiten unterscheiden sich die Systeme erheblich. Systeme, die auf OLAP-Datenbanken basieren, sind hier deutlich im Vorteil, da die zugrundeliegende Regelsprache zur Definition von Abhängigkeiten verwendet werden kann.

Die Abbildung von Abhängigkeiten ist bei Tabellenkalkulationen durch Verknüpfung von Zellen sehr einfach. Allerdings ist eine solche Verlinkung immer eine Einzelproduktion, d. h. für genau ein Element in einer Periode gültig. Wird diese Abhängigkeit mehrfach benötigt, muss kopiert werden. Dies ist insbesondere fehleranfällig bei nachträglichen Anpassungen der Strukturen.

12.3.16 Nachvollziehbarkeit

Einen letzten Punkt betrifft die Nachvollziehbarkeit der Plandaten. Die Planung ist die Grundlage der Lagebeurteilung der externen Berichterstattung. Es muss zu jedem Zeitpunkt sichergestellt sein, wer wann ein bestimmtes Informationsobjekt geändert hat. Dies sicherzustellen, ist zunächst einmal eine organisatorische Aufgabe. Allerdings kann das Planungssystem hier unterstützen, indem die Manipulation mit Datum und Anwender unveränderlich gespeichert wird.

ERP-Systeme speichern in der Regel Zusatzinformationen wie Anwender und Zeitpunkt der Änderung. Hierdurch lässt sich in der Regel sehr gut nachvollziehen, wer welche Daten eingegeben hat. OLAP-Systeme unterstützen ebenfalls das Logging-Konzept. Allerdings werden diese Informationen häufig nicht in der Datenbank selber sondern in einer eigenen Datei gespeichert. Eine Auswertung erfolgt dann über Zusatzwerkzeuge. Die Nachvollziehbarkeit ist damit grundsätzlich gesichert, allerdings ist die Auswertung meistens nicht sehr komfortabel. Viele Planungssysteme verfügen ebenso über einen Logging-Mechanismus. Die Nachvollziehbarkeit wird über integrierte Berichte sichergestellt. Die Nachvollziehbarkeit ist in Tabellenkalkulationen häufig schwierig. Zum einen sind verwendete Abhängigkeiten zum Teil nur schwer identifizierbar. Zum anderen werden Änderungsvermerke nur für die gesamte Datei geführt.

12.4 Zusammenfassende Bewertung

Die folgende Grafik zeigt die Leistungsfähigkeit der Ansätze. Der Grad der Unterstützung erfolgte subjektiv im Rahmen diverser Workshops des Arbeitskreises „Moderne Budgetierung“ des Internationalen Controller Vereins. Es wird auf die Ausführungen des vorherigen Abschnitts verwiesen.

Die Tabellenkalkulation schneidet insgesamt recht schwach ab. Lediglich bei Simulationsrechnungen kommen die Stärken der hohen Flexibilität zum Tragen (Abb. 12.2).

Das ERP-System erscheint in einigen Bereichen durchaus gut für die Planungsunterstützung geeignet zu sein. Schwachstellen sind immer dort zu identifizieren, wo es um Flexibilität geht (Abb. 12.3).

OLAP kann demgegenüber in den Bereichen der Flexibilität punkten. Schwachstellen bestehen in Bezug auf individuelle Teilplanungen (Abb. 12.4).

Spezifische Planungsapplikationen sind in Bezug auf die Funktionsabdeckung am leistungsfähigsten. Das überrascht nicht, schließlich werden diese Werkzeuge genau für diese Aufgabe entwickelt (Abb. 12.5).

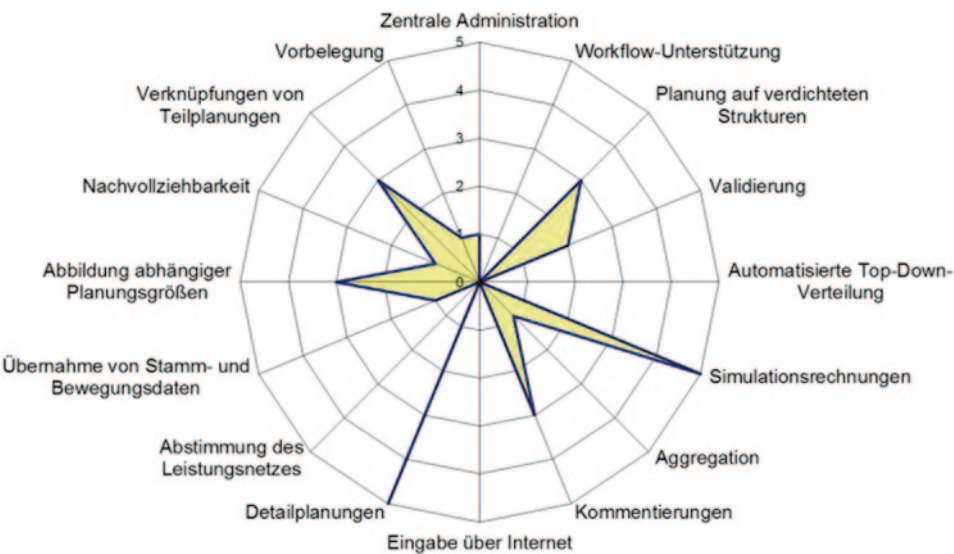


Abb. 12.2 Einschätzung der Tabellenkalkulation

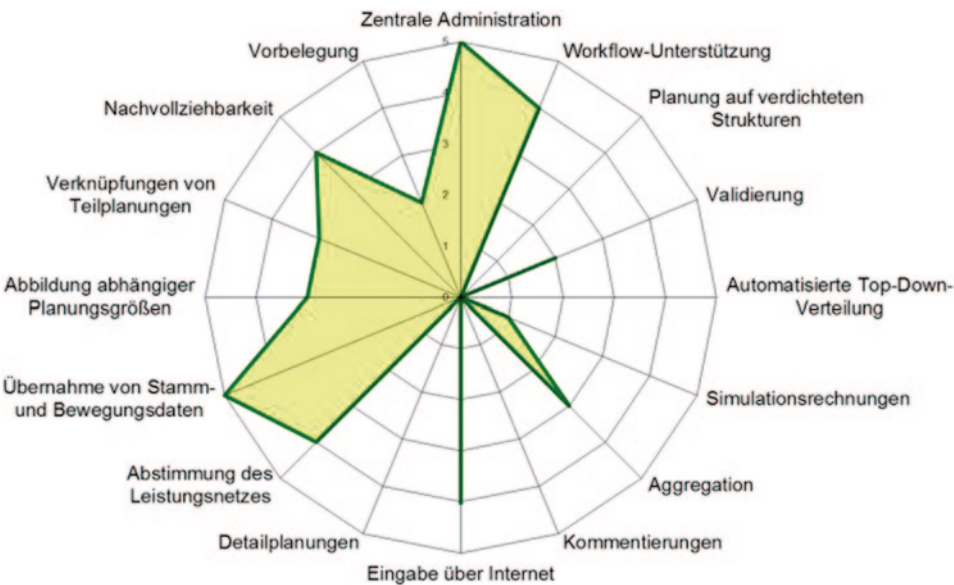


Abb. 12.3 Einschätzung des ERP-Systems

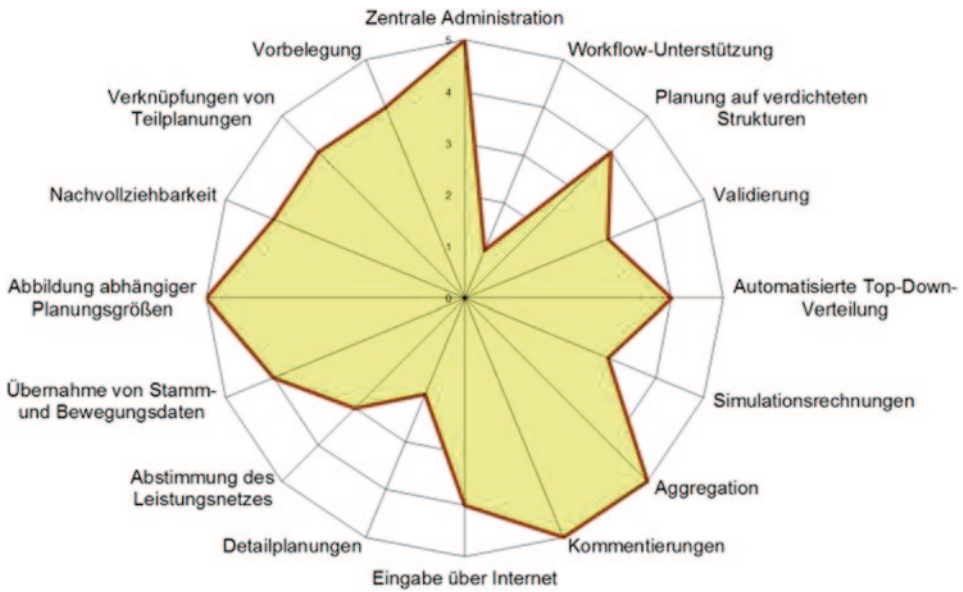


Abb. 12.4 Einschätzung von OLAP

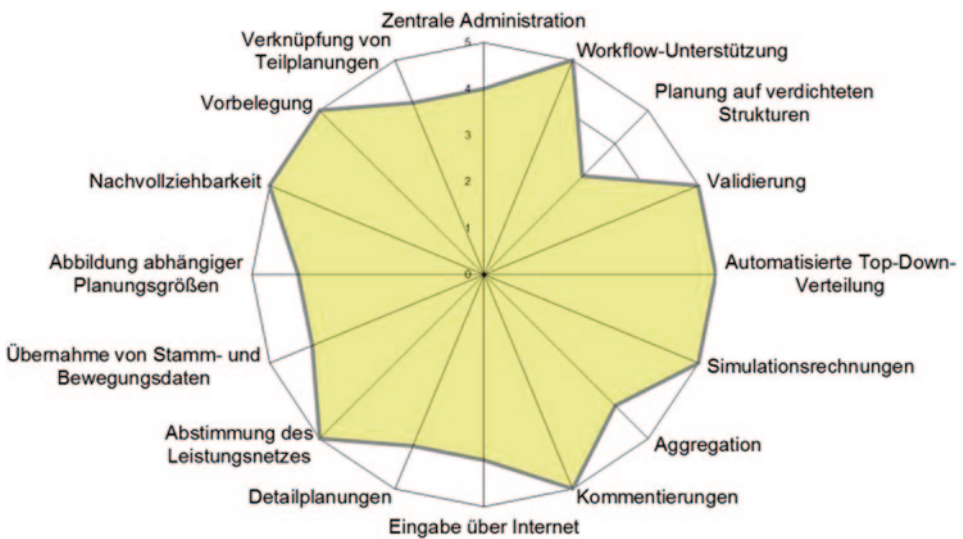


Abb. 12.5 Einschätzung der spezifischen Planungslösung

12.5 Ausblick

Spezielle Planungslösungen erscheinen am besten für Unternehmensplanungen geeignet zu sein. Allerdings ist das Ergebnis bei der konkreten Auswahl zu relativieren: Die Umfeldbedingungen (insbesondere Komplexität und Dynamik) müssen berücksichtigt werden und den anfallenden Kosten gegenübergestellt werden (im Detail (Oehler et al. 2011a) und (Oehler et al. 2011b)). Bei einer niedrigen Komplexität oder Dynamik können ERP oder Tabellenkalkulation durchaus funktional ausreichend sein. Schließlich sind ERP, Tabellenkalkulation und OLAP bei vielen Unternehmen bereits im Einsatz, so dass keine zusätzlichen Lizenzkosten anfallen bzw. auch ausreichend Expertise zur Einführung vorhanden ist.

Literatur

- Oehler, K.: Corporate performance management mit business intelligence technologien. Hanser, München (2006)
- Oehler, K. et al.: Systematische Auswahl von IT-Lösungen zur Planung (Teil 1). Control. Mag. 4, 28–34 (2011a)
- Oehler, K. et al.: Systematische Auswahl von IT-Lösungen zur Planung (Teil 2). Control. Mag. 5, 22–27 (2011b)
- Seufert, A., Oehler, K.: Unternehmenssteuerung und Wertbeitrag von Business Intelligence – Interview. In: Controllermagazin, vol. 34, Sonderheft Business Intelligence, S. 16–19 (2009)

Tom Hänel und Carsten Felden

Inhaltsverzeichnis

13.1 Einleitung 260

13.2 Entwicklung der IT in der Produktion bis zur Industrie 4.0 262

 13.2.1 Unterstützung technischer Aufgaben und Funktionen 262

 13.2.2 Integration von Aufgaben und Funktionen 263

 13.2.3 Integration von Prozessen und Wertschöpfungsketten 265

 13.2.4 Digitalisierung und globale Vernetzung 268

 13.2.5 Status Quo der Operational Business Intelligence 269

13.3 Industrie 4.0 270

 13.3.1 Technologische Grundlagen 271

 13.3.2 Das Zukunftsszenario „Smart Factory“ 275

 13.3.3 Operational Business Intelligence in der Industrie 4.0 278

13.4 Fazit 279

Literatur 280

Zusammenfassung

Steigende Dynamik und Komplexität sind dauerhafte Herausforderungen, denen sich Industrieunternehmen stellen müssen. Im deutschsprachigen Raum ist als Lösungsüberschrift Industrie 4.0 in die Diskussion getreten. Kerngedanke ist die technische

T. Hänel (✉) · C. Felden
Informationswirtschaft/Wirtschaftsinformatik, Technische Universität Bergakademie Freiberg,
Freiberg, Deutschland
E-Mail: tom.haenel@bwl.tu-freiberg.de

C. Felden
E-Mail: carsten.felden@bwl.tu-freiberg.de

Integration von cyberphysischen Systemen sowie von Technologien des Internets der Dinge und Dienste in die Arbeiten industrieller Prozesse. Dieser Einfluss richtet sich auch auf IT-Systeme, die für eine Analyse industrieller Prozesse und eine daran gekoppelte Entscheidungsunterstützung Verwendung finden. Neben industriegetriebenen Ansätzen wie Manufacturing Execution Systems oder Advanced Process Control wird auch Operational Business Intelligence in diesem Kontext als Nutzen stiftender IT-Ansatz diskutiert. Eine Einordnung der Operational Business Intelligence in das Zukunftsbild der Industrie 4.0 ist zum gegenwärtigen Zeitpunkt noch nicht erfolgt. Der Beitrag nimmt daher ausgehend vom aktuellen Status Quo eine zukünftige Positionsbestimmung vor. Es wird aufgezeigt, dass die konzeptionellen Bestandteile der OpBI auch in der Industrie 4.0 der Entscheidungsunterstützung beitragen.

13.1 Einleitung

Die Leistungsfähigkeit produzierender Unternehmen hängt von ihrer Wettbewerbsfähigkeit auf zunehmend globalisierten Märkten ab. Aus diesem internationalen Wettbewerbsdruck ergibt sich die Herausforderung, fortwährend innovativ zu sein. Eine Konsequenz daraus sind komplexe Produkte sowie aufwändige Herstell- und Logistikprozesse, die einem stetigen Wandel unterliegen. Diese Dynamik wird im Industriesektor durch eine Variantenvielfalt der angebotenen Produkte verstärkt, um individuellen und differenzierten Kundenbedürfnissen gerecht zu werden. Hinzu kommen volatile Bedarfe und abnehmende Produktlebenszyklen. Bundesministerium für Bildung und Forschung (2014a) Als möglicher Lösungsansatz für die skizzierten Herausforderungen wird im deutschsprachigen Raum ein Zukunftsprojekt unter der Überschrift Industrie 4.0 diskutiert. Dieser Ausdruck wird mit einer revolutionären Umgestaltung der industriellen Produktion verbunden. Industrie 4.0 ist aus der Perspektive der Informationstechnik (IT) konzipiert und thematisiert ein Zusammenwachsen von zeitgemäßen IT-Technologien mit industriellen Prozessen (Kagermann et al. 2013). Die Nutzung von informationsverarbeitenden Systemen und deren Integration ist dabei im Produktionskontext von einer dauerhaften Entwicklung gekennzeichnet. Der gegenwärtige Entwicklungsstand schließt Automatisierungssysteme, rechnergestützte Werkzeuge und ganzheitliche Ansätze zur Produktentwicklung sowie operative und dispositive IT-Systeme ein (Westkämper 2013). Darüber hinaus existieren Konzepte für eine IT-Unterstützung produktionsspezifischer Entscheidungen (Lasi 2012). Auf Grund dynamischer und vernetzter Produktionsstrukturen benötigen auch Entscheidungsträger im Produktionsumfeld IT-Systeme, die eine Bearbeitung komplexer Aufgabenstellungen systematisch mit Informationen anreichern. In diesem Kontext wird neben ingenieurgetriebenen Ansätzen wie Manufacturing Execution Systems (MES) auch Operational Business Intelligence (OpBI) als IT-basiertes Konzept zur Entscheidungsunterstützung in der Produktion diskutiert (Hänel und Felden 2011). OpBI schließt Funktionalitäten für eine Erfassung, Konsolidierung und Auswertung von Produktionsdaten ein, um betriebswirtschaftliche und technische Analysen im Hinblick auf operative oder

planerische Entscheidungen zu unterstützen (Hänel und Felden 2012). Ein Zukunftsthema wie Industrie 4.0 impliziert neben veränderten Produktionsabläufen und zeitgemäßen Technologien ebenso eine Neuorganisation oder Weiterentwicklung dieser bestehenden Systementwürfe für eine Entscheidungsunterstützung. Das Ziel des Beitrages ist daher die Untersuchung von möglichen Auswirkungen und Veränderungen hinsichtlich der Relevanz und der informationstechnischen Gestaltung von OpBI-Systemen.

Industrie 4.0 thematisiert die intelligente Vernetzung der an der industriellen Produktion beteiligten Ressourcen, Systeme und Prozesse durch Technologien des Internets (Kagermann et al. 2013). Zur Steuerung komplexer Industrieprozesse werden cyberphysische Produktionssysteme (CPPS) geschaffen, deren Komponenten, Objekte und Geräte eigenständig miteinander kommunizieren können. Bereits vorhandene Produktionssysteme (z. B. Maschinen, Roboter) und Betriebsmittel werden mit Sensoren, Aktoren, Mikrocontrollern und Leistungselektronik ausgestattet, um dezentrale, sich selbst regulierende Steuerungs- und Regelungsprozesse zu ermöglichen (Bauernhansel 2014). Im Rahmen des Zukunftsthemas Industrie 4.0 sind damit Auswirkungen auf die Gestaltung und das Management von Information- und Datenflüssen in den Produktionsunternehmen zu erwarten. Kennzeichnend für die zu schaffenden CPPS-Lösungen sind horizontal und vertikal integrierte Produktionssysteme sowie durchgängige Technologiekonzepte über den gesamten Lebenszyklus der entstehenden Produkte. Die Grundlage für eine Planung, Simulation, Beschreibung und Bewertung anwenderspezifischer CPPS-Lösungen bilden interdisziplinäre Datenmodelle, die sowohl den Ist- als auch den Soll-Zustand eines Produktionsunternehmens darstellen. Die entsprechende Datenorganisation umfasst Methoden, um entscheidungsrelevante Informationen systemübergreifend zu sammeln, zu analysieren und Entscheidern bereitzustellen (Bundesministerium für Bildung und Forschung 2014b). Diese Herangehensweise ist durch eine Überlappung mit dem OpBI-Konzept gekennzeichnet. OpBI unterstützt zunächst einmal, unabhängig von einem bestimmten Anwendungsgebiet, die Abwicklung von täglichen Geschäftsprozessen infolge einer direkten und zeitnahen Verbindung dispositiver und operativer IT-Systeme (Gluchowski et al. 2009). Eine derartige Systemintegration unterstützt im Kontext einer hierarchischen Gliederung von Produktionssystemen die Verarbeitung, Verwaltung und den Austausch komplexer Informationen in der Produktions- und Prozessleitebene (Langmann 1996). Infolge der dezentralen und eigenständigen Organisation von CPPS werden diese Ebenen für eine informationstechnische Strukturierung und damit die aktuelle Einordnung von OpBI infrage gestellt (Schlick et al. 2013). Der Beitrag nimmt daher eine zukünftige Positionsbestimmung von OpBI in der IT-Systemlandschaft der Industrie 4.0 vor. Er trägt der sachgebietsübergreifenden Diskussion um intelligent vernetzte Produktionsstrukturen auf Basis von internetbasierten Technologien aus der Perspektive einer operativen Entscheidungsunterstützung bei.

Im Verlauf des Beitrags wird in Kap. 2 die Entwicklung der IT-Landschaft im Produktionsumfeld vorgestellt. Weiterführend erfolgt die Darstellung des Status Quo von OpBI für eine produktionsspezifische Entscheidungsunterstützung. Kapitel 3 geht auf die technologischen Grundlagen und das Zukunftsbild der Industrie 4.0 ein. OpBI wird in dieses

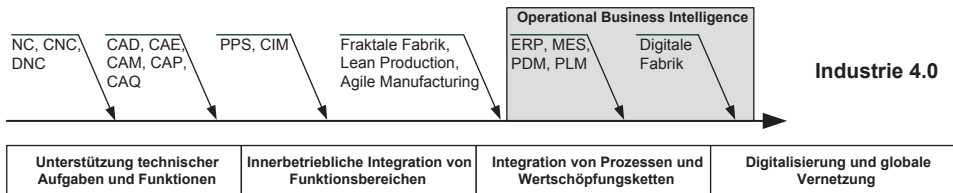


Abb. 13.1 IT-Unterstützung in der Produktion im Zeitverlauf

futuristische Szenario eingeordnet. Kapitel 4 fasst den Beitrag zusammen und liefert einen Ausblick.

13.2 Entwicklung der IT in der Produktion bis zur Industrie 4.0

In der Industrie kommen verschiedenartige IT-Systeme zum Einsatz, die typische Funktionen der Produktion und der Produktentwicklung ausführen, aufzeichnen, modellieren oder beeinflussen. Diese IT-Systeme haben im Zeitverlauf verschiedene Begriffe und Ansätze begleitet, die in einer groben zeitlichen Einordnung in der nachfolgenden Abbildung dargestellt sind. Der Zweck dieser Darstellung ist die überblicksartige Vorstellung der IT-Unterstützung in der Produktion im Vorfeld der Diskussion um Industrie 4.0 (vgl. Westkämper 2013; Jacobi 2013; Bracht et al. 2011). Es werden keine ineinander greifenden zeitlichen Entwicklungsstufen veranschaulicht. Für die nachfolgende Erläuterung sind die IT-Systeme und konzeptionellen Ansätze vier Phasen zugeordnet (Abb. 13.1).

13.2.1 Unterstützung technischer Aufgaben und Funktionen

Die computerbasierte Maschinensteuerung kennzeichnet frühe Ansätze zur Nutzung von IT für die industrielle Fertigung. Sie hat ihren Ursprung in den 1950er Jahren. Die maschinelle Bearbeitung von Werkstücken ist eine charakterisierende Aufgabe in der Fertigung. Eine Maschine benötigt Schalt- und Weginformationen, um mithilfe von geometrischen und technologischen Funktionen ein Werkstück zu bearbeiten. In diesem Zusammenhang steht NC für Numerical Control und ist eine Methode zur automatischen Maschinensteuerung auf Basis numerischer Daten (Buchstaben, Ziffern, Sonderzeichen). Diese Daten spiegeln Anweisungen zur Fertigung eines Werkstückes wider und ergeben entsprechend der Bearbeitungsreihenfolge ein NC-Programm (Kief und Roschiwal 2013). Derartige Steuerungen lassen sich in drei Kategorien einteilen (Werner 2010):

- Konventionelle numerische Steuerung (NC),
- Computerisierte numerische Steuerung (CNC) und
- Verteilte numerische Steuerung (DNC).

Konventionelle NC-Steuerungen sind verbindungsprogrammiert und bestehen aus elektronischen oder elektromechanischen Bauteilen (z. B. Relais oder kontaktlose Schaltungen), die fest miteinander verdrahtet sind. Die Anpassungsfähigkeit dieser Steuerungen ist eingeschränkt, weil neue Schaltungen und damit verbunden Änderungen an den festen Verbindungen erforderlich werden (Karaali 2013). CNC-Steuerungen weisen eine höhere Flexibilität als konventionelle NC-Steuerungen auf. Zur Ausübung der Steuerungsfunktion verwenden CNC einen lokalen Rechner, der die numerischen Daten abspeichert. Im Zuge von modifizierten Aufgabenstellungen müssen nur die Speicherinhalte geändert werden. Die Anforderungen der jeweiligen Steuerung werden durch Softwareprogramme gedeckt. Das Steuerungsprinzip von DNC ist ähnlich. Verteilte numerische Steuerungen nutzen im Unterschied zu CNC einen gemeinsamen Rechner für eine Vielzahl an numerisch gesteuerten Maschinen (ISO 2806 1996).

NC, CNC und DNC haben sich als Automatisierungstechniken zur Maschinensteuerung bewährt und werden unter dem Begriff des *Computer Aided Manufacturing* (CAM) zusammengefasst. Vorgelagert zur Fertigung sind Aufgaben der Konstruktion und Arbeitsplanung von Bedeutung, die sich ebenfalls durch Rechentechnik unterstützen lassen. CAD steht in diesem Zusammenhang für *Computer Aided Design* und unterstützt die Anfertigung von Konstruktionszeichnungen. Ziel ist die Erstellung, Veränderung oder Erweiterung eines Produktmodells. Die rechnergestützte Auslegung von Produkten sowie deren Komponenten während der Entwicklungs- und Konstruktionsphase wird als *Computer Aided Engineering* (CAE) bezeichnet. In den Bereich CAE fallen auch die Programme der *Finite-Elemente Methode* (FEM). Dies dient der Berechnung von Strukturparametern (Festigkeit, Stabilität) von geometrischen Körpern. Durch Zerlegung dieser Körper entsteht das Finite-Elemente Modell. CAE hat eine enge Verbindung zu CAD und wird daher als Ergänzung zu CAD-Systemen verstanden. *Computer Aided Planning* (CAP) unterstützt die Arbeitsplanerstellung, indem mithilfe von IT die Reihenfolge der Fertigungsschritte für ein Produkt festgelegt wird. CAD, CAP und CAM werden von einer computergestützten Qualitätssicherung permanent begleitet. Dieses präventive Qualitätsmanagement ist Aufgabe der *Computer Aided Quality Assurance* (CAQ). Diese aufgabenorientierten Komponenten werden unter dem Begriff der CAx-Systeme zusammengefasst (Hehenberger 2011).

13.2.2 Integration von Aufgaben und Funktionen

Neben der alleinigen Nutzung zur Ausführung von einzelnen Fertigungsaufgaben wurde IT im Zeitverlauf für eine übergreifende Aufgabenunterstützung relevant. Gemäß der Definition von Gutenberg kombiniert die Produktion die elementaren Faktoren Arbeit, Material und Betriebsmittel im Rahmen von Planungs- oder Organisationsaufgaben um Waren und Dienstleistungen herzustellen (Gutenberg 1983). Zusätzlich zur eigentlichen Leistungserstellung kennzeichnen die Verfügbarkeit und der Verbrauch von Eingangsressourcen sowie der Absatz der Sachgüter und Dienstleistungen Aufgabenbereiche in der

Produktion (Plümer 2003). In diesem Zusammenhang verfolgt die Produktionsplanungs- und Steuerung (PPS) eine Planung, Durchführung, Überwachung und Korrektur des Produktionsablaufs unter Berücksichtigung von betrieblichen Zielen (Zäpfel 2001).

Die Computerunterstützung im Rahmen von PPS-Systemen für die operative Planung und Steuerung von Produktionsabläufen reicht bis in die 1960er Jahre zurück. Erste Anwendungssysteme dienten der Materialbedarfsplanung (*Material Requirements Planning*) und wurden als MRP-Systeme bezeichnet. Durch die Erweiterung um Stufen für eine hierarchische Planung von Absatz-, Produktions- und Bestandsmengen entstand das *Manufacturing Resource Planning* (MRP II). Dies wurde im deutschsprachigen Raum als PPS aufgefasst (Kurbel 2005).

PPS-Systeme setzen auf einer Grunddatenverwaltung von Stücklisten, Arbeitsplänen und Betriebsmitteldaten auf und gehen in der Planung sukzessiv vor. Die Planungsaufgaben werden in verschiedene Ebenen untergliedert und stufenweise durchlaufen. Zunächst ermittelt eine Hauptproduktionsprogrammplanung aus Kundenaufträgen Primärbedarfe. Diese bilden gemeinsam mit Stücklisten die Grundlage für die Mengenplanung von Endprodukten oder Baugruppen. Damit entstehen Losgrößen für die Produktion, die gemeinsam mit den Arbeitsplänen in die Terminplanung einfließen. Nach Abschluss der Planung liegen Produktionsaufträge für eine Produktionssteuerung vor (Stadtler 2008). Der Übergang zwischen Planung und Steuerung ist fließend und kennzeichnet sich durch die Durchsetzung von Planvorgaben. Die Produktionssteuerung veranlasst, überwacht und sichert die Durchführung von Aufgaben in Bezug auf Mengen, Termine, Qualität und Kosten. Korrekte und aktuelle Rückmeldungen können über eine Betriebsdatenerfassung (BDE) erreicht werden (Zäpfel 2001).

PPS erfüllen vorrangig betriebswirtschaftliche Aufgaben im Kontext des Computer Integrated Manufacturing (CIM). CIM ist ein Ansatz zur Integration der Informationsverarbeitung im Industriebetrieb und hat seinen Ursprung zu Beginn der 1980er Jahre (Jacobi 2013). Ein konzeptioneller Rahmen ist das Y-Integrationsmodell nach Scheer.

CIM sieht vor, während der Planung und der Realisierung von Produktionsaufgaben betriebswirtschaftliche und technische Funktionen (CAx) miteinander zu verknüpfen. Die Verbindung kommt durch gemeinsame Aufgabenobjekte zustande. Die Vorgänge der Produkterstellung durchlaufen verschiedene kaufmännische und ingenieurtechnische Aufgaben, deren Bearbeitungszusammenhänge die Anforderungen an eine Integration der jeweiligen IT-Systeme begründen. Damit ist der Aufbau einer Datenbasis verbunden, die alle Funktionen gemeinsam nutzen können. Durch die Integration der IT-Systeme verfolgt CIM die ganzheitliche Betrachtung von Produktionsprozessen mit dem Ziel sowohl die Kundenorientierung als auch die Reaktionsfähigkeit eines Unternehmens zu verbessern (Scheer 1990).

Die Integration der Systeme im Rahmen von CIM kann als anspruchsvoll für Unternehmen bezeichnet werden. Eine vollständige Umsetzung von CIM wurde durch technische, organisatorische und wirtschaftliche Gründe beeinträchtigt. Technische Hürden lassen sich auf fehlende leistungsfähige Systeme zur Datensammlung und -übertragung sowie eine unzureichende Standardisierung zurückführen. Organisatorisch verlangt CIM

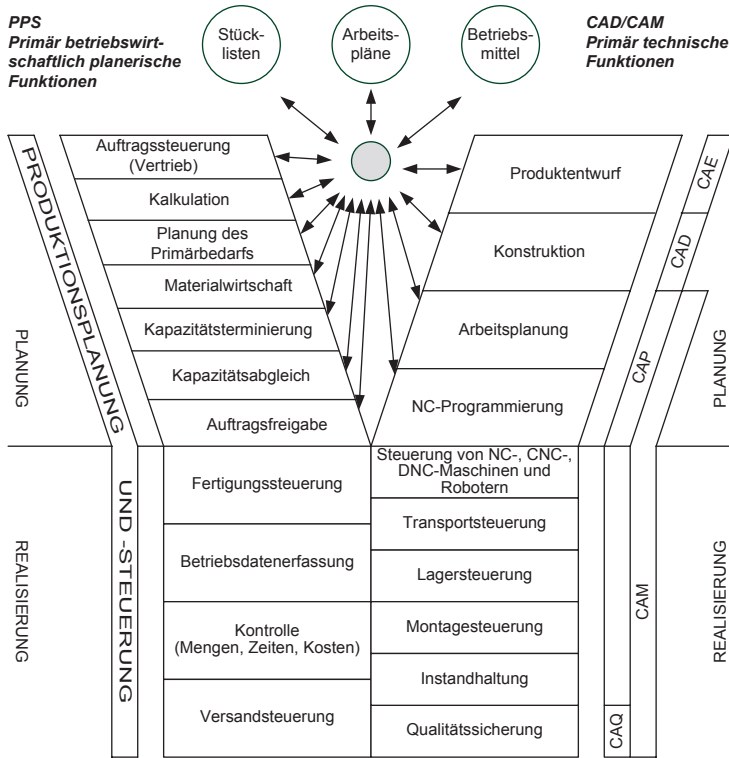


Abb. 13.2 Y-Integrationsmodell. (Scheer 1990)

bereichsübergreifende Arbeitsorganisationen und interdisziplinäre Herangehensweisen. Als schädigend wird in diesem Zusammenhang die Vernachlässigung des Faktors Mensch in CIM-Umsetzungen diskutiert. Aus wirtschaftlichen Gesichtspunkten ist die Inflexibilität hochautomatisierter Fertigungsstrecken, wie sie im Rahmen von CIM-Bestrebungen verfolgt wurde, nachteilig. Der Betrieb und die Instandhaltung derartiger Anlagen stellten einen unverhältnismäßigen Aufwand dar, weil Änderungen an Produkten oder Produktionsprozessen mit kostenintensiven Umstellungen verbunden sind (Abb. 13.2) (Bracht et al. 2011; Soder 2014).

13.2.3 Integration von Prozessen und Wertschöpfungsketten

Im weiteren Zeitverlauf sind im Anschluss an CIM organisatorische Konzepte sowie die separate Weiterentwicklung der betriebswirtschaftlichen und technischen IT-Systeme in den Mittelpunkt getreten (Abramovici und Schulte 2005). Durch die Diskussion des *Business Reengineering* (Hammer und Champi 1996) Anfang der 1990er Jahre hat die Prozessorientierung sich sowohl auf die Organisationsstrukturen als auch auf die Gestaltung der IT-Systeme ausgewirkt.

Organisationskonzepte CIM hat seinen Schwerpunkt auf die Nutzung von Werkzeugen und Methoden der Informationstechnik zur Vernetzung von Funktionsbereichen im Zusammenhang mit der Produktion gelegt. Das Ziel ist die Steigerung der Produktivität und Effizienz von Industriebetrieben. Ein Grund, dass CIM diesem Ziel nur teilweise beitragen konnte, ist die unzureichende Berücksichtigung von personellen und organisatorischen Aspekten.

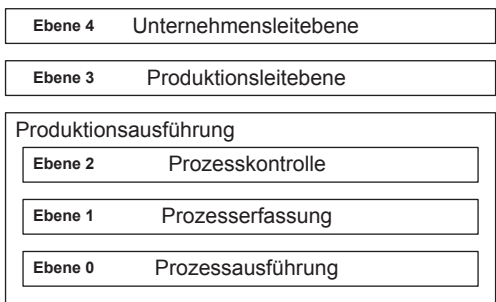
Die Diskussion um *Lean Production* (LP) greift diese Gesichtspunkte bezüglich der Arbeitsorganisation Anfang der 1990er Jahre im europäischen Raum auf (Kuhn 2008). LP stellt dabei ein integriertes sozio-technisches System dar. Das wesentliche Ziel ist die Vermeidung von Verschwendungen in der Produktion. Dazu wird die gleichzeitige Reduktion und Minimierung von internen, kundenspezifischen und lieferantenseitigen Schwankungen verfolgt (Shah und Ward 2007). LP kennzeichnet sich durch ein Prozessdenken über die gesamte Wertschöpfungskette, eine Prozessverantwortung und Kompetenz der Mitarbeiter, einer umfassenden Qualitätssicherung, einer beständigen Orientierung nach Märkten und Kunden, einer Einbindung von Zulieferern sowie einer kontinuierlichen Verbesserung der Produkte und Prozesse (Bullinger 2001).

Der Ursprung von LP geht auf das Produktionssystem des japanischen Automobilherstellers Toyota aus den 1950er Jahren zurück. Als US-amerikanische Reaktion ist der Produktionsansatz des *Agile Manufacturing* (AM) entstanden. Während LP operative Methoden für einen optimalen Ressourceneinsatz anbietet, stellt Agilität einen unternehmensweiten strategischen Ansatz dar (Sanchez und Nagi 2001). Das Ziel ist die möglichst zeitnahe Erfüllung von sich ständig ändernden Kundenbedürfnissen. Die zu Grunde liegenden Konzepte des AM sind die Konzentration auf Kernkompetenzen, Anpassungsfähigkeit, wissensbasierte Methoden sowie die unternehmensübergreifende Zusammenarbeit durch eine Virtualisierung von Unternehmensanwendungen (Yusuf et al. 1999).

Im europäischen Raum boten LP und AM gleichermaßen eine Orientierung für Produktionsunternehmen. Darauf aufbauend wurde 1991 mit der fraktalen Fabrik ein weiteres Organisationskonzept zur Gestaltung und Steuerung von produzierenden Unternehmen vorgestellt. Fraktale sind selbständige Unternehmenseinheiten, die eindeutige Ziele und Leistungen eigenverantwortlich organisieren und optimieren. Sie agieren im Einklang mit den Zielen des Gesamtunternehmens und sind durch Informationssysteme untereinander vernetzt. Die Organisationsstrukturen der fraktalen Fabrik sind dynamischen Einflussgrößen unterworfen, die wechselseitig ineinander greifen (Warnecke 1993). Eine flächendeckende Durchdringung konnte das Konzept der fraktalen Fabrik allerdings nicht erreichen (Jacobi 2013).

IT-Systeme in der Dispositions- und Steuerungshierarchie Die Funktionalitäten der von CIM ursprünglich adressierten IT-Systeme wurden in neo-logische Systemkonzepte übernommen und weiterentwickelt. Aus den PPS-Systemen sind im Zeitverlauf Enterprise Resource Planning (ERP)-Systeme entstanden. Dies stellt für Industriebetriebe eine integrierte Standardsoftware dar, die modular aufgebaut ist und auch an die Produktion angrenzende Aufgaben unterstützt. Damit ermöglichen ERP-Systeme eine gemeinsame Sicht auf Produktion, Finanz- und Rechnungswesen, Personalwesen sowie Vertrieb und Rechnungswesen. In einem unternehmensweiten Anwendungssystem werden Geschäftsprozesse transaktionsorientiert über eine gemeinsame Datenbank koordiniert und integriert (Kurbel 2005).

Abb. 13.3 Funktionale Hierarchie für Aktivitäten in der Produktion. (entnommen und modifiziert nach: ISO 62264–3 2008)



ERP-Systeme sind nicht nur auf produktionsspezifische Aufgaben beschränkt. Bei der Einordnung von Leittechnik in die industrielle Fertigung werden ERP-Systeme in die Unternehmensleitebene eingeordnet, weil diese bereichsübergreifend eine mittel- bis langfristige Planung unterstützen. ERP-Systeme haben jedoch keine Steuerungsmechanismen, die eine Rückkopplung zwischen den Planvorgaben und dem tatsächlichen Fertigungsprozess herstellen. Manufacturing Execution Systems (MES) sind durch die Weiterentwicklung elektronischer Leitstände hervorgegangen, um diese Lücke zu schließen. MES werden als Informationssysteme zwischen Unternehmensplanung und der Fertigungsausführung in die Produktionsleitebene eingeordnet. Sie stellen eine Verbindung zwischen den technischen Steuerungssystemen der Fertigung und der betriebswirtschaftlichen Planung her. Die Grobpläne aus den ERP-Systemen werden auf Stunden- bis Minutenebene herunter gebrochen. MES erfassen und bereiten die Ist-Daten der Fertigung auf, um eine fortlaufende Kontrolle und Anpassung der Planungsgrößen zu erreichen. Unterstützt werden Funktionen in den Bereichen der Produktionsausführung, Lagerverwaltung, Instandhaltung und Qualitätsmanagement (Thiel et al. 2008).

Die Fertigungsausführung wird durch ingenieurtechnische Informationssysteme unterstützt und lässt sich in drei generische Ebenen unterteilen. Die unterste Ebene charakterisiert den realen Herstell- oder Produktionsprozess. Darüber werden Sensoren und Aktoren zur Messung, Überwachung und Einflussnahme auf das Prozessgeschehen eingeordnet. Die oberste Ausführungsebene fasst Instrumente und Rechneinheiten zur Stabilisierung und Kontrolle der Prozessabläufe zusammen. Die hierarchische Strukturierung findet sich insbesondere bei der Automatisierung komplexer technischer Prozesse wieder. Im MES-Kontext ergibt sich daraus eine Vielzahl heterogener Schnittstellen, die für eine Integration von betriebswirtschaftlicher Planung und technischer Ausführung berücksichtigt werden müssen (Abb. 13.3) (ISO 62264–3 2008).

Produktorientierte IT-Systeme Ebenso wie die dispositiven Systeme haben sich die rechnergestützten Verfahren im Kontext der Produktentwicklung im Zeitverlauf weiterentwickelt. Dabei wurden die aufgabenorientierten CAX-Systeme durch zentrale IT-Lösungen in ein einheitliches Produktdatenmanagement (PDM) integriert. Gemäß der VDI 2219 steht der Begriff synonym zu Engineering Data Management (EDM). EDM/PDM-Systeme dienen der Speicherung, Verwaltung und Kommunikation von produktdefinierenden Daten (z. B. technische Zeichnungen, digitale Modelle, Textdokumente).

Produkte, zugehörige Entstehungsprozesse sowie deren Lebenszyklen werden über ein Produktdatenmodell informationstechnisch abgebildet. Dies verbindet schnittstellenbasiert die notwendigen Anwendungs- und CAx-Systeme zu einem Gesamtsystem für die Produktentwicklung. EDM/PDM-Systeme kommunizieren als Bestandteil der IT-Infrastruktur mit den dispositiven und steuerungstechnischen Systemen (VDI 2219 2002).

Product Lifecycle Management (PLM) ist eine Erweiterung der EDM/PDM-Systeme und berücksichtigt sämtliche Informationen aus dem Lebenslauf von Produkten. PLM wird als Sammelbegriff für Management-Ansätze und IT-Systeme im ingenieurtechnischen Bereich verstanden (Abramovici und Schulte 2005). Es führt die Entwicklungen im Bereich des Produkt- und Prozessdatenmanagements mit den Techniken der Produktmodellierung und Prozessdigitalisierung zusammen. PLM umfasst Methoden, Modelle und IT-Werkzeuge, um Produktinformationen, ingenieurtechnische Prozesse und Anwendungen über alle Phasen des Produktlebenszyklus zu organisieren. Es handelt sich um einen übergreifenden Ansatz für eine interdisziplinäre Zusammenarbeit zwischen Produzenten, Zulieferern, Partnerunternehmen und Kunden (Abramovici 2007).

13.2.4 Digitalisierung und globale Vernetzung

Ende der 1990er Jahre entstand der Begriff der digitalen Fabrik (DF), der gemäß VDI 4499-1 wie folgt definiert ist:

Die Digitale Fabrik ist der Oberbegriff für ein umfassendes Netzwerk von digitalen Modellen, Methoden und Werkzeugen – u. a. der Simulation und der dreidimensionalen Visualisierung –, die durch ein durchgängiges Datenmanagement integriert werden.

Ihr Ziel ist die ganzheitliche Planung, Evaluierung und laufende Verbesserung aller wesentlichen Strukturen, Prozesse und Ressourcen der realen Fabrik in Verbindung mit dem Produkt (VDI 4499-1 2008).

Die DF wird als Bestandteil des PLM aufgefasst und unterstützt die Produktlebenszyklusphasen der Produktentwicklung, Produktionsplanung sowie der Produktherstellung und Produktion (Schack 2008). Einen Teilbereich der DF bildet der digitale Fabrikbetrieb, der durch das Zusammenwirken von Fertigungsmanagementsystemen, virtuellen oder realen Anlagensteuerungen sowie virtuellen oder realen Maschinen und Anlagen gekennzeichnet ist. Von der Montage- und Fertigungsprozessplanung bis zur Serienproduktion werden die Phasen eines Produktsystems horizontal über ein durchgängiges Datenmanagement integriert (VDI 4499-2 2011).

Die DF bildet die reale Fabrik mithilfe von rechnergestützten Werkzeugen, Methoden und Technologien digital ab. Das Konzept der *Virtuellen Fabrik* ermöglicht es, Aussagen über das Verhalten des erzeugten Abbilds auf Grundlage von Simulationen zu generieren. Auf diese Weise können Produktionsabläufe in der tatsächlichen Komplexität für eine Erprobung oder Verbesserung virtuell nachgebildet werden. Die Ergebnisse der Simulation ermöglichen Adaptionen des digitalen Modells sowie eine Umsetzung zukünftiger

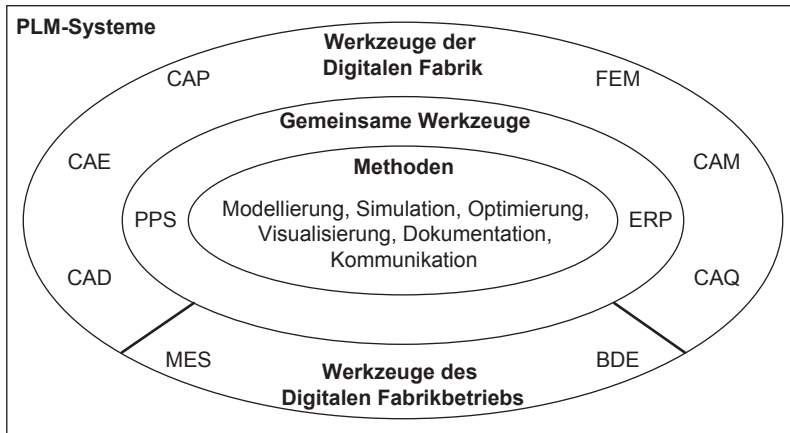


Abb. 13.4 Methoden und Werkzeuge der digitalen Fabrik. (entnommen und modifiziert nach: Landherr et al. 2013)

Produktions- und Fabrikstrukturen. Die Methoden der DF adressieren Aspekte der Modellierung, Simulation, Optimierung, Visualisierung, Dokumentation und Kommunikation, um organisatorische, technische sowie betriebswirtschaftliche Zielstellungen zu erreichen. In den Werkzeugen der DF finden sich zur Unterstützung dieser Methoden die bis hierher dargestellten rechnerbasierten Ansätze aus dem Produktionsumfeld gemeinsam wieder (Abb. 13.4) (Landherr et al. 2013).

13.2.5 Status Quo der Operational Business Intelligence

Business Intelligence (BI) umfasst Anwendungen, Technologien und Prozesse zum Sammeln, Speichern und Analysieren von Daten, damit Fachanwender bessere Entscheidungen treffen können (Watson 2009). Operational BI (OpBI) ist eine begriffliche Erweiterung, die durch folgende Definition präzisiert wird:

Operational BI bezeichnet integrierte geschäftsprozessorientierte Systeme, die mithilfe klassischer Business-Intelligence-Methoden auf der Basis zeitnaher, prozessualer Ablaufdaten und in aller Regel auch historischer, harmonisierter Daten eine Realtime-(Near-Realtime-) Unterstützung für zeitkritische Entscheidungen während des Prozessablaufes bieten. (Gluchowski et al. 2009)

Produktionsumgebungen eröffnen ein breites Anwendungsfeld für OpBI. Die Zweckmäßigkeit einer derartigen Entscheidungsunterstützung wird durch vergleichbare Funktionen industriegetriebener Ansätze verdeutlicht. So unterstützen MES z. B. die Sammlung, Aufbereitung und Darstellung von Produktionsdaten für eine Koordination der Betriebsabläufe (Thiel et al. 2008). Die Integration von Prozessdaten ist gleichermaßen charakteristisch für Advanced Process Control (APC) im Bereich der Halbleiterindustrie.

APC-Lösungen umfassen Module zur Durchführung von Ablaufkontrollen, Fehlerklassifikationen und Effizienzberechnungen im Produktionsumfeld sowie statistische Methoden zur Analyse und Überwachung von technischen Prozessen (Yugma et al. 2014).

Eine Extraktion von Informationen aus produktionsspezifischen Daten zum Zweck der Entscheidungsunterstützung spiegelt sich auch in der Terminologie einer Manufacturing Intelligence wider. Im Mittelpunkt steht ein durchgängiger Gebrauch von Datenintegrationstechniken zur Analyse, Modellierung und Simulation, um alle im Unternehmen befindlichen Produktionsprozesse in Echtzeit zu organisieren. Die Anwendung von Manufacturing Intelligence und eine entsprechende Informationsversorgung von Entscheidungsträgern werden aktuell im Kontext eines Smart Manufacturing diskutiert. Diese Diskussion zeigt auf, dass die bestehenden IT-Lösungen in der Industrie nur vereinzelte Prozessverbesserungen und eine lückenhafte Entscheidungsunterstützung mit sich bringen (Davis et al. 2012).

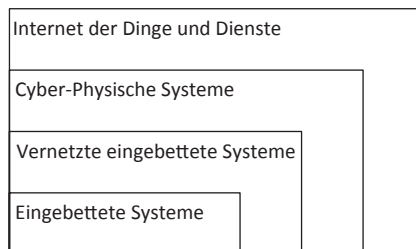
Die thematisierten Aspekte hinsichtlich einer durchgängigen Nutzung und Strukturierung von Produktionsdaten erlauben jedoch eine OpBI-getriebene Erweiterung der bestehenden Ansätze für eine produktionsspezifische Entscheidungsunterstützung. So können mit OpBI im Produktionsumfeld generierte Daten konsolidiert und visualisiert werden. Ein Anwendungsbeispiel ist die Analyse von Fertigungsaufträgen, um den Durchlauf von Produkten durch die Anlagen und Arbeitsplätze zu charakterisieren. Die Analyseergebnisse lassen sich als Kennzahlen in Produktionscockpits zusammenführen und veranschaulichen. Die im Netzwerk generierten Kennzahlen geben dann z. B. Auskunft über die Durchlaufzeiten von Fertigungsfolgen, Auslastungsgrade, Engpässe, Materialverbräuche oder die Gesamteffizienz (Laqua 2010).

OpBI ist in der Lage, die Analyse- und Berichtsfunktionen z. B. von MES zu verbessern und die erzeugten Erkenntnisse in eine unternehmensweite Entscheidungsarchitektur einfließen zu lassen (Hänel und Felden 2011; Koch et al. 2010). Damit ermöglicht OpBI eine Integration von analysierbaren Daten aus technischen und betriebswirtschaftlichen Zusammenhängen (Lasi 2012). Dies ist mit einer Verknüpfung von dispositiven und produktorientierten IT-Systemen in industriellen Unternehmen verbunden. OpBI unterstützt dadurch eine multidimensionale Sichtweise auf die Flexibilitätsanforderungen der Produktionsprozesse (Hänel und Felden 2013). Modularisierung und serviceorientierte Architekturen bilden mit Blick auf die zu erreichende Flexibilität den Ausgangspunkt für eine technische Umsetzung von OpBI (Hänel et al. 2012).

13.3 Industrie 4.0

Industrie 4.0 umfasst verschiedene Technologien und Konzepte. Das Thema wird als Zukunftsprojekt im deutschsprachigen Raum diskutiert. Der Kerngedanke ist eine Anwendung moderner IT-Techniken in der produzierenden Industrie, um eine intelligent vernetzte Produktion im Kontext einer Smart Factory zu erreichen (Kagermann et al. 2013).

Abb. 13.5 Vom eingebetteten System zum Internet der Dinge und Dienste. (entnommen und modifiziert Geisberger und Broy 2012)



13.3.1 Technologische Grundlagen

Die technologische Basis für Industrie 4.0 bilden Cyber-Physische Systeme (CPS) sowie das Internet der Dinge und Dienste. Diese Begriffe thematisieren die Vernetzung und Synchronisation von Informationen aus physischen und digitalen Umgebungen. Den Ausgangspunkt dieser Technologien bilden eingebettete Systeme, die informationsverarbeitende Einheiten mit physischen Prozessen oder Gegenständen verknüpfen. Die eingebetteten Systeme wurden in einer weiteren Entwicklungsstufe untereinander vernetzt. Der Erweiterungsaspekt von CPS ist die Vernetzung über das Internet. CPS bilden einen Teilbereich des Internets der Dinge und Dienste. In diesem Zusammenhang werden visionäre Konzepte wie z. B. Smart City oder Smart Factory diskutiert (Abb. 13.5) (Geisberger und Broy 2012).

Cyber-Physische Systeme CPS sind in physische Prozesse eingebettete Computer, die über ein Netzwerk miteinander kommunizieren (Lee 2008). Dies erlaubt eine regelkreisbasierte Überwachung und Steuerung von realen Prozessen mit digitalen Methoden. Eine acatech Studie aus dem Jahr 2012 präzisiert den Begriff wie folgt:

CPS umfassen eingebettete Systeme, Logistik-, Koordinations- und Managementprozesse sowie Internetdienste, die mittels Sensoren unmittelbar physikalische Daten erfassen und mittels Aktoren auf physikalische Vorgänge einwirken, mittels digitaler Netze untereinander verbunden sind, weltweit verfügbare Daten und Dienste nutzen und über multimodale Mensch-Maschine-Schnittstellen verfügen. (Geisberger und Broy 2012)

Eine beispielhafte CPS-Struktur ist in der nachfolgenden Abbildung dargestellt. Dabei werden die Kernkomponenten eines CPS deutlich. Dies ist zunächst die physikalische Umgebung, die unabhängig von Computern oder digitalen Netzwerken existiert. Über Schnittstellen wird eine Verbindung zur digitalen Welt hergestellt. Dort finden sich Module mit Rechneinheiten, Sensoren und Aktoren wieder, die über ein Netzwerk miteinander verbunden sind (Abb. 13.6) (Lee und Seshia 2011).

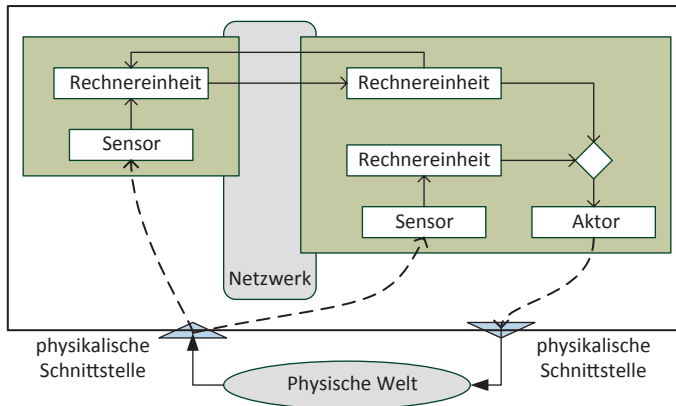


Abb. 13.6 Beispielhafte CPS-Struktur. (entnommen und modifiziert nach: Lee und Seshia 2011)

Im Zusammenhang mit CPS wird von offenen soziotechnischen Systemen gesprochen, die sich durch folgende Eigenschaften kennzeichnen (Geisberger und Broy 2012):

- **Verschmelzung von physikalischer und virtueller Welt:** Die Zusammenführung von physikalischen und digitalen Umgebungen erfordert die Kombination von kontinuierlichen Regelungs- und Steuerungssystemen mit diskreten Informationssystemen. Die mithilfe von Sensoren oder Sensorverbünden erfassten Umgebungsdaten können so parallel verarbeitet und interpretiert sowie anschließend zur Steuerung der physischen Prozesse in Echtzeit benutzt werden.
- **System of Systems mit dynamischen Systemgrenzen:** Dieses Merkmal spiegelt wider, dass CPS unterschiedliche Funktionen unterstützen sollen. Dazu entsteht situativ eine temporäre Kommunikation mit zusätzlichen Diensten und CPS-Komponenten.
- **Kontextadaptive und (teil-)autonom handelnde Systeme:** Die Fähigkeit zu kontextadaptiven Handeln erfordert eine umfängliche Kontexterfassung, um die gegenwärtige Anwendungssituation umfassend zu modellieren. Das System ist damit autonom in der Lage, sich an wechselnde oder unvorhergesehene Situationen oder Ereignisse anzupassen.
- **Kooperative Systeme mit verteilter und wechselnder Kontrolle:** CPS unterliegen oftmals keiner zentralen Kontrolle. Das Systemverhalten ergibt sich durch die Interaktion und Koordination von softwaregesteuerten Maschinen, Systemen, Diensten, Menschen und sozialen Gruppen. Die Kontrollmechanismen können in Mensch-Maschine-Interaktion, geteilte Kontrolle sowie Kontrolle zwischen verteilten Hardware- und Softwaresystemen eingeteilt werden.
- **Mensch-System-Kooperation:** CPS unterstützen die Interaktion zwischen Mensch und Maschine beispielsweise über multimodale Schnittstellen und können sowohl die Wahrnehmung als auch die Handlungsfähigkeit von einzelnen Personen und Gruppen vergrößern. Eine weitere Fähigkeit stellt die Analyse und Deutung menschlicher Gefühlszustände oder Verhaltensmuster dar. CPS besitzen eine Intelligenz um Entschei-

dungen oder Handlungen aus der Interaktion von Systemen und Menschen abzuleiten. Dies schließt eine Lernfähigkeit ein.

Internet der Dinge und Dienste Das Schlagwort Internet der Dinge (Internet of Things) thematisiert seit 1999 die Vernetzung von Gegenständen und Prozessen, die eigenständig interagieren und sich selbst organisieren (Brand et al. 2009). Die Erweiterung der Diskussion hinsichtlich Dienste (Services) hat das gemeinsame Begriffsumfeld Internet der Dinge und Dienste geprägt (Sendler 2013). Dies kennzeichnet im Zusammenhang mit Industrie 4.0 die internetbasierte Vernetzung von CPS mit anderen Computersystemen beliebiger Art. Die Begriffe Internet der Dinge und Internet der Dienste haben jedoch eigenständige Erläuterungen und Basistechnologien, die Gegenstand der nachfolgenden Betrachtung sind.

Das *Internet der Dinge* bindet eindeutig identifizierbare physikalische Objekte (Dinge) in das Internet oder vergleichbare Netzinfrastrukturen ein (Geisberger und Broy 2012). Die ist eng verknüpft mit den Begriffen des *Ubiquitous Computing* (UC) und des *Pervasive Computing* (PC). UC thematisiert, dass sich IT und Computersysteme in sämtlichen Bereichen des Alltags wiederfinden, um dort Hintergrundfunktionen auszuüben. PC hat eine vergleichbare Intention wie UC, richtet sich jedoch verstärkt auf die Vernetzung von mobilen oder eingebetteten Computersystemen. Physikalische Objekte werden so in die Lage versetzt, ihre Umgebung wahrzunehmen und miteinander über das Internet zu kommunizieren. Den Übergang zum *Internet der Dinge* kennzeichnet ein selbständiger Umgang der Objekte mit den Umgebungsdaten, um eigenständig Aktionen abzuleiten. Diese Fähigkeit wird auch als *Ambient Intelligence* (Umgebungsintelligenz) bezeichnet (Brand et al. 2009).

Die Verwirklichung des Internets der Dinge setzt Technologien (Atzori et al. 2010) voraus, die eine Identifikation, Erfassung und Kommunikation von physikalischen Gegenständen unterstützen. RFID (*Radio Frequency Identification*) ermöglicht in diesem Zusammenhang eine elektronische Objektidentifizierung. Externe Lesegeräte initiieren eine kontaktlose Übertragung von Daten, die auf einem Transponder hinterlegt sind. Derartige Transponder sind direkt an den Gegenständen angebracht und bestehen aus einem Chip sowie einer Antenne, um mit den externen Lesegeräten kommunizieren zu können. Die Chips enthalten einen elektronischen Produktcode (EPC) zum Zweck der eindeutigen Identifikation und zusätzliche Produktinformationen.

Sensornetzwerke unterstützen die Erfassung von Umgebungsdaten im Kontext des Internets der Dinge. Drahtlose Multihop-Netzwerke verknüpfen eine Vielzahl von Sensorknotenpunkten, deren Messwerte direkt oder über mehrere Knotenpunkte hinweg zu einer oder auch mehreren Basisstationen weitergegeben werden. Derartige Mikrosysteme können über Logik- und Speicherkomponenten verfügen, um Aktoren anzusteuern. Eine Besonderheit stellt die Anwendung von RFID in sensorischen Mikrosystemen dar. Die RFID-Transponder werden mit einer programmierbaren Steuereinheit und einem Sensor ausgestattet. Das Lesegerät fungiert als Basisstation, die sowohl Sensor- als auch Konfigurationsdaten empfangen kann. Der Transponder arbeitet als passiver Messpunkt, so

lange er nur über das Lesegerät mit Energie versorgt wird. Im Falle einer eigenen Energieversorgung, beispielsweise einer Batterie, können die Sensordaten auch im Transponder erfasst werden (Brand et al. 2009).

Die Mikrosysteme stellen mit ihren Kommunikationsfähigkeiten verteilte Anwendungssysteme dar, die über eine Middleware vereinheitlicht werden können (Lucke 2013). Eine Middleware stellt Dienste zur Kommunikation und Integration der Anwendungen bereit, die mithilfe der Infrastruktur des Internets der Dinge entwickelt wurden (Atzori et al. 2010). Die Eigenschaften und Funktionen der Mikrosysteme werden über Adapter gekapselt, sodass ein zentrales IT-System über bestimmte Schnittstellen auf die Daten der verteilten Anwendungssysteme zugreift, diese zwischen den Systemen austauscht und die Struktur der Daten vereinheitlicht. Middleware-Architekturen folgen im Zusammenhang mit dem Internet der Dinge dem Prinzip der Serviceorientierten Architektur (SOA). Diese wird als Basiskonzept des Internets der Dienste betrachtet.

Unter dem Internet der Dienste wird ein Teilbereich des Internets verstanden, in dem Dienstleistungen und Funktionalitäten in Form eigenständiger Softwarekomponenten mit webbasierten Technologien abgebildet werden (Geisberger und Broy 2012). In diesem Zusammenhang wird SOA als zu Grunde liegendes Architekturkonzept herangezogen. Es ist ein Ansatz zur Gestaltung und Umsetzung von Unternehmensanwendungen, der eine Vernetzung von lose gekoppelten IT-Services zur Unterstützung von Geschäftsprozessen adressiert. Ein IT-Service unterstützt eine abgrenzbare Funktion oder führt diese aus. Als Konsequenz der losen Kopplung ist es möglich, funktionale Komponenten je nach fachlichem Bedarf unterschiedlich zu kombinieren. Dies wird als Orchestrierung bezeichnet. Die Services sollen in diesem Zusammenhang wiederverwendbar sein und über eine plattformunabhängige Schnittstelle verfügen (Erl 2008).

Das Konzept einer SOA ist mithilfe von Web Services umsetzbar. Ein Web Service ist ein Softwaresystem, das eine vollständig kompatible Maschine-zu-Maschine-Interaktion über ein Netzwerk unterstützt. Dieser verfügt über eine in maschinenlesbarer Form beschriebene Schnittstelle. Dafür hat sich die Web Service Description Language (WSDL) etabliert. Die WSDL beschreibt die Schnittstelle eines Web Services und definiert seine Funktionalitäten sowie technische Details. Derartige Beschreibungen werden in einem Service Repository hinterlegt. UDDI (Universal Description, Discovery and Integration) stellt in diesem Zusammenhang einen Verzeichnisdienst dar, der die Veröffentlichung von Web Services ermöglicht. Andere Systeme können mit Web Services interagieren. Für die Kommunikation und Interaktion von Web Services ist ein Nachrichtenaustauschformat, das eine Suche, Erkennung und Nutzung unterstützt. Dafür kommt mit SOAP (Simple Open Network Protocol) ein Netzwerkprotokoll zum Einsatz, das sich auf XML (Extensible Markup Language) und weitere webbezogene Standards stützt. Der Transport erfolgt durch zu Grunde liegende Protokolle wie dem Hypertext Transfer Protocol (HTTP) oder dem File Transfer Protocol (FTP) (Abb. 13.7) (Melzer 2010).

Ein Web Service eines Anbieters wird durch ein WSDL-Dokument beschrieben und in einem Service Repository (UDDI) veröffentlicht. Ein potentieller Dienstanutzer durchsucht das Service Repository mithilfe von SOAP und findet durch das passende WSDL-Doku-

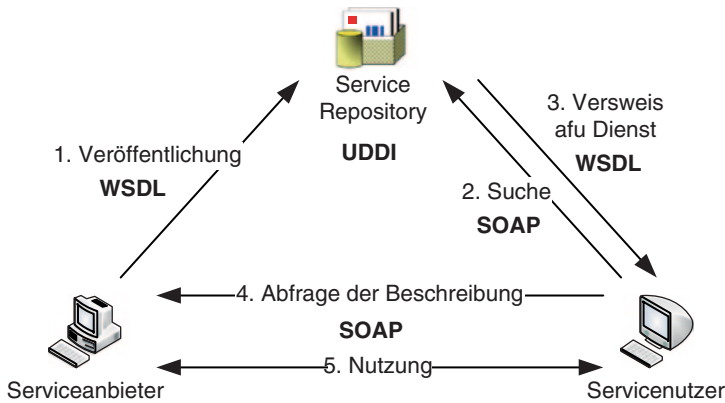


Abb. 13.7 Funktionsweise einer SOA. (Entnommen und modifiziert nach: Melzer 2010)

ment einen Verweis auf den gewünschten Dienst. Damit wird auf Basis der SOAP-Nachricht eine Kommunikation zwischen Dienstanbieter und Dienstnutzer hergestellt (Melzer 2010).

Cloud Computing ist ein weiterer Ansatz, der zur Umsetzung des Internets der Dienste diskutiert wird (Dufft 2010). IT-Infrastrukturen, Anwendungen oder Plattformen verteilen sich dabei auf verschiedene Systeme, die der jeweilige Anwender nicht lokal betreibt. Die Nutzung erfolgt über Dienste auf die über das öffentliche Internet (Public Cloud) oder über spezifische Intranets (Private Cloud) zugegriffen wird. Mithilfe von virtuellen Ressourcen sind die Dienste dynamisch skalierbar und abhängig vom Bedarf flexibel einsetzbar. Aus Anwendersicht bleibt der technische Unterbau verborgen, sodass der Zugriff über Schnittstellen oder standardisierte Programme erfolgen muss. Im Zusammenhang des Cloud Computing lassen sich drei generische Ebenen unterscheiden (Baun et al. 2011):

- Infrastructure as a Service (IaaS) umfasst bestimmte Dienste, um Kapazitäten für Server, Speicher und Netzwerke sowie Sicherheitsaspekte abzudecken.
- Platform as a Service (PaaS) adressiert die Bereitstellung von Entwicklungs- und Laufzeitumgebungen von webbasierten Anwendungen.
- Software as a Service (SaaS) bezeichnet webbasierte Anwendungen, die als Dienstleistung bezogen werden.

13.3.2 Das Zukunftsszenario „Smart Factory“

Das Ziel von Industrie 4.0 ist die Smart Factory. Die Systeme der Smart Factory sind vertikal mit den betriebswirtschaftlichen Prozessen integriert und verbinden horizontal alle Bereiche der Wertschöpfung in den Unternehmen und entlang der Wertschöpfungskette. Einzelne Unternehmen oder Unternehmensverbünde nutzen IT zur Entwicklung von Produkten, für die Gestaltung von Produktionssystemen, für Logistik- und Produktionsaufgaben sowie zur Interaktion mit ihren Kunden. Die Smart Factory zielt auf die

flexible Bearbeitung von individuellen Kundenwünschen, die Beherrschung komplexer Produktionsaufgaben, die Minimierung der Störanfälligkeit und die Effizienzsteigerung in der Produktion. Den Schwerpunkt bildet dabei die Kommunikation von Menschen, Maschinen und Ressourcen in einem internetbasierten Netzwerk. Intelligente Produkte (Smart Products) stellen Informationen über ihren Kontext (Identifikation, Ort, Zustand, Historie, alternative Wege zum Zielzustand) selbst zur Verfügung. Die Ressourcen der Smart Factory umfassen Maschinen, Roboter, Förderanlagen, Lagersysteme und Betriebsmittel einschließlich der angegliederten Planungs- und Steuerungssysteme. In der Smart Factory bilden diese Einheiten ein komplexes miteinander vernetztes System. Eine jeweils inhärente Intelligenz ermöglicht den Ressourcen, sich selbst zu steuern und zu konfigurieren sowie Informationen zu erfassen und diese über räumliche Entfernungen zu kommunizieren (Kagermann et al. 2013).

Technologisch basiert die Netzinfrastruktur der Smart Factory auf dem Internet der Dinge und Dienste. Die Intelligenz der Produktionsressourcen ist eine Folge des Einsatzes von CPS in Produktionssystemen. Aus der Anwendung von CPS in der produzierenden Industrie resultiert der Begriff Cyber-Physisches Produktionssystem (CPPS). Dieser Term beschreibt flexible, adaptive, selbstorganisierende und selbstkonfigurierende Produktionsanlagen. Daten, Dienste und Funktionen werden zu Gunsten einer effizienten Entwicklung und Produktion an einer geeigneten Stelle in der Smart Factory vorgehalten, abgerufen und ausgeführt. Die damit verbundene Vernetzung sowie die ubiquitäre Verfügbarkeit von Daten, Diensten und Funktionen bringen Veränderungen für die Produktionsstrukturen mit sich. Die bestehenden funktionalen Hierarchieebenen in der Produktion brechen auf und werden durch vernetzte, dezentrale Systeme ersetzt (Schlick et al. 2013).

Neben den technischen Veränderungen erfolgt im Rahmen der Industrie 4.0 eine intensive Auseinandersetzung mit der Rolle des Menschen in der Smart Factory im Hinblick auf die Steuerung, Regelung und Gestaltung der Produktionsressourcen. Der Mensch berücksichtigt in der Rolle des Entscheiders kontextbezogene Zielvorgaben zur Kontrolle der Produktionsabläufe und für die Qualitätssicherung (Abb. 13.8) (Kagermann et al. 2013).

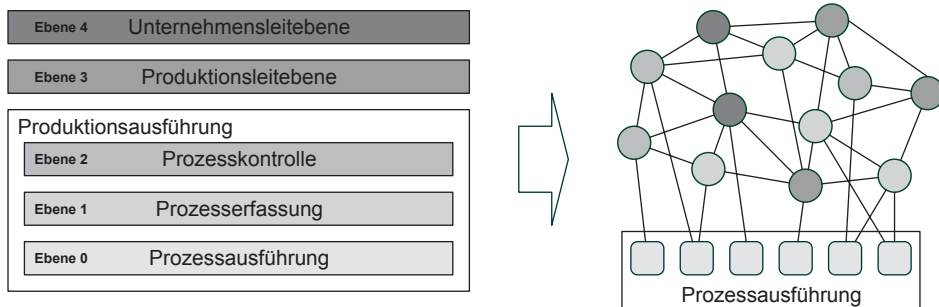


Abb. 13.8 Auflösung der hierarchischen Struktur in der Produktion (entnommen und modifiziert nach VDI 2013)

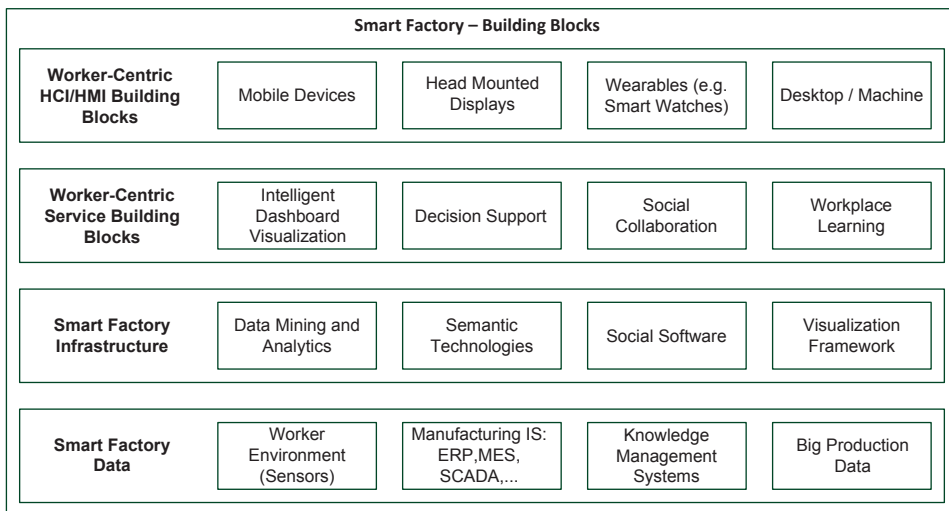


Abb. 13.9 Mensch-zentrierte IT-Architekturbausteine einer Smart Factory. (entnommen und modifiziert nach Stocker et al. 2014)

In diesem Zusammenhang wird die Relevanz von Wissensmanagement und Wissenstransfer in den Vordergrund gestellt. Die Architektur einer Smart Factory ist modular aufgebaut, um das notwendige Wissen mithilfe von Mensch-zentrierten IT-Lösungen anwendungsfallspezifisch bereitzustellen (Abb. 13.9) (Stocker et al. 2014).

Die schematisch dargestellte Architektur umfasst vier Ebenen. Die unterste Ebene charakterisiert die Daten der Smart Factory, die mithilfe von Sensoren im Arbeitsumfeld gesammelt werden. Eine Verknüpfung mit weiteren strukturierten sowie unstrukturierten Daten ist über Informationssysteme der Fertigung und Wissensmanagementsysteme möglich. „Big Production Data“ thematisiert, dass durch die steigende Digitalisierung die Datenmenge im Produktionsumfeld wächst. Die Daten werden mit hoher Geschwindigkeit generiert und liegen in verschiedenen Formaten vor. Die Infrastruktur der Smart Factory adressiert die Mustererkennung und Analyse von Produktionsdaten. Semantischen Technologien unterstützen eine IT-basierte Interpretation von Bedeutung und Kontext dieser Daten zur Nutzbarmachung von digitalem Wissen. In der Smart Factory kommt zum Wissensaustausch zwischen den Beschäftigten Social Software zum Einsatz. Einen weiteren Aspekt der Infrastruktur bildet die Visualisierung der Produktionsdaten. Die Dienste der Smart Factory sind in die Bereiche Visualisierungsdashboards, Entscheidungsunterstützung, Zusammenarbeit von Beschäftigten sowie das Lernen am Arbeitsplatz eingeteilt. Die oberste Ebene kennzeichnet Schnittstellen und Interaktionsmittel zwischen Mensch und Maschine, z. B. Anzeigen an Geräten, mobile Geräte oder intelligente Endgeräte (Stocker et al. 2014).

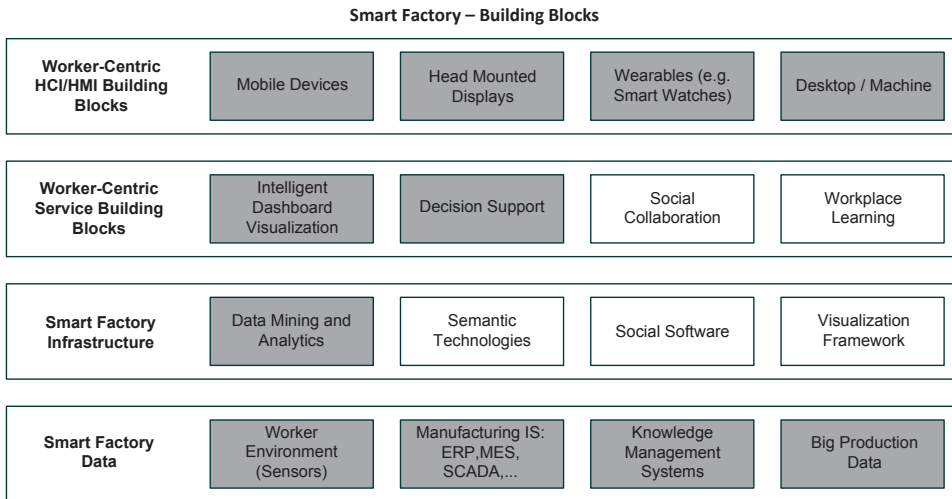


Abb. 13.10 OpBI in der IT-Architektur einer Smart Factory

13.3.3 Operational Business Intelligence in der Industrie 4.0

Unter dem Gesichtspunkt der Datenintegration lässt sich OpBI in das Zukunftsbild der Industrie 4.0 einordnen. Im Zusammenhang mit CPS und dem Internet der Dinge richtet sich das Potenzial von OpBI auf die Analyse und Weiterverarbeitung von Sensordaten. Die erzeugten Ergebnisse können dann von Entscheidungsträgern und technischen Akteuren gleichermaßen genutzt werden. Die bereits existierende Diskussion von OpBI im Rahmen von SOA stellt einen Zusammenhang zum Internet der Dienste her. Die Funktionen von OpBI können demzufolge als Dienste im Kontext einer Datensammlung, Datenaufbereitung, Leistungsanalyse und Informationsdarstellung umgesetzt werden. Damit fungiert OpBI als Bestandteil der Kommunikationsinfrastruktur von CPPS in der zukünftigen Smart Factory. Die nachfolgende Abbildung stellt die von OpBI unterstützten Architekturbausteine diesbezüglich dar (Abb. 13.10).

Die Architekturbausteine, die sich mit OpBI unterstützen lassen, sind in der obigen Abbildung durch eine graue Markierung hervorgehoben. Die Datengrundlage bilden Sensordaten aus dem Arbeitsumfeld der Mitarbeiter, aufgabenspezifische IT- und Wissensmanagementsysteme sowie eine Menge an Produktionsdaten. OpBI unterstützt mit Funktionen hinsichtlich Data Mining, Datenanalyse, Visualisierung und Entscheidungsunterstützung eine Mensch-Maschine-Interaktion in der Smart Factory. Die nachfolgende Darstellung ordnet diese Bausteine in eine serviceorientierte Plattform im Hinblick auf das Internet der Dienste ein (Abb. 13.11).

Die Funktionalitäten von OpBI werden über ein vielschichtiges Netzwerk angeboten, das eine Menge von heterogenen Daten verarbeitet. Die Analyseergebnisse können über Mensch-Maschine-Schnittstellen Entscheidungsträgern präsentiert werden oder direkt in

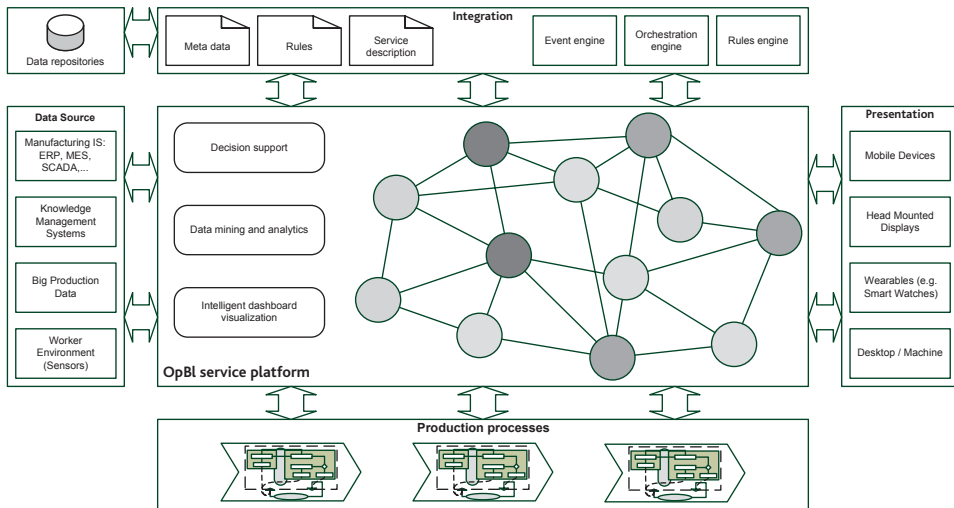


Abb. 13.11 OpBI-Serviceplattform in einer Smart Factory

die CPPS der Produktionsprozesse einfließen. Die Datenintegration und -kommunikation erfolgt internetbasiert über webserviceorientierte Strukturen.

13.4 Fazit

Der Beitrag nimmt eine Einordnung und Positionsbestimmung von OpBI im Rahmen des Zukunftsprojekts Industrie 4.0 vor. Diese Diskussion basiert auf der Darstellung des Entwicklungsstandes von IT-Systemen in Produktionsumgebungen und der Beschreibung des Status Quo von OpBI für eine produktionsspezifische Entscheidungsunterstützung. Der Beitrag zeigt daraufhin das Zukunftsbild der Industrie 4.0 als komplexes Zusammenspiel von CPPS über das Internet der Dinge und Dienste auf. In einem derartigen Netzwerk aus intelligenten Ressourcen, Produktionsmaschinen und Entscheidungsträgern besteht die Notwendigkeit für eine Datenverarbeitung und Informationsstrukturierung. OpBI kann vor diesem Hintergrund dienstbasiert eine Mensch-Maschine-Interaktion von Entscheidungsträgern unterstützen. Dabei wird eine Infrastruktur der Smart Factory zu Grunde gelegt, die aus Mensch-zentrierten IT-Bausteinen besteht.

Industrie 4.0 und die damit verbundene CPPS-Strategie adressiert eine horizontale und vertikale Integration von produktionsbezogenen IT-Systemen sowie ein durchgängiges Datenmanagement über die Produktlebenszyklen. Im Verlauf des Beitrages hat sich gezeigt, dass diese Aspekte sich ebenso im OpBI-Konzept widerspiegeln. Die aktuell umfassende Diskussion zur Industrie 4.0 mindert jedoch nicht die Relevanz einer OpBI-getriebenen Entscheidungsunterstützung im Produktionsumfeld. Diese Erkenntnis stützt sich auf die durch den Beitrag initiierte Gegenüberstellung der Technologien einer Industrie 4.0 mit OpBI. Die Betrachtung aus Perspektive der operativen Entscheidungsunterstützung

regt eine Neugestaltung von bestehenden Systementwürfen an. Die Berücksichtigung von serviceorientierten Architekturen ist dabei aus Sicht der IT schon länger Gegenstand der Diskussion. Diese wird aber von den sich in der Entwicklung befindlichen CPPS-Lösungen beeinflusst. Industrie 4.0 blickt in dieser Hinsicht ungefähr 10 bis 15 Jahre in die Zukunft (Bundesministerium für Bildung und Forschung 2014a). Der bis dahin erreichte Entwicklungsstand wird ein Gestaltungsfaktor für entscheidungsunterstützende Systementwürfe wie OpBI sein. Im weiteren Verlauf der Forschung sollten daher prototypische OpBI-Umsetzungen die Entwicklung der CPPS-Lösungen und damit die Diskussion um Industrie 4.0 begleiten.

Literatur

- Brand, L., Hülser, T., Grimm, V., Zweck, A.: Internet der Dinge – Perspektiven für die Logistik. VDI-Verlag, Düsseldorf (2009)
- Bracht, U., Geckler, D., Wenzel, S.: Digitale Fabrik – Methoden und Praxisbeispiele. Springer, Heidelberg (2011)
- Bullinger, H.-J.: Lean production. In: Mertens, P. (Hrsg.) Lexikon der Wirtschaftsinformatik, 4. Aufl. Springer, Heidelberg (2001)
- Gutenberg, E.: Grundlagen der Betriebswirtschaftslehre, Bd. 1, Die Produktion, 24. Aufl., Springer, Berlin, Heidelberg (1983)
- Hehenberger, P.: Computergestützte Fertigung. Springer, Heidelberg (2011)
- ISO 2806: Numerische Steuerung von Maschinen, Beuth, Berlin (1996)
- ISO 62264–3: Integration von Unternehmens-EDV und Leitsystemen, Teil 3: Aktivitätsmodelle für das operative Produktionsmanagement. Beuth, Berlin (2008)
- Jacobi, H.-F.: Computer Integrated Manufacturing (CIM), In: Westkämper, E., Spath, D., Constantinescu, C., Lentjes J. (Hrsg.) Digitale Produktion S. 51–92. Springer, Heidelberg (2013)
- Kagermann, H., Wahlster, W., Helbig, J.: Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. acatech, Frankfurt am Main (2013)
- Karaali, C.: Grundlagen der Steuerungstechnik, 2. Aufl., Springer, Wiesbaden (2013)
- Kief, H.B., Roschiwal, H.A.: CNC-Handbuch 2013/2014. Hanser, München (2013)
- Kurbel, K.: Produktionsplanung und -steuerung im Enterprise Resource Planning und Supply Chain Management, 6. Aufl., Oldenbourg, München (2005)
- Langmann, R.: Prozesslenkung – Grundlagen zur Automatisierung technischer Prozesse. Vieweg, Wiesbaden (1996)
- Landherr, M., Neumann, M., Volkmann, J., Constantinescu, C.: Digitale Fabrik. In: Westkämper, E., Spath, D., Constantinescu, C., Lentjes J. (Hrsg.) Digitale Produktion, S. 107–132. Springer, Heidelberg (2013)
- Lee, E.A.: Cyber physical systems: design challenges. 11th IEEE International Symposium on Object Oriented Real-Time Distributed Computing, Orlando (2008)
- Lee, E.A., Seshia, S.A.: Introduction to embedded systems – a cyber-physical systems approach. LeeSeshia.org (2011), Zugriff am 16.03.2015
- Lucke D.: Smart Factory. In: Westkämper, E., Spath, D., Constantinescu, C., Lentjes J. (Hrsg.) Digitale Produktion, S. 251–270. Springer, Heidelberg (2013)
- Melzer, I.: Service-orientierte Architekturen mit Web Services, 4. Aufl., Spektrum, Heidelberg (2010)
- Plümer, T.: Logistik und Produktion. Oldenbourg, München (2003)
- Schack, R.J.: Methodik zur bewertungsorientierten Skalierung der Digitalen Fabrik. Herbert Utz, München (2008)

- Scheer, A.-W.: Computer integrated manufacturing, 4. Aufl. Springer, Berlin (1990)
- Sendler, U.: Industrie 4.0 – Beherrschung der industriellen Komplexität mit SysLM (Systems Lifecycle Management). Springer, Heidelberg (2013)
- Soder, J.: Use Case Production: Von CIM über Lean Production zu Industrie 4.0. In: Bauernhansel, T., ten Hompel, M., Vogel-Heuser, B. (Hrsg.) Industrie 4.0 in Produktion, Automatisierung und Logistik. S. 85–102. Springer, Wiesbaden (2014)
- Stadtler, H.: Hierarchische Systeme der Produktionsplanung und -steuerung. In: Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Hrsg.) Handbuch Logistik, 3. Aufl. S. 194–214. Springer, Heidelberg (2008)
- Thiel, K., Meyer, H., Fuchs, F.: MES – Grundlage der Produktion von morgen. Oldenbourg, München (2008)
- Warnecke, H.-J.: Revolution der Unternehmenskultur – Das Fraktale Unternehmen, 2. Aufl. Springer, Heidelberg (1993)
- Werner, H.: Supply chain management – Grundlagen, Strategien, Instrumente und Controlling, 4. Aufl. Gabler, Wiesbaden (2010)
- Westkämper, E.: Struktureller Wandel durch Megatrends. In: Westkämper, E., Spath, D., Constantinescu, C., Lentjes J. (Hrsg.) Digitale Produktion, S. 7–10. Springer, Heidelberg (2013)
- Zäpfel, G.: Grundzüge des Produktions- und Logistik-Management, 2. Aufl. Oldenbourg, München (2001)

Henning Baars und Heiner Lasi

Inhaltsverzeichnis

14.1 Entscheidungsunterstützung in Logistik und Produktion – auf dem Weg zur Industrie 4.0 284

14.2 Relevante Entwicklungen in der Business Intelligence 285

14.3 Logistik und Produktion: Rahmenbedingungen, Herausforderungen und Entwicklungen 287

 14.3.1 Logistik 287

 14.3.2 Produktion 290

14.4 Innovative BI-Anwendungen in der Logistik 292

 14.4.1 Produktionslogistik und Ladungsträgermanagement 293

 14.4.2 Cloud- und RFID-basierte Retail Supply Chain 294

14.5 Innovative BI-Anwendungen in der Produktion 295

 14.5.1 Produktentwicklung und Produktdatenmanagement 296

 14.5.2 MES als Datenquelle für Produktions- und Qualitätsoptimierung 297

14.6 Diskussion und Ausblick 298

Literatur 299

Zusammenfassung

Mit der zunehmenden Digitalisierung von Logistik- und Produktionsprozessen und der steigenden Relevanz einer integrierten Entscheidungsunterstützung wächst auch

H. Baars (✉) · H. Lasi
Lehrstuhl für Allgemeine Betriebswirtschaftslehre und Wirtschaftsinformatik I,
Universität Stuttgart, Stuttgart, Deutschland
E-Mail: baars@wi.uni-stuttgart.de

H. Lasi
E-Mail: lasi@wi.uni-stuttgart.de

der Bedarf für leistungsfähige BI-Lösungen in diesen Domänen. Dieser Beitrag widmet sich innovativen Anwendungen in diesem Umfeld – vor dem Hintergrund der vorzufindenden Rahmenbedingungen, Herausforderungen und Entwicklungen. Ein besonderes Augenmerk wird dabei auf neue Technologien zur Datenerhebung und -aufbereitung gelegt: Es wird aufgezeigt, wie mit Radio Frequency ID und Manufacturing Execution Systems entscheidungsorientierte Daten gewonnen, wie diese mit Cloud- und Big-Data-Lösungen in die Entscheidungsunterstützung integriert werden können sowie welche Potentiale daraus resultieren.

14.1 Entscheidungsunterstützung in Logistik und Produktion – auf dem Weg zur Industrie 4.0

Eine effiziente Ausgestaltung der betrieblichen Funktionen Logistik und Produktion zählt seit jeher zu den Kernaufgaben der Betriebswirtschaft. Für die Planung und Steuerung dieser Bereiche wurde im Laufe der Zeit ein reichhaltiger Fundus an Methoden und Konzepten aufgebaut und in unterschiedlichen Typen von Standard-Anwendungssystemen umgesetzt. Das Spektrum reicht dabei von einfachen Dispositionssystemen bis hin zu komplexen lieferkettenumspannenden Supply-Chain-Management-Systemen (Musa et al. 2013; Arnold et al. 2008; Finkenzeller 2012; Gienke und Kämpf 2007).

Ein umfassendes Bild des aktuellen Produktions- und Liefergeschehens sowie das Verständnis und die Bewertung der dahinter stehenden Mechanismen und Zusammenhänge sind bis heute für viele Unternehmen erfolgsentscheidend. Das gilt in besonderem Maße für Industrieunternehmen und den Handel (Schuh und Gierth 2006; Wildemann 2005). Unter der Überschrift *Industrie 4.0* erfahren hier derzeit Konzepte zur horizontalen und vertikalen Integration sowie zur Steigerung der Digitalisierung und Vernetzung starke Beachtung. Beispiele hierfür sind der Einsatz Cyber-physischer Systeme im Kontext der Produktionssteuerung, die Durchdringung der physischen Umwelt mit Internettechnologien bis auf die Einzelobjekt- und Maschinensteuerungsebene (Internet der Dinge) oder die Erweiterung physischer Produkte um digitale Funktionalitäten (Embedded Systems) (Bauernhansl et al. 2014; Lasi et al. 2014).

Dass trotz des langjährigen und methodisch fundierten IT-Einsatzes immer noch Potenziale brachliegen und Industrie-4.0-Ansätze somit einen betriebswirtschaftlich relevanten Bedarf ansprechen, resultiert nicht zuletzt aus den Besonderheiten des Managements physischer bzw. hybrider Güter:

- Die diskutierten Aktivitäten befinden sich an der *Nahtstelle zwischen technischen und betriebswirtschaftlichen Kompetenzbereichen* und müssen heterogene und lange Jahre unabhängig voneinander entstandene Systeme und Konzepte zusammenführen (Mertens und Meier 2009).
- Für deren Unterstützung ist es erforderlich, eine *Abbildung der physischen Realität in der Sphäre der Informationsverarbeitung zu gewährleisten* (Bi et al. 2014; Fleisch und

Dierkes 2003). Obwohl bereits seit einiger Zeit Ansätze etwa zur Maschinendatenerfassung oder zur barcodegestützten Objektkennzeichnung etabliert sind, ist die Erfassung korrekter, zeitnaher und lückenloser Ist-Daten immer noch mehr Wunsch als Realität (Tellkamp 2006). Dies liegt auch daran, dass hierfür eine Kombination unterschiedlicher Technologien erforderlich ist (Bi et al. 2014; Musa et al. 2013).

- Sollen entsprechende Daten zur Managementunterstützung aufbereitet werden, kommt es schnell zu Engpässen aufgrund der anfallenden Datenvolumina, der Datenkomplexität und der Datenbereitstellungsgeschwindigkeit (Ma et al. 2013).
- Relevante Potentiale im Umgang mit physischen Gütern sind zudem mit *Anwendungen verknüpft, die Unternehmensgrenzen überschreiten* und die aufgrund volatiler globaler Rahmenbedingungen eine gleichsam integrierte und agile Informationsversorgung erfordern (DeGroote und Mar 2013; BVL 2013).

Hier setzen neue Konzepte der Datenerfassung und der Managementunterstützung an, die das Methodenspektrum der Industrie 4.0 komplettieren: Einerseits erlangen zunehmend neue Technologien zur automatisierten Erfassung von Maschinen- und Objektdaten Anwendungsreife (Günther et al. 2008). Andererseits wurde für die integrierte IT-gestützte Managementunterstützung (Business Intelligence, BI) eine erhebliche Integrationskompetenz aufgebaut. Von besonderer Relevanz für die Positionierung der BI in Logistik und Produktion sind dabei die immer stärker betriebene Verzahnung strategischer, taktischer und operativer BI-Systeme (Golfarelli et al. 2004; Gluchowski et al. 2009) sowie die Flankierung klassischer BI-Ansätze durch Cloud-Computing-Ansätze (Thomson und Van Der Walt 2010; Baars und Kemper 2010) und Technologien zur massiv-parallelen Verarbeitung von Daten (Big Data) (Chen et al. 2012; [BaFe14]). Die Kombination der neuen Datenerfassungsmöglichkeiten mit entsprechenden BI-Ansätzen eröffnet Potenziale, die weit über die der etablierten und oftmals isolierten Produktions- und Logistiksysteme hinausgehen.

Der folgende Beitrag widmet sich den entsprechenden Möglichkeiten. Hierfür werden zunächst kurz die relevanten Entwicklungen im BI-Bereich aufgezeigt, um dann Rahmenbedingungen, Herausforderungen und Entwicklungen in Logistik und Produktion vorzustellen. Im Folgenden wird die Diskussion anhand von vier exemplarischen Einsatzgebieten konkretisiert. Der abschließende Abschnitt diskutiert offene Fragen hinsichtlich des Zentralisierungsgrades und der Datenharmonisierung sowie Möglichkeiten für einen weitergehenden Ausbau des Integrationsgrads in der Managementunterstützung.

14.2 Relevante Entwicklungen in der Business Intelligence

Motive für eine Zusammenführung von BI mit Logistik und Produktion, insbes. bei Verfolgung eines Industrie-4.0-Ansatzes, resultieren aus dem zunehmend konsequenter verfolgten Anspruch von BI, einen möglichst *weitgehend integrierten Ansatz* zur Entscheidungsunterstützung aufzubauen sowie aus dem Trend zur *Verzahnung von BI mit operativen Prozessen* [BaFe14]. Zusätzlich unterstützen mehrere neuere Entwicklungen in der BI unmittelbar eine solche Kombination.

Mit dem Begriff *Business Intelligence* wird üblicherweise zunächst eine Reihe von Werkzeugen für die Bereitstellung, Aufbereitung und Analyse von Daten assoziiert – allen voran das Data Warehouse (DWH) als integriertes Datenhaltungssystem für die Managementunterstützung (Inmon 2005), OLAP für die Navigation in multidimensionalen Datenräumen (Codd et al. 1993), Data Mining für die Datenmustererkennung (Hand et al. 2001), Reporting-Tools zur Unterstützung des Berichtswesens sowie konzeptspezifische Anwendungen etwa zur Planung und zur Konsolidierung. Anwendungssysteme und -komponenten mit ähnlichen Aufgaben finden sich seit geraumer Zeit auch in Logistik und Produktion, etwa im Rahmen der Bedarfsprognose, der Produktionsplanung oder bei der Simulation von Lieferstrategien. Um den Innovationsgehalt möglicher neuer BI-Anwendungen ausloten zu können, ist es vor diesem Hintergrund notwendig, den im Business-Intelligence-Begriff enthaltenen *konzeptionellen, technischen und organisatorischen Integrationsaspekt* herauszustellen (Kemper et al. 2010).

Weitreichend integrierte Business-Intelligence-Infrastrukturen sind mittlerweile in vielen Unternehmen gelebte Praxis. Hiermit einher geht die Etablierung unternehmensweiter Governance- und Strategie-Konzepte für die BI sowie von *Business Intelligence Competence Centern* (BICC) als dedizierte organisatorische Einheiten (Miller et al. 2006; Gansor und Totok 2015). Gleichzeitig wird unter der Überschrift *Operational BI* der Einsatz von BI-Werkzeugen für die Unterstützung operativer und taktischer Aufgaben diskutiert – bei gleichzeitiger Integration der entsprechenden Systeme in existierende BI-Infrastrukturen (Gluchowski et al. 2009; Golfarelli et al. 2004). Hierbei wird gezielt auf die Stärken der BI-Werkzeuge abgehoben, Daten zu integrieren, anzureichern und gegebenenfalls auch zu historisieren. So wird es möglich, prozess- und systemübergreifende Monitoring-, Steuerungs- und auch Analysewerkzeuge aufzubauen (Sun 2014).

Dies trifft auf zunehmende Anforderungen zur Prozess- und Systemintegration in Logistik und Produktion. Dabei entstehende, neue Herausforderungen im Umgang mit Daten und Infrastrukturen, fallen mit weiteren Trends der BI zusammen, v. a. einer zunehmenden Kompatibilität der BI mit unternehmensübergreifenden Ansätzen, einer verstärkten Einbindung großvolumiger und heterogen strukturierter Datenquellen auf Basis massiv-paralleler Infrastrukturen („Big Data“) bzw. hauptspeicherbasierter Datenhaltungssysteme („InMemory-BI“) sowie einer systematischen Berücksichtigung kurzfristiger Anforderungsänderungen („agile BI“) [BaFe14]. Relevante Potentiale zur Erhöhung der Agilität in der BI und zur Unterstützung unternehmensübergreifender BI-Lösungen bietet zudem der Einsatz von Cloud-Lösungen, bei denen mit Internettechnologien auf mehrmandantenfähige, skalierbare und virtualisierte Infrastrukturen, Plattformen oder Anwendungen zugegriffen wird (Thomson und Van Der Walt 2010; Baars und Kemper 2010).

Insgesamt lässt sich konstatieren, dass mit dem Wachstum des Hoheitsbereichs der BI sowie mit deren technologischen und konzeptionellen Weiterentwicklung zunehmend breitere Überlappungsflächen mit den Systemen in Logistik und Produktion entstehen. Die daraus resultierenden BI-Anwendungen berühren dabei die unmittelbare Unterstützung des Transports, der Handhabung und der Umformung physischer Güter. Der folgende Abschnitt detailliert das hierbei zu berücksichtigende Anwendungs- und Systemumfeld.

14.3 Logistik und Produktion: Rahmenbedingungen, Herausforderungen und Entwicklungen

Bei einer Diskussion von Einsatzmöglichkeiten von BI-Ansätzen in Logistik und Produktion ist zu berücksichtigen, dass hier bereits ausgereifte Systemlandschaften bestehen, die oftmals spezialisierte Management-Unterstützungsfunktionalität mitbringen.

Dabei sind die Fragen der Logistik und der Produktion aufgrund des gemeinsamen Prozesszusammenhangs stark interdependent. *Produktion* wird im Folgenden als betriebliche Leistungserstellung bei Sachgüterproduktion abgegrenzt (Gienke und Kämpf 2007). Die *Logistik* beschäftigt sich hingegen mit der Gestaltung von Güter- sowie den assoziierten Wert- und Informationsflüssen (Arnold et al. 2008; Wildemann 2005). Dies umfasst neben der Beschaffungs-, Distributions- und Entsorgungslogistik auch die Produktionslogistik, die naturgemäß eng mit der Produktion abzustimmen ist (Pawellek 2007; Wildemann 2005).

14.3.1 Logistik

Der Stellenwert der Logistik für Industrie und Handel spiegelt sich bereits an dem hohen Anteil der Logistik- an den Gesamtkosten wider, der im Schnitt bei 8,6% liegt (BVL 2013) liegt. Darüber hinaus wird auch der Outputseite – die *Logistikleistung* mit ihren Komponenten Lieferzeit, Lieferservice, Liefertreue und Lieferflexibilität – eine große Bedeutung zugesprochen (BVL 2013; Wildemann 2005; Arnold et al. 2008). Je nach Branche und Unternehmensausrichtung finden sich hier unterschiedliche Schwerpunktsetzungen. In den letzten Jahren haben insbes. Nachhaltigkeitsziele und Agilitätsanforderungen zunehmendes Gewicht erhalten (BVL 2013).

Da der Logistik per Definition ein Querschnittscharakter inhärent ist, verwundert es nicht, dass dort – anders als in der Produktion – häufig bereits eine grundsätzlich prozessorientierte Sicht vorgefunden werden kann. Dies zeigt sich daran, dass seit jeher neben der funktionsspezifischen Differenzierung in Beschaffungs-, Produktions- und Distributionslogistik ein besonderes Augenmerk auf die Optimierung unternehmensübergreifender Logistikketten und netzwerke im Rahmen des *Supply Chain Management* (SCM) gelegt wird (Chopra und Meindl 2012; Christopher 2010).

Die besondere Bedeutung eines prozessübergreifenden Informationsaustausches wurde dabei wiederholt und nachdrücklich belegt (Christopher 2010; BVL 2013). Hiermit verbunden sind Potenziale zur Vermeidung des Aufschaukelns von Nachfrageschwankungen („Bullwhip-Effekt“) (Lee et al. 1997), zur Reduzierung von Sicherheitsbeständen sowie zum Wechsel zu effektiveren Lieferstrategien.

Der Stellenwert der Informationstechnik für die Logistik ist somit evident. Im Laufe der Zeit ist eine Vielzahl spezifischer Einzellösungen entwickelt worden, etwa für die Lagerverwaltung (Warehouse Management), für die Kommissionierung, für das Flotten-, Fahrzeug- und Transportmanagement oder für die Warenwirtschaft. Flankiert werden

diese Systeme durch intelligente und (teil-)automatisierte Transport- und Lagersysteme (z. B. Sorter, Stetigförderer, Hochregallager) (Helo und Szekely 2005; Gleißner und Femerling 2008; Arnold et al. 2008). Aufgrund der stärkeren Prozessausrichtung wurden bereits frühzeitig Lösungen für die unternehmensübergreifende Koordination sowie für den Datenaustausch entwickelt – speziell bei Supply Chains mit einem dominierenden („fokalen“) Unternehmen. So ist der Einsatz von Barcodes mit entsprechenden internationalen Standards (EAN, UPC) seit langem etabliert. Bereits seit den 80er Jahren werden Electronic-Data-Interchange-(EDI)-Lösungen im Handel („SEDAS“ bzw. später „EAN-COM“) und in der Automobilindustrie („ODA/ ODETTE“) eingesetzt, die durch XML-basierte Austauschformate und Frameworks abgelöst werden (Gleißner und Femerling 2008; Arnold et al. 2008; Finkenzeller 2012).

Die deutlich umfassender ausgelegten SCM-Lösungen bringen üblicherweise ein Bündel unterschiedlicher Komponenten zusammen, etwa Transaktionssysteme für das operative SCM (d. h. für die Abwicklung von Einzelaktivitäten wie Bestellungen und Lieferungen), Monitoring-Lösungen für das Supply-Chain-Event-Management, Komponenten zum elektronischen Datenaustausch sowie Komponenten zur Planung und Simulation von Supply-Chains – sogenannte Advanced Planning Systems (APS) (Meyr et al. 2008; Betge 2006). APS sind managementunterstützende Systeme, die intensiv auf BI-Technologien aufsetzen, unmittelbar von Innovationen im BI-Bereich profitieren und idealerweise auch in die BI-Infrastruktur integriert sind. So ermöglichen APS-Lösungen auf Basis von In-Memory-BI-Komponenten häufigere, interaktivere und komplexere Planungsläufe (Schmalzried et al. 2013).

Trotz der beschriebenen Anstrengungen verbleibt weiter Optimierungspotenzial. Dieses resultiert zum einen aus der Komplexität der heutigen, oftmals globalen und volatilen Supply-Chain-Netzwerke. Zum anderen stoßen Konzepte zur Prozesssteuerung und -optimierung aufgrund unzureichender Datenqualität und -verfügbarkeit an Grenzen. Der hohe Aufwand einer manuellen oder halbmanuellen Datenerfassung führt zu Informationslücken, einer unzureichenden Zuordnung physischer Objekte zu übergreifenden Güterflüssen sowie insgesamt fehlerhaften Datenbeständen. Beispiele für Bereiche, in denen noch ein Verbesserungsbedarf gesehen wird, betreffen die Vermeidung von Lieferengpässen („Out-of-Stock-Situationen“), Schwund, Verderb und Beschädigung von Waren, die Genauigkeit und Korrektheit von Bestandsdaten, die Rückverfolgbarkeit von Waren sowie das Management von Transportmaterial (Brewer et al. 1999; Leclerc 2002; Tellkamp 2006).

Abhilfe versprechen neue Identifikationstechnologien, speziell die funkbasierte Übertragung von Objektdaten mittels *Radio Frequency Identification* (RFID) (Alexander et al. 2002; Kambil und Brooks 2002; Tellkamp 2006). Wesentliche Treiber des RFID-Einsatzes sind dabei die Standardisierungsbemühungen des EPC-Global-Konsortiums, das Standards für die Hardware (RFID-Tags und Reader), die Codierung, die Middleware sowie einen unternehmensübergreifenden Datenaustausch definiert hat (GS 1 2014).

Die mit einem RFID-Einsatz einhergehenden Möglichkeiten, Objekte unterhalb der Logistikeinheit verfolgen zu können, die Objektidentifikation vollautomatisiert und als

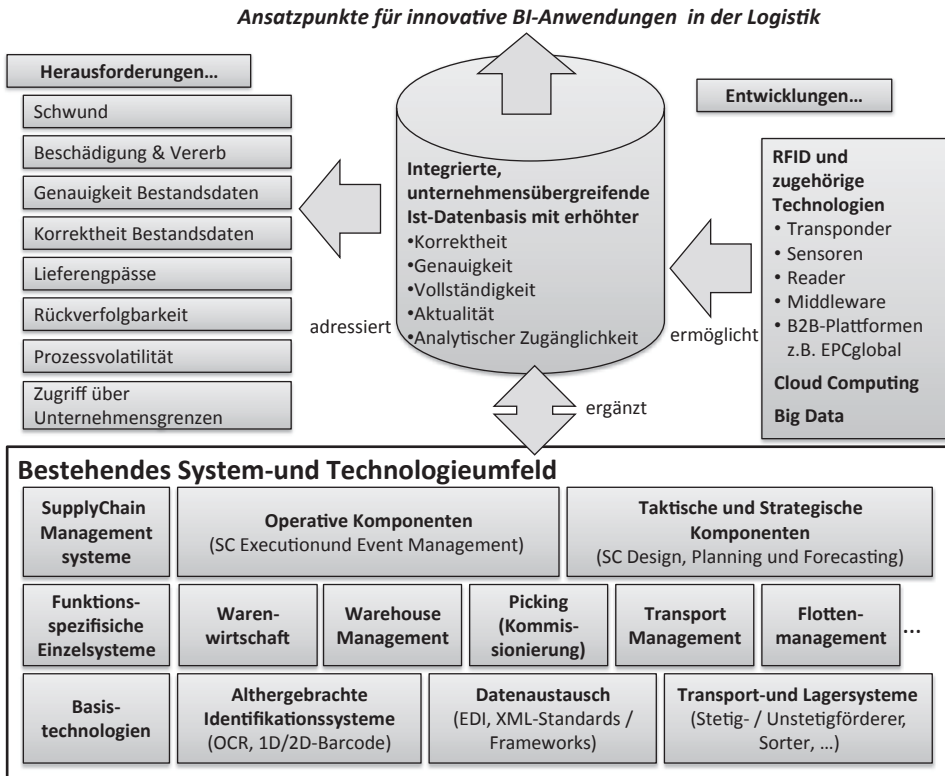


Abb. 14.1 Ansatzpunkte für innovative Systeme in der Logistik

Pulkerfassung durchzuführen, effizient die räumliche und zeitliche Auflösung der Daten zu verfeinern sowie deren Bereitstellungszeit in den Near-Time-Bereich zu drücken, versprechen einen wesentlichen Beitrag zum Aufbau einer gänzlich neuartigen Ist-Daten-Basis mit neuen Möglichkeiten einer prozessorientierten Ursachenanalyse [BaSu09] (Ilic et al. 2009; Sun 2014).

Sofern derartige analytische Lösungen realisiert werden müssen, stellen sich allerdings Fragen hinsichtlich des Umgangs mit den in großer Menge und in hoher Frequenz einströmenden Daten sowie mit der Verortung der erforderlichen Infrastrukturen. Den datenbezogenen Problemen kann mit Big-Data-Lösungen auf Basis massiv-paralleler Infrastrukturen begegnet werden, die auf die Nutzung relationaler Schema verzichten („NoSQL“) (Sandler 2014). Für einen schnellen Aufbau der entsprechenden Infrastrukturen zur Datenvorhaltung und -analyse können hingegen Cloud-Lösungen genutzt werden (Yan et al. 2014; Baars und Kemper 2011). Die kombinierte Nutzung von Big-Data- und Cloud-Lösungen stellt somit die analytische Zugänglichkeit der per RFID-Infrastrukturen erfassten Datenbestände sicher. Abbildung 14.1 liefert eine Übersicht über die aufgezeigten Zusammenhänge.

14.3.2 Produktion

Die industrielle Produktion, die oftmals als Rückgrat der deutschen Wirtschaft gesehen wird, muss sich in einem anspruchsvollen Spannungsfeld von Anforderungen im internationalen Wettbewerb behaupten. Neben dem permanenten Strukturwandel und der Notwendigkeit hoher Investitionen in Forschung- und Entwicklung (F&E) ergibt sich ein Grunddilemma des Produktionsmanagements aus den sich verkürzenden Produktlebenszyklen, einer zunehmend geforderten Kundenindividualität bei einem gleichzeitigen Zwang zur Standardisierung sowie kurzen Lieferzeiten und geringer Lagerhaltung (Friedli 2006).

Die Aufbauorganisation industrieller Betriebe (Sachgüterproduktion) folgt häufig funktionalen Kriterien. Klassische Organisationseinheiten eines Industriebetriebes sind daher die Verrichtungsbereiche Beschaffung, F&E, Produktion und Absatz (Bea und Göbel 2002). Dies schlägt sich auch in der etablierten Systemlandschaft vieler industrieller Unternehmen nieder:

- Es dominieren bereichsspezifische *Insellösungen*, beispielsweise unterschiedliche Qualitätsmanagementsysteme für Produktentwicklung und Erprobung, Fertigung und Qualitätskontrolle.
- Häufig herrscht eine *heterogene Landschaft* mit teilweise redundanten Anwendungssystemen in unterschiedlichen Bereichen vor (z. B. isolierter Einsatz unterschiedlicher Systeme für die Maschinendatenerfassung in unterschiedlichen Produktionsbereichen).
- Die historisch gewachsenen Systeme sind *unzureichend integriert* (z. B. Medienbrüche zwischen Konstruktions-, Planungs- und Steuerungstools).

Die veränderte Wettbewerbssituation und die Entwicklungen auf den Absatzmärkten, z. B. die Globalisierung, die zunehmende Geschwindigkeit der Migration von Produkten oder der rasante technologische Fortschritt, stellen verschärfte Anforderungen an die Optimierung von Produktionsprozessen. Dies führt zu einer Divergenz zwischen den etablierten Strukturen der ausführenden Verrichtungseinheiten und dem Ablauf der Leistungsprozesse. Um nachhaltige Effektivitäts- und Effizienzsteigerungen zu erreichen, bedarf es einer Optimierung des gesamten *Systems* der Produktion und nicht nur einzelner Maschinen, Anlagen oder Bereiche (Subsysteme) (Westkämper 2007).

Konzepte für ein entsprechend ganzheitliches Produktionsmanagement werden üblicherweise in eine operative, eine taktische und eine strategische Ebene unterteilt (Corsten 2007).

Das *operative Produktionsmanagement* beinhaltet schwerpunktmäßig die Feinplanung, Steuerung und Kontrolle einzelner Fertigungsschritte auf der Fertigungsebene (Shop Floor) (Corsten 2007). Ausführende Personen sind z. B. Schichtführer, Meister oder Gruppenleiter. Zielsetzung des operativen Produktionsmanagements ist die Optimierung der Ausbringungsmenge durch eine hohe Kapazitätsauslastung und eine Minimierung der Anlagenstillstände. Hierfür benötigte Informationen sind Maschinen-, Betriebs-, Prozess-, Produktions- und Personaldaten. Aufgrund der Zeitkritikalität werden diese Daten in der

Regel in Echtzeit und auf einer feingranularen Ebene benötigt (z. B.: Bei welcher Maschine tritt gerade eine Störung auf oder wie groß ist die Ausbringungsmenge einer gegebenen Maschine in dieser Minute).

Zur Unterstützung der Shop-Floor-Ebene kommen *Manufacturing Execution Systems* (MES) zum Einsatz. Diese verfügen in der Regel über eine Echtzeitanbindung von Produktionsmaschinen (Maschinen-/Betriebsdatenerfassung – MDE/BDE) (Günther et al. 2008; VDI Richtlinie 5600 2007). Die entsprechenden Systeme verfügen daher vielfach über umfangreiche Schnittstellenbatterien (für Maschinen, Zeiterfassung, Fördersysteme, RFID-Systeme etc.) und Analyse- und Reportingfunktionalität (Kletti 2006; IT&Production 2014).

Der operativen Ebene fällt im Kontext von *Industrie 4.0* eine Schlüsselrolle zu. Dies ist der Fall, da die Erhöhung der Flexibilisierung und Automatisierung vielfach über die Implementierung autonomer, sich selbst steuernder Produktionssysteme erfolgen soll. Wesentliche Ansätze beinhalten hierbei den Einsatz Cyber-physischer Systeme zur Kommunikation von Maschinen, Bauteilen und Steuerungssystemen. Die Steuerungssysteme können entweder als agentenbasierte Embedded Systems in Maschinensteuerungen oder als übergeordnete Steuerungsinstanzen realisiert werden. Auf Basis dieser Entwicklungen ist davon auszugehen, dass MES-Systeme sich weg von Monitoring- und entscheidungsunterstützenden Systemen in Richtung leistungsfähiger Steuerungssysteme entwickeln, die auf Basis von Planungskorridoren aus der taktischen und der strategischen Ebene Fertigungseinheiten in Echtzeit steuern. Damit gewinnen aus Sicht der Entscheidungsunterstützung die taktische und die strategische Ebene zunehmend an Bedeutung.

Die *taktische Ebene* des Produktionsmanagements umfasst die Lenkung des Wertschöpfungsprozesses auf der Planungsebene. Hierzu gehören z. B. die bereichsübergreifende Produktionsplanung und -steuerung (Gienke 2007). Die IT-basierte Unterstützung erfolgt dabei mittels Produktionsplanungs- und -steuerungssystemen (PPS), die in modernen Enterprise-Resource-Planning-(ERP)-Systemen häufig beinhaltet sind.

Im Kontext von *Industrie 4.0* fällt der taktischen Ebene zusätzlich die Gestaltung der Korridore zum Aufspannen von Lösungsräumen, in denen sich die autonomen Fertigungssysteme selbst steuern können, zu. Hierdurch entstehen neue Planungsaufgaben, die zur zielführenden Ausführung integrierte Informationen aus der taktischen sowie der automatisierten operativen Ebene benötigen. Daher entsteht ein Bedarf an integrierten Entscheidungsunterstützungssystemen auf der taktischen Ebene. Zu berücksichtigen ist hierbei, dass innerhalb der automatisierten operativen Ebene die erfasste Datenmenge aufgrund der weitreichenden Durchdringung der Fertigungseinheiten mit Sensorik rasant anwächst. Daher entsteht in der taktischen Ebene zunehmend der Bedarf an Big Data-Lösungen.

Das *strategische Produktionsmanagement* zielt auf die langfristige Ausrichtung der Produktion als komplexes System ab (Friedli 2006). Aufgabenträger hierbei ist das Top-Management. In der Literatur und Praxis finden sich viele innovative Konzepte wie z. B. die Digitale Fabrik oder die Fraktale Fabrik, die als Lösungsansätze für die Gestaltung der Fabrik als Gesamtsystem konzipiert sind.

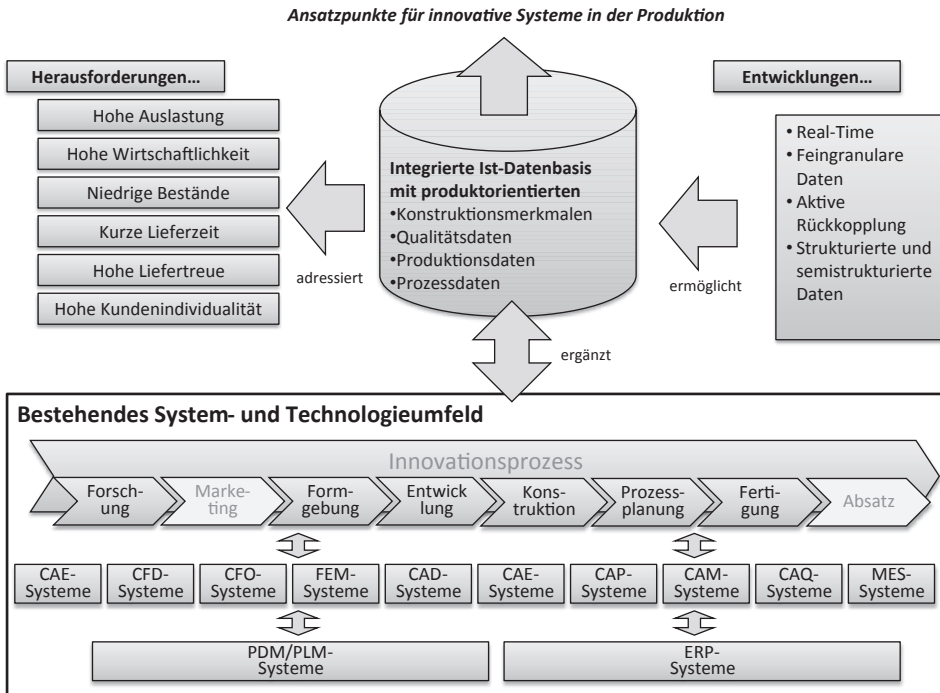


Abb. 14.2 Ansatzpunkte für innovative Systeme in der Produktion

Im Zuge der IT-basierten Entscheidungsunterstützung des Top-Managements werden bereits klassische BI-Systeme genutzt. Diese enthalten Daten aus transaktions- und wertorientierten Systemen wie z. B. ERP- oder PPS-Systemen. Die Extraktion, Transformation und Aufbereitung von produktionsorientierten Daten aus Shop-Floor-Systemen wird hierbei bislang wenig betrachtet. Damit fehlt dem Top-Management ein wesentliches Instrument, das in Fragen der strategischen Produktionsausrichtung eine ganzheitliche Entscheidungsunterstützung bietet. Abbildung 14.2 visualisiert dies vor den dargestellten Hintergründen.

14.4 Innovative BI-Anwendungen in der Logistik

Im Folgenden wird anhand zweier Anwendungsdomänen aufgezeigt, welche Potenziale in der Logistik durch eine Kopplung von BI-Infrastrukturen mit einer RFID-basierten Objektdatenerfassung erschlossen werden können. Hierbei spielen integrierte DWHs, die im Logistikkontext primär für Planungs- und Simulationsaufgaben eingesetzt werden, eine Schlüsselrolle. Die Ergebnisse basieren dabei auf einer Studienreihe, die von einer Prototypenentwicklung begleitet wurde.

Es ist zu betonen, dass die Bedeutung von DWHs in der Logistik generell mit der Reichweite und damit den Integrationsanforderungen der Lösungen zunimmt (vgl. z. B.

Behler et al. 2003). Die Potenziale zur Erhöhung von Datenqualität, -granularität und -aktualität in RFID-Lösungen sind als besondere Treiber zu sehen. Eine zielgerichtete Zusammenführung, Anreicherung und Historisierung roher Güterbewegungsereignisse mit Daten aus zahlreichen benachbarten Funktionsbereichen (etwa Produkt-, Produktions-, Kosten-, oder Kundendaten) eröffnen dabei neue Möglichkeiten zum Monitoring des Liefergeschehens sowie zu dessen umfangreicher Analyse (Niederman et al. 2007; Ilic et al. 2009; Baars und Sun 2009). Die entsprechenden Daten können u. a. mit OLAP-Systemen, Geovisualisierungs- oder speziellen Process-Mining-Werkzeugen aufbereitet und für Analysten und Entscheider bereitgestellt werden (Ilic et al. 2009; Sun 2014; Baars und Sun 2009). Darüber hinaus kann das DWH auch über eine automatisierte Mustererkennung Ereignisse in den operativen SCM-Systemen auslösen.

14.4.1 Produktionslogistik und Ladungsträgermanagement

Im Kontext der Produktionslogistik spielt das Management von Mehrwegladungsträgern (Paletten, Gitterboxen u. ä.) eine besondere Rolle. So wird im Transport von Bauteilen im Automobilbau oftmals eine erhebliche Zahl kostenintensiver (Spezial-)Ladungsträger eingesetzt, deren Nutzung bislang üblicherweise nur unzureichend überwacht, gesteuert und optimiert wird (Leclerc 2002; Strassner et al. 2005).

Dies überrascht nicht, da üblicherweise die Voraussetzungen für solche Managementaktivitäten fehlen. Hierfür können zwei Ursachen ausgemacht werden (Koch und Baars 2009):

1. Die Bestände werden oftmals nur manuell, in relativ großen Zeitintervallen (z. B. durch tägliche Zählung) und an einzelnen Punkten erfasst (etwa an einem zentralen „Pool-Platz“, an dem Ladungsträger gesammelt und geprüft werden).
2. Die für das Management des Ladungsträgerkreislaufs benötigten Daten werden – wenn überhaupt – verteilt in separaten Anwendungen vorgehalten. Beispielsweise werden Frachtbrief- und Lieferscheindaten üblicherweise im ERP-System abgelegt, im PPS- oder ERP-System geplante Produktionsmengen gespeichert und in operativen Insellösungen für das Ladungsträgermanagement Angaben zu Reservierungen von Ladungsträgern, zu den Bedarfen von Lieferanten oder zur Menge der Ladungsträger im Pool.

Die unzureichende Informationsversorgung hat mehrere Konsequenzen, etwa eine Prozessintransparenz, die einen hohen Abstimmungsaufwand nach sich zieht, erhöhte Kosten aufgrund einer getrennten Verwaltung von Bestands- und Bewegungsdaten, Engpässe in der Ladungsträgerbereitstellung, unnötig hohe Sicherheitsbestände, lange Umlaufzeiten sowie eine schlechte Auslastung (Hofmann und Bachmann 2006).

Der hier vorgestellte Lösungsansatz basiert auf einer Ausstattung von Ladungsträgern mit RFID-Chips in Kombination mit einer DWH-basierten Sammlung, Anreicherung und Historisierung der Daten (Koch und Baars 2009). Es ist zu betonen, dass der Einsatz von

RFID-Technologie für das Ladungsträgermanagement derzeit von vielen Unternehmen pilotiert oder gar operativ betrieben wird – allerdings üblicherweise dabei die integrative Wirkung eines DWH noch nicht berücksichtigt wird.

Auf der Grundlage einer Zusammenführung der beschriebenen Daten in einem DWH können eine Reihe analytischer Anwendungen unterstützt werden. Hier seien beispielhaft drei genannt (Koch und Baars 2009):

1. Ein weitreichendes Problem stellt der Schwund von Ladungsträgern im Nutzungskreislauf dar. Üblicherweise wird ein Ladungsträgerverlust erst bei Sortiervorgängen im Pool-Platz festgestellt. Mit einer feinmaschigen Erfassung sowie einer DWH-basierten historisierten Vorhaltung der Bestandsmengen werden Analysen zu Fehlmengen im Zeitverlauf möglich. Dies unterstützt nicht nur eine effektive Beschaffung neuer Ladungsträger, sondern hilft auch bei der Verortung der Ursachen von Schwund.
2. Bedarf für analytische Systeme besteht auch hinsichtlich der Überwachung der Performance in der Ladungsträgerversorgung, etwa zum Erkennen eines ungewollten Bestandsaufbaus. Relevant ist zudem das Erkennen einer Unterversorgung mit Ladungsträgern trotz eigentlich ausreichender Bestände, da dies auf Probleme in der Auftragsabwicklung hindeuten kann.
3. Die Disponenten können bei der Frachtenplanung unterstützt werden, indem ihnen Daten zur zeitlichen Entwicklung des disponierbaren Bestands je Ladungsträgertyp in Relation zu den zugehörigen Bedarfen bereitgestellt werden – als real erhobene Ist-Daten. Damit kann Engpässen entgegengewirkt werden, die aus einer ungenauen, prognosebasierten Disposition resultieren. Basierend auf historischen Daten wird es darüber hinaus möglich, Tendenzen in der Bedarfsentwicklung über den Jahresverlauf zu erkennen.

14.4.2 Cloud- und RFID-basierte Retail Supply Chain

Aufgrund entsprechender Initiativen großer Handelshäuser wurden die Potenziale eines Einsatzes von RFID in der Supply Chain des Einzelhandels (*Retail Supply Chain*) bereits ausführlich untersucht (vgl. z. B. Morán et al. 2004; METROGroup 2004; Tellkamp 2006).

Grundsätzlich wird der Technologie in dieser Domäne weitreichendes Transformationspotenzial zugesprochen (METROGroup 2004). Dieses resultiert vor allem aus der zeitnahen Erfassung und der über EPC standardisierten Weitergabe von Objektdaten über die Supply Chain.

Eine entsprechende Lieferkette muss dabei mehrere unterschiedliche Beteiligte zusammenführen, zu denen insbesondere die Verkaufsfilialen, Transportdienstleister, Warenverteilzentren sowie Hersteller zählen – die jeweils heterogene Systeme mitbringen (Baars et al. 2008). Beispiele für Datenquellen, die zu integrieren sind, umfassen Herstelldaten (z. B. Produktionscharge), Produktdaten (z. B. Haltbarkeitsdaten), Kapazitätsdaten (z. B. Regalkapazitäten in Warenverteilzentrum und in den Filialen) sowie Tracking- und Tracing-Daten. Eine DWH-basierte Zusammenführung von Daten dieser Beteiligten wird

dabei nicht zuletzt durch die standardisierte Kennzeichnung von Objekten auf Einzelartikelebene ermöglicht. Der EPC-Code ist dabei ein natürlicher Kristallisationskern für eine Datenintegration.

Hier drei Beispiele für Anwendungen, die auf eine solche RFID-Cloud-Lösung aufsetzen können (Baars et al. 2008):

1. Die Identifikation von Mängeln an Gütern während des Transports (beispielsweise auf Basis von Temperatur-, Feuchtigkeits- oder Drucksensoren) sowie darauf aufsetzend die Lokalisierung von Mangelursachen und die Analyse möglicher Regelmäßigkeiten beim Auftreten der Mängel im Zeitverlauf.
2. Eine effizientere Versorgung der Filialen, etwa durch einen vom Warenverteilzentrum koordinierten Bestandsausgleich zwischen den Filialen unter Einbezug verfügbarer Regalflächenkapazitäten. Eine im DWH hergestellte Übersicht über Ist-Mengen im Zeitverlauf erlaubt es, die Beschaffungspolitik fortlaufend zu optimieren.
3. Ein selektiver Rückruf von Produkten – basierend auf der betroffenen Chargennummer. Dieser Anwendung wird insgesamt großes Potenzial zugesprochen, sie setzt aber eine umfangreiche Datenintegration voraus.

Wie bereits im Ladungsträgerfall wird in den Beispielen ein operatives Monitoring mit taktisch-strategischen Analyseaufgaben verbunden. Auch ist es sinnvoll, die Systeme im Sinne eines Operational BI an transaktionsorientierte SCM-Komponenten anzuschließen, etwa für automatisiert angestoßene Bestellungen.

Eine Cloud-basierte Vorhaltung des entsprechenden DWH- und Analysesystems ist hierbei von Vorteil, da das System damit in kurzer Zeit aufgebaut, getestet und vor allem auch angepasst werden kann. Letzteres betrifft erstens wiederkehrende Herausforderungen wie die Ein- und Anbindung neuer Partner (Hersteller wie auch third- und forth-party logistics provider). Zweitens können unvorhersehbare Ereignisse kurzfristig funktionale Erweiterungen des Systems erfordern, etwa als Reaktion auf unterbrochene Lieferwege aufgrund von Streiks, Unfällen oder Katastrophen. Hier werden ad-hoc analytische Komponenten erforderlich, die in einem begrenzten Zeitraum intensiv genutzt werden. Drittens kann es auch zu grundlegend überarbeiteten Anforderungen kommen, etwa wenn aufgrund veränderter Energiepreisniveaus neue Liefer- und Lagerstrategien verfolgt werden: Je geringer die Lagerbestände und je kleiner somit die Puffer, desto kurzfristiger und detaillierter müssen die RFID-Daten bereitgestellt und vorgehalten werden, was wiederum eine Re-Skalierung des DWH-Systems erforderlich machen kann.

14.5 Innovative BI-Anwendungen in der Produktion

Das Optimierungspotenzial von Produktionsprozessen auf Shop-Floor-Ebene ist begrenzt. Auf Basis eines entsprechend ausgebauten MES sowie von Analysen der damit gesammelten Daten sind zwar in flexiblen Produktionssystemen sehr wohl laufende Feinanpassungen möglich (z. B. hinsichtlich des Werkzeugwechsels oder der Materialauswahl). Wie

jedoch bereits dargestellt wurde, wird der Erfolg industrieller Unternehmen zunehmend dadurch determiniert, inwieweit das *System* Produktion als Ganzes an die sich permanent veränderten Gegebenheiten angepasst wird (Wandlungsfähigkeit).

Eine grundlegende Voraussetzung hierbei ist eine adäquate Versorgung des Top Managements mit entscheidungsrelevanten Informationen. Um valide Entscheidungen im Bereich des strategischen Produktionsmanagements treffen zu können, muss dieses neben den betriebswirtschaftlichen Unternehmenskennzahlen auch semantisch aufbereitete und verdichtete Informationen über den gesamten Produktlebenszyklus enthalten.

Im Folgenden werden zwei Anwendungsbereiche dargestellt, in denen BI-Infrastrukturen einen Beitrag zur Verbesserung der Entscheidungsunterstützung liefern können. In die Ausführungen sind dabei Ergebnisse aus zwei laufenden Forschungsvorhaben eingeflossen.

14.5.1 Produktentwicklung und Produktdatenmanagement

Klassische DWH-Lösungen zielen oftmals auf eine Analyse von Finanz-, Controlling- und Vertriebskennzahlen ab. Speziell bei strategischen Entscheidungen zum Produktionsmanagement bedarf es jedoch zusätzlicher Informationen aus produktorientierten Unternehmensbereichen. Hierzu gehören z. B. Angaben zur Anzahl an Produktvarianten über den Produktlebenszyklus, zu deren Stücklisten und zum Anteil einzelner Teile an der Wertschöpfung.

Speziell innerhalb der Funktionsbereiche F&E finden sich dabei in der Regel eine Vielzahl an isolierten CAx-Systemen, die in unterschiedlichen Phasen des Produktlebenszyklus (z. B. in der Konstruktion, im Qualitätsmanagement oder in der Produktionsmittelpflege) zum Einsatz kommen (Vajna et al. 2009). Je nach Anwendung enthalten die Systeme z. B. Konstruktions-, Qualitäts-, Simulations- oder Steuerungsdaten. Um die Informationen im Entwicklungsprozess konsistent vorhalten zu können, haben sich *Produktdatenmanagement (PDM)* bzw. *Produktlebenszyklusmanagement (PLM)*-Systeme etabliert. Diese stellen in der Regel spezielle Content Management Systeme dar, in denen strukturierte und semistrukturierte Daten wie z. B. Konstruktionszeichnungen, Prüfberichte oder Protokolle gespeichert und verwaltet werden (Eigner und Stelzer 2001).

Damit enthalten diese Systeme für Entscheidungen auf strategischer Ebene relevante Informationen, die für die Entscheidungsträger bisher nur schwer zugänglich sind. Über eine Extraktion dieser Daten aus PDM/PLM-Systemen und deren Aufbereitung in der BI-Infrastruktur wird ein erhebliches Informationsdefizit behoben. Dies kann anhand zweier Beispiele illustriert werden (Lasi 2009):

1. Durch die Verknüpfung von Informationen zu Produktmerkmalen (Produkteigenschaften aus Kundensicht) mit Daten zu dem zugehörigen Entwicklungs- und Produktionsauf-

wand sowie dem bislang erreichten wirtschaftlichen Erfolg können gezielt die Produktgestaltung sowie übergreifende Produktstrategien abgeleitet werden.

2. Treten Reklamationen zu einer Funktion eines Produktes auf, so können Zusammenhänge zwischen den Reklamationen und Konstruktionsmerkmalen bzw. Produktionsverfahren analysiert werden. Dies ermöglicht eine permanente Verbesserung von Produkten und Produktionsprozessen.

14.5.2 MES als Datenquelle für Produktions- und Qualitätsoptimierung

Die schon in Abschn 3.2 vorgestellten MES versprechen eine Zusammenführung von Datenerfassung, Produktionsplanung und -steuerung mit einem integrierten Reporting für das Management des Shop Floor (Kletti 2006). Wie bereits ausgeführt, unterstützen die Systeme neben den unmittelbar operativen Produktionsaufgaben auch managementunterstützende Aufgaben, etwa im Kontext der Materialdisposition, der Auftragsfreigabe, des Qualitäts- oder des Störungsmanagements. Die z. T. optional einbindbaren Reporting- und Analysekomponenten ermöglichen ein laufendes Monitoring des Produktionsgeschehens sowie eine Analyse von Störungsursachen. Ein entsprechend ausgestaltetes und integriertes MES kann damit letztlich als Baustein eines bereichsspezifischen Operational BI verstanden werden.

Die Ausrichtung von MES-Systemen ist jedoch grundsätzlich mengen- und zeitorientiert, was zur Befriedigung der Informationsbedarfe des lokalen Shop-Floor-Managements auch durchaus ausreichend ist. Aufgrund dieser Positionierung ist der Integrationsgrad der Systeme jedoch oftmals begrenzt – Auswertungen über Fabriken, Gewerke, Werke und Standorte sind bislang nicht gängig – womit interessante Analysemöglichkeiten wegfallen (z. B. zu Werkzeug- und Beschaffungsstrategien). Eine Anreicherung mit betriebswirtschaftlichen Kennzahlen (z. B. Beschaffungskosten) erfolgt ebenfalls nicht.

Entsprechend ist es sinnvoll, die Weiterverarbeitung der MES-Daten in einem übergeordneten DWH in Betracht zu ziehen, das entsprechende Integrationsaufgaben leistet und gleichzeitig Kennzahlen für die übergeordnete strategische Steuerung des Unternehmens liefert.

Auch hier sollen zwei Beispiele die Potenziale verdeutlichen:

1. Im MES aufgebaute Datensammlungen können genutzt werden, um Werkzeug- und Materialentscheidungen mit Daten zu durchgeführten Serviceleistungen abzugleichen, um so die Folgen einzelner Produktionsentscheidungen umfassend abwägen zu können.
2. Eine umfassend historisierte Datenbasis auf Basis von MES- und ERP-Daten macht es möglich, einzelne Produktionsanlagen oder ganze Produktionsverbünde auf Konfigurationsparameter hin zu untersuchen und unter Berücksichtigung von deren wertmäßigen Konsequenzen zu optimieren. Dies liefert die Grundlage für Entscheidungen etwa beim Make-or-Buy von Teilen oder Baugruppen.

14.6 Diskussion und Ausblick

Sowohl in der Produktion wie auch in der Logistik finden sich bereits ausdifferenzierte Systemlandschaften. Hierbei werden durchaus Teilaufgaben abgedeckt, die üblicherweise dem BI zugeordnet werden. Dennoch kann festgehalten werden, dass in beiden Bereichen noch Verbesserungspotenziale brach liegen, die auf ähnliche Ursachen zurückzuführen sind:

1. Eine *unzureichende Datenerfassung* – in der Logistik aufgrund des hohen Aufwands zur Objekterfassung mit herkömmlichen Datenerfassungstechnologien, in der Produktion durch eine Vielzahl unabhängiger Insellösungen. Diesem wird u. a. im ersteren Fall mit RFID-basierten Lösungen begegnet, im zweiten durch neue integrative Systeme, insbesondere MES.
2. Ungenutzte Potenziale zur prozess- und funktionsübergreifenden Integration. In der Produktion können *Lösungen* Abhilfe schaffen, die Daten über Fertigungslinien, Werke und Standorte hinweg reichen und diese mit betriebswirtschaftlichen Daten anreichern. In der Logistik ist für eine zielgerichtete Verbesserung eine Qualifizierung von Bestandsdaten durch ergänzende Produkt-, Produktions- und Bedarfsdaten erforderlich.

Hierbei kommt es schnell zu einer Situation, in der ausgereifte DWH-basierte Lösungen relevant werden, wie sie klassischerweise aus dem BI-Bereich bekannt sind. Speziell in größeren Lösungen legen es dabei der zunehmende Integrationsgrad, das durch Technologien wie MES und RFID rasant zunehmende Datenaufkommen sowie die mit der Integration einhergehende funktionsübergreifende Ausrichtung nahe, die entsprechenden Systeme eng mit einer zentralen BI-Einheit, etwa in Form eines BICC, abzustimmen.

Dies bedeutet allerdings weder zwangsläufig, dass damit auch eine technische Zusammenführung der DWHs einhergeht, noch dass die Verantwortung für alle reporting- und analyserelevanten Komponenten abgegeben werden sollte. Im Gegenteil sprechen die hohe Spezialisierung der Aufgaben, die vorhandene Infrastruktur und Expertise und nicht zuletzt auch die Near-Time-Anforderungen der Lösungen für ein gewisses Maß an Dezentralität. So wird ein zentrales und unternehmensweites DWH schwerlich die im MES verankerten Maschinenschnittstellen vereinnahmen und auch nicht in Konkurrenz zu ausgereiften SCM-Lösungen mit deren komplexen Planungskomponenten treten. Zentrale Lösungen können die dezentralen Systeme jedoch punktuell ergänzen – sofern die entsprechenden Datenharmonisierungs-Anforderungen sorgfältig ausgelotet und die Grenzen der jeweiligen Hoheitsgebiete gezielt abgesteckt werden. Hierbei sollten auch die Möglichkeiten Cloud-basierter Lösungen zur Steigerung der Agilität gerade in den unternehmensübergreifenden Szenarios der Logistik sowie der Einbindung von Big-Data-Ansätzen zur Aufbereitung der heterogenen und komplex strukturierten Datenbestände aus der Produktion geprüft werden. Abbildung 14.3 illustriert diese Schlussfolgerungen.

Auf dieser Grundlage lassen sich dann auch deutlich weiterführende Anwendungen konzipieren – beispielsweise durch die Einbindung von Kundenanforderungen in Logis-

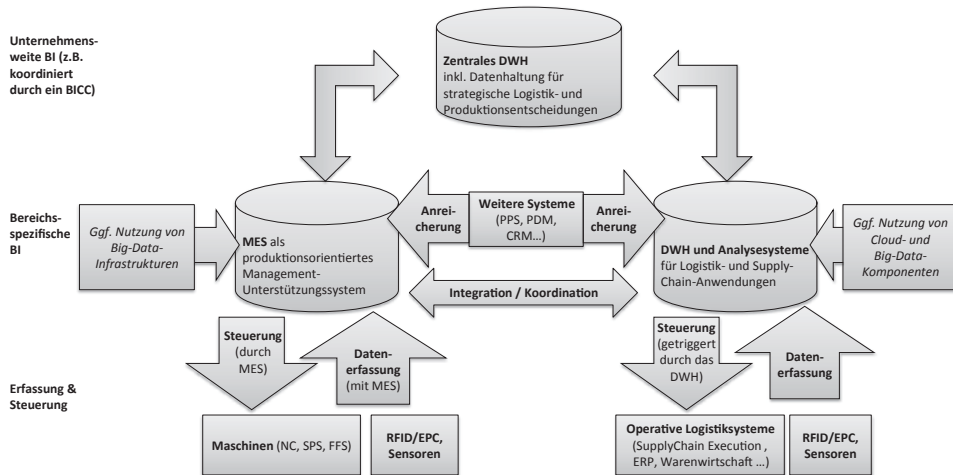


Abb. 14.3 Zentrale und dezentrale BI in Produktion und Logistik

tikanalysen, die Untersuchung von Wechselwirkungen zwischen Produktions- und Lieferentscheidungen oder eine Abstimmung von Produktions- und Servicedaten zur Optimierung von Produktionsprozessen.

Literatur

- Alexander, K., Gilliam, T., Gramling, K., Kindy, M., Moogimane, D., Schultz, M., Woods, M.: Focus on the Supply Chain: Applying Auto-Id Within the Distribution Center. AutoID-Labs Whitepaper, Cambridge (2002)
- Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Hrsg.): Handbuch Logistik, 3. Aufl. VDI Buch, Berlin (2008)
- Baars, H., Kemper, H.G.: Business intelligence in the cloud? Proceedings of the 14th Pacific Asia Conference on Information Systems (PACIS), 09.12.07, Taipeh, Taiwan, S. 1529–1539 (2010)
- Baars, H., Kemper, H.G.: Ubiquitous computing – an application domain for business intelligence in the cloud? Proceedings of the 17th Americas Conference on Information Systems (AMCIS), 04.08.–07.08.2011, Detroit, USA (2011)
- Baars, H., Sun, X.: Multidimensional analysis of RFID data in logistics. Proceedings of the 42nd Hawaii International Conference on System Sciences (HICCS-42), Hawaii (2009)
- Baars, H., Kemper, H.G., Siegel, M.: Combining RFID technology and business intelligence for supply chain optimization – scenarios for retail logistics. Proceedings of the 41st Hawaii International Conference on System-Sciences (HICCS-41), Hawaii (2008)
- Baars, H., Felden, C., Gluchowski, P., Hilbert, A., Kemper, H.G., Olbrich, S.: Shaping the next incarnation of business intelligence. BISE. 6(1), 11–16 (2014a)
- Baars, H., Funke, K., Müller, P.A., Olbrich, S.: Big Data als Katalysator für die Business Intelligenz – Das Beispiel der informa Solutions GmbH. HMD – Prax Wirtsch. 51(4), 436–446 (2014b)
- Bauernhansl, T., ten Hompel, M., Vogel-Heuser, B. (Hrsg.): Industrie 4.0 in Produktion, Automatisierung und Logistik – Anwendung, Technologien und Migration, Wiesbaden (2014)

- Bea, F.X., Göbel, E.: *Organisation – Theorie und Gestaltung*, 2. Aufl. Utb Fuer Wissenschaft, Stuttgart (2002)
- Behler, A., Römer, S., Trunte, T.: Automotive supply chain monitoring – Einsatzfeld für data warehouse Lösungen. *Inf. Manag. Consult.* **18**(4), 52–57 (2003)
- Betge, D.: *Koordination in Advanced Planning and Scheduling-Systemen*. Deutscher Universitätsverlag, Wiesbaden (2006)
- Bi, Z., Xu, L.D., Wang, C.: Internet of things for enterprise systems of modern manufacturing. *IEEE Trans. Ind. Inform.* **10**(2), 1537–1546 (2014)
- Brewer, A., Sloan, N., Landers, T.: Intelligent tracking in manufacturing. *J. Intell. Manuf.* **20**(10), 245–250 (1999)
- BVL (Bundesvereinigung Logistik): *Trends and Strategies in Logistics and Supply Chain Management 2013*. BVL, Hamburg (2013)
- Chandra, C., Grabis, J., Tumanyan, A.: Problem taxonomy: a step towards effective information sharing in supply chain management. *Int. J. Prod. Res.* **45**(11), 2507–2544 (2007)
- Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics: from big data to big impact. *MIS Quart.* **36**(4), 1165–1188 (2012)
- Chopra, S., Meindl, P.: *Supply Chain Management – Strategy, Planning & Operations*, 5. Aufl. Pearson Prentice Hall, Upper Saddle River (2012)
- Christopher, M.: *Logistics and Supply Chain Management – Strategies for Reducing Cost and Improving Service*, 4. Aufl. London (2010)
- Codd, E.F., Codd, S.B., Salley, C.T.: Providing OLAP to User-Analysts: An IT Mandate, o. O. (1993)
- Corsten, H.: *Produktionswirtschaft – Einführung in das industrielle Produktionsmanagement*, 11. Aufl. Oldenbourg, München (2007)
- DeGroote, S.E., Mar, T.G.: The impact of IT on supply chain agility and firm performance: an empirical investigation. *Int. J. Inf. Manage.* **33**, 909–916 (2013)
- Eigner, M., Stelzer, R.: *Produktdatenmanagement-Systeme: Ein Leitfaden für Product Development und Life Cycle Management*. Springer, Berlin (2001)
- Finkenzeller, K.: *RFID-Handbuch – Grundlagen und praktische Anwendungen von Transpondern, kontaktlosen Chipkarten und NFC*, 6. Aufl. München und Wien (2012)
- Fleisch, E., Dierkes, M.: Ubiquitous computing aus betriebswirtschaftlicher sicht. *Wirtschaftsinformatik* **45**(6), 611–620 (2003)
- Friedli, T.: *Technologiemanagement: Modelle zur Sicherung der Wettbewerbsfähigkeit*. Springer, Berlin (2006)
- Gansor, T., Totok, A.: *Von der Strategie zum Business Intelligence Competency Center (BICC) – Von der Strategie zum Business Intelligence Competency Center (BICC)*, 2. Aufl. dpunkt, Heidelberg (2015)
- Gienke, H.: Ziele, Strategien und Aufgaben der Produktionsplanung und -steuerung. In: Gienke, H., Kämpf, R. (Hrsg.) *Handbuch Produktion: Innovatives Produktionsmanagement: Organisation, Konzepte, Controlling*, S. 551–744. Hanser, München (2007)
- Gienke, H., Kämpf, R.: *Handbuch Produktion: Innovatives Produktionsmanagement. Organisation, Konzepte, Controlling*. Hanser, München (2007)
- Gleißner, H., Femerling, C.: *Logistik – Grundlagen – Übungen – Fallbeispiele*. Gabler, Wiesbaden (2008)
- Gluchowski, P., Kemper, H.G., Seufert, A.: Innovative Prozess-Steuerung – Was ist neu an Operational BI? *BI-Spektrum* **4**(1), 8–12 (2009)
- Golfarelli, M., Rizzi, S., Cella, I.: Beyond data warehousing: what's next in business intelligence? *Proceedings of the 7th ACM International Workshop on Data Warehousing and OLAP*, S. 1–6. New York (2004)
- GS 1: EPCglobal-Standards Overview. <http://www.gs1.org/gsm/kc/epcglobal> (2014). Zugriffen: 11 Nov. 2014
- Günther, O., Kletti, W., Kubach, U.: *RFID in Manufacturing*. Springer, Berlin (2008)

- Hand, D., Mannila, H., Smyth, P.: Principles of Data Mining. The MIT Press, A Bradford Book, Cambridge (2001)
- Helo, O., Szekely, B.: Logistics information systems – an analysis of software solutions for supply chain co-ordination. *Ind. Manage. Data Syst.* **105**(1), 5–18 (2005)
- Hofmann, E., Bachmann, H.: Behälter-Management in der Praxis: Die St. Galler Behälter-Management-Studie. State of the Art und Entwicklungstendenzen bei der Steuerung von Ladungsträgerkreisläufen, DVV Media Group, Hamburg (2006)
- Ilic, B., Anderson, T., Michahelles, F.: EPCIS-Based Supply Chain Visualization Tool. Auto-ID-Labs Whitepaper, St. Gallen (2009)
- Inmon, W.H.: Building the Data Warehouse. John Wiley & Sons, New York (2005)
- IT & Production: Manufacturing Execution Systems (MES) 2014/15. Marburg (2014)
- Kambil, A., Brooks, J.D.: Auto-ID Across the Value Chain: From Dramatic Potential to Greater Efficiency & Profit. AutoID-Labs Whitepaper, Cambridge (2002)
- Kemper, H.G., Baars, H., Mehanna, W. Business Intelligence – Grundlagen und praktische Anwendungen, 3. Aufl. Vieweg+Teubner, Wiesbaden (2010)
- Kletti, J. (Hrsg.): MES Manufacturing Execution System – Moderne Informationstechnologie zur Prozessfähigkeit der Wertschöpfung. Springer, Berlin (2006)
- Koch, M., Baars, H.: Analyzing RFID data for the management of reusable packaging. Proceedings of the 4th Mediterranean Conference on Information Systems (MCIS 09), 25–27.09.2009, Athens (2009)
- Lasi, H.: Aufbau eines IT-basierten Integrationskonzepts zur Unterstützung von Produktentwicklungs- und Produktionsprozessen. Eul, Lohmar (2009)
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M.: Industrie 4.0. *Wirtschaftsinformatik* **4**, 261–264 (2014)
- Leclerc, M.: Behälterlogistik als zentrale Aufgabe des Supply Chain Management. *Suppl. Chain Manage.* **2**(3), 33–36 (2002)
- Lee, H.L., Padmanabhan, V., Whang, S.: Information distortion in a supply chain: the bullwhip effect. *Manage. Sci.* **43**(4), 546–558 (1997)
- Ma, M., Wang, P., Chu, C.H.: Data management for Internet of things: challenges, approaches and opportunities. Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing (2013)
- Mertens, P., Meier, M.C.: Integrierte Informationsverarbeitung 2 – Planungs- und Kontrollsysteme in der Industrie, 10. Aufl. Gabler, Wiesbaden (2009)
- METROGroup: RFID Uncovering the Value Applying RFID Within the Retail and Consumer Package Goods Value Chain. o. O. (2004)
- Meyr, H., Rosič, H., Seipl, C., Wagner, M., Wetteraure, U.: Architecture of selected APS. In: Stadler, H., Kilger, C. (Hrsg.) Supply Chain Management and Advanced Planning – Concepts, Models, Software, and Case Studies, 4. Aufl., S. 349–366. Springer, Berlin (2008)
- Miller, G.J., Brautigam, D., Gerlach, S.V.: Business Intelligence Competency Centers: A Team Approach to Maximizing Competitive Advantage. Wiley, Hoboken (2006)
- Morán, H.J., Ayub, S., McFarlane, D.: Auto-ID Use Case: Improving Inventory Visibility in a Retail Company – Impact on existing Procedures and Information Systems. AutoID-Labs Whitepaper, Cambridge (2004)
- Musa, A., Gunasekaran, A., Yusuf, Y.: Supply chain product visibility – methods, systems and impact. *Expert. Syst. Appl.* **41**, 176–194 (2013)
- Niederman, F., Mathieu, R.G., Morley, R., Kwon, I.W.: Examining RFID applications in supply chain management. *Commun. ACM* **50**(7), 92–101 (2007)
- Pawellek, G.: Produktionslogistik: Planung – Steuerung – Controlling. Carl Hanser, München (2007)

- Sandler, N.R.: Big Data Driven Supply Chain Management: A Framework for Implementing Analytics and Turning Information into Intelligence. Pearson FT Press, Pearson Education, Upper Saddle River (2014)
- Schmalzried, D., Cundius, C., Franke, R., Lambeck, C., Alt, R., Zimmermann, W., Groh, W.: In-Memory basierte Real-Time Supply Chain Planung. Proceedings of the 11th International Conference on Wirtschaftsinformatik, 27.02.–01.03. 2013. Leipzig, Germany (2013)
- Schuh, G., Gierth, A.: Grundlagen der Produktionsplanung und -steuerung. In: Schuh, G. (Hrsg.) Produktionsplanung und -steuerung – Grundlagen, Gestaltung und Konzepte, 3. Aufl., S. 11–27. Springer, Berlin (2006)
- Strassner, M., Plenge, C., Stroh, S.: Potenziale der RFID-Technologie für das Supply Chain Management in der Automobilindustrie. In: Fleisch, E., Mattern, F. (Hrsg.) Das Internet der Dinge – Ubiquitous Computing und RFID in der Praxis: Visionen, Technologien, Anwendungen, Handlungsanleitungen, S. 177–196 (2005)
- Sun, X.: Ein szenario- und prototypenbasiertes Konzept zur Informationsbedarfsanalyse für Business-Process-Intelligence-Systeme. Dissertation, Universität Stuttgart, Eul, Lohmar (2014)
- Tellkamp, C.: The impact of Auto-ID technology on process performance – RFID in the FMCG supply chain. Dissertation, Universität St. Gallen (2006)
- Thomson, W.J.J., Van Der Walt, J.S.: Business intelligence in the cloud. South Afr. J. Inf. Manage. 2(1), 1 (2010)
- Vajna, S., Weber, C., Bley, H., Zeman, K.: CAx für Ingenieure – Eine praxisbezogene Einführung. Springer, Berlin (2009)
- VDI Richtlinie 5600: Fertigungsmanagementsystem: Manufacturing Execution Systems (MES). VDI, o. O. (2007)
- Westkämper, E.: Nachhaltige Strategien der Produktion am Standort Deutschland. In: Zahn, E. (Hrsg.) Erfolgreich produzieren am Standort Deutschland – Eine strategische Herausforderung, S. 25–52. Lemmens Medien, Bonn (2007)
- Wildemann, H.: Logistik Prozessmanagement. TCW, München (2005)
- Yan, J., Xin, S., Liu, Q., Xu, W., Yang, L., Fan, L., Chen, B., Wang, Q.: Intelligent supply chain integration and management based on cloud of things. Int. J. Distrib. Sens. Netw. Article ID 624839, 15 S. (2014)

Jörn Kohlhammer, Dirk U. Proff und Andreas Wiener

Inhaltsverzeichnis

15.1 Einführung 304

15.2 Das Visual-Business-Analytics-Modell 305

15.3 Information Design und Reporting 307

 15.3.1 Dekoration 307

 15.3.2 3-D-Diagramme 309

 15.3.3 Skalierung 309

 15.3.4 Einsatz von Farben, schlanke Visualisierung und hohe Informationsdichte. . . 309

15.4 Visual Business Intelligence und Dashboarding 311

15.5 Visual Analytics und Big Data 315

 15.5.1 Big Data 315

 15.5.2 Visual Analytics 317

 15.5.3 Aktueller Einsatz in Unternehmen 318

 15.5.4 Anwendungsbeispiel 321

15.6 Zusammenfassung 322

Literatur 323

J. Kohlhammer (✉)
Fraunhofer-Institut für Graphische Datenverarbeitung, Darmstadt, Deutschland
E-Mail: Joern.Kohlhammer@igd.fraunhofer.de

D. U. Proff
blueforte GmbH, Hamburg, Deutschland
E-Mail: dirk.proff@blueforte.com

A. Wiener
reportingimpulse GmbH, Hamburg, Deutschland
E-Mail: Andreas.Wiener@reportingimpulse.com

Zusammenfassung

Visualisierung wird im Bereich Business Intelligence (BI) und Business Analytics immer wichtiger. Heutige BI-Software, gerade im Bereich der Self-Service-BI, wird von den Anwendern ohne ein interaktives Benutzerinterface, das ihnen verschiedene Visualisierungsmöglichkeiten bietet, nicht mehr akzeptiert. Dennoch muss Unternehmen und Anwendern noch besser dargelegt werden, wie sie von solchen Visualisierungslösungen am Ende tatsächlich profitieren können. Dieser genaueren Betrachtung der Vorteile widmet sich der Bereich Visual Business Analytics (VBA), der drei Teilthemen umfasst: Information Design (ID), Visual Business Intelligence (VBI) und Visual Analytics (VA). Dieser Beitrag soll anhand illustrativer Beispiele in diese Teilthemen einführen.

15.1 Einführung

Dank der modernen Informations- und Sensortechnologie werden in der heutigen Business-Welt Datenmengen generiert und gespeichert, die exponentiell wachsen. In vielen datengetriebenen Unternehmen haben das Volumen, die Vielfalt und die Dynamik der Daten ein solches Ausmaß angenommen, dass zu ihrer Beschreibung zurecht der Begriff Big Data verwendet wird (Beyer und Lancy 2012). Gelingt es Unternehmen, das mit solchen großen Datenmengen verbundene Potenzial über Business Intelligence (BI) bzw. ein entsprechendes System zu erschließen, kann die übliche periodische Berichterstattung für Fach- und Führungskräfte um wertvolle Informationen ergänzt werden. So können Entscheider dabei unterstützt werden, rechtzeitig Trends zu erkennen und Planungs- und Kontrollaufgaben wahrzunehmen. BI-Anwender wie Controller und Business-Analysten bekommen mehr Möglichkeiten, Informationen zu kombinieren, zu analysieren und daraus aussagekräftige Berichte für die Entscheider zu generieren.

Oft mangelt es jedoch an Ressourcen wie Zeit oder der richtigen Informationstechnologie, um alle verfügbaren Daten auswerten und sinnvoll für anstehende Entscheidungen aufbereiten zu können [KoPW13]. Bis heute reizen viele Unternehmen das Potenzial, das gespeicherte Daten ihnen bieten, nur unzureichend aus. Um aus internen und externen Daten einen Wert schöpfen zu können, sind sie auf zeitgemäße und menschengerechte Systeme angewiesen, die es ihnen erlauben, in kurzer Zeit entscheidungsrelevante Rohdaten zu generieren, zu strukturieren und für das gesamte Adressatenspektrum von BI-Anwendern und Entscheidern aufzubereiten.

Der Schlüssel hierzu ist das richtige Visualisieren der Daten. Denn der Mensch wird weiterhin im Mittelpunkt stehen und darauf angewiesen sein, in Daten bedeutende Muster, Strukturen und Trends zu erkennen und hieraus die richtigen Entscheidungen abzuleiten. Je nach Rolle stellen die unterschiedlichen Adressaten im Unternehmen jedoch sehr individuelle Anforderungen an die Struktur und Aufbereitung ihrer Daten und Berichte.

Datenanalysten (sogenannte Data Scientists) erschließen mithilfe von spezieller Analysesoftware und komplexen statistischen Methoden polystrukturierte Daten (Manyika et al. 2011). Sie bedienen sich dabei häufig visueller Methoden, die sie dabei unterstützen, Muster in großen Datenmengen visuell zu erfassen. Dieser Vorgang ist die Voraussetzung für das Extrahieren relevanter Informationen, die den BI-Anwendern als strukturierte Daten zur Verfügung gestellt werden. BI-Anwender analysieren im Anschluss daran die Daten und stellen sie Entscheidern in Form von Berichten zur Verfügung. In allen drei Disziplinen kommen Visualisierungen zum Einsatz.

Jede dieser drei Nutzerrollen stellt sehr individuelle Anforderungen an die Visualisierung von Daten. An dieser Stelle setzt dieser Beitrag an. Der Leser erhält einen Überblick über das Visual-Business-Analytics-Modell (VBA-Modell), das die drei verschiedenen Disziplinen der Visualisierung von Daten innerhalb von analytischen Informations- und Entscheidungssystemen beschreibt und mit den Nutzerrollen verknüpft [KoPW13]. Anhand typischer Fehler im Berichtswesen soll der Leser für die Bedeutung einer hohen Qualität bei Visualisierungen sensibilisiert werden. Ausgehend davon, befasst sich der Beitrag mit einem Praxisbeispiel der Otto Group. Im Konditionenmanagement der Firma kommen nach einem Visualisierungsprojekt Dashboards zum Einsatz, die sich an den Regeln für Information Design (ID) nach Prof. Dr. Hichert orientieren. Der Leser erhält neben klassischen Dashboards einen zusätzlichen Vorschlag für interaktive Visualisierungen, die sich gut für Dashboards in Unternehmen einsetzen lassen. Anschließend geht der Buchbeitrag auf das Thema Visual Analytics (VA) ein. Es wird erläutert, wie mithilfe von modernen Auswertungsmethoden und Technologie in Verbindung mit den Erkenntnissen menschlicher Wahrnehmung der Zugang zu unstrukturierten Daten von enormer Menge hergestellt werden kann. Eine abschließende Zusammenfassung stellt die wesentlichen Aussagen des Beitrags noch einmal auf prägnante Weise heraus und zeigt Perspektiven für die zukünftige Entwicklung des Themas für die interne Informationsverarbeitung und für das Berichtswesen in Unternehmen auf.

15.2 Das Visual-Business-Analytics-Modell

Das VBA-Modell geht von drei Wegen der Entscheidungsfindung aus, die sich in erster Linie bezüglich vier Parametern unterscheiden. Diese Parameter lauten: Nutzer, Einsatzgebiete, Daten und Visualisierungen (Vgl. Abb. 15.1).

Die bekannteste und am weitesten verbreitete Vorgehensweise, Daten und Informationen in Unternehmen visuell darzustellen, ist das ID. Dabei werden Daten in Form von Diagrammen und Texten in Tabellenkalkulations- und Präsentationsprogrammen aufbereitet, um damit den Entscheidern die Entscheidungsfindung zu erleichtern. In Unternehmen übernimmt diese Aufgabe üblicherweise das Controlling. In regelmäßigen Zeitintervallen (Tages-, Wochen-, Monatsberichten) erhalten die Entscheider über ein internes Berichtswesen oder Reporting aktuelle Geschäftszahlen. Die betreffenden Berichte sind meist statisch und werden als PowerPoint-Folien oder in Papierform erstellt. Somit hat der

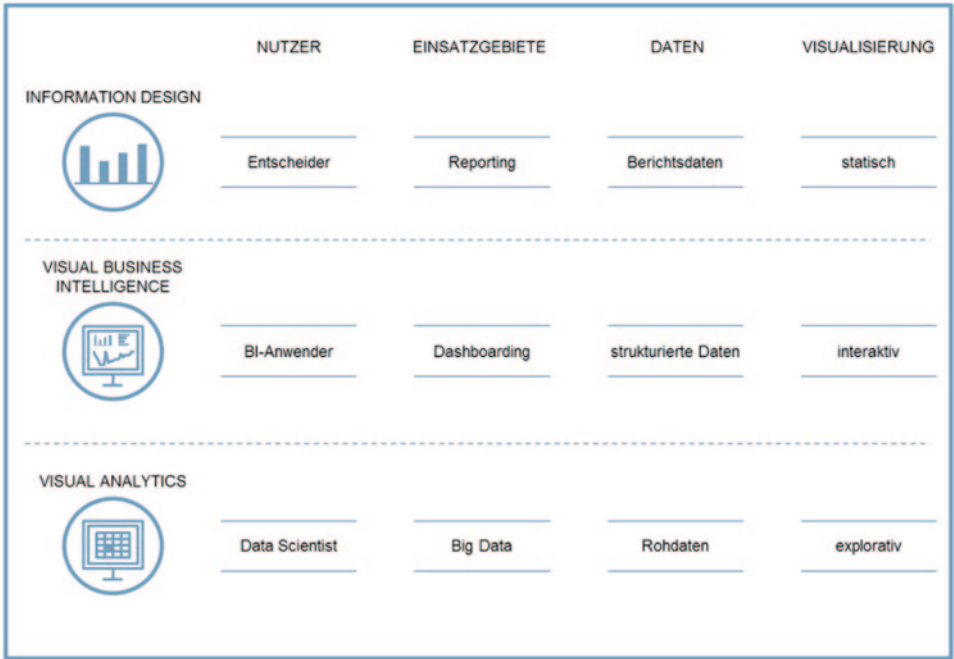


Abb. 15.1 Das VBA-Modell [KoPW13]

Empfänger keine Möglichkeit, Einfluss auf die Art und Weise der Darstellung zu nehmen [KoPW13]. Welche Visualisierungen am besten für ein gutes Reporting geeignet sind und welche „Regeln“ es zu beachten gilt, soll im späteren Verlauf des Artikels genauer beleuchtet werden.

In den letzten Jahren wird das klassische Reporting durch die Visual Business Intelligence (VBI) ergänzt. Dazu tragen vor allem Dashboards (Informationscockpits) bei, die es den BI-Anwendern ermöglichen, sich interaktiv durch die Daten zu bewegen. BI-Anwender sind in der Regel Mitarbeiter, die sehr analytisch arbeiten. Sie sind häufig in einem bestimmten Fachbereich als Business Analyst oder als Financial Analyst im Controlling tätig. Sie treffen eigenständig Entscheidungen oder versorgen Fach- und Führungskräfte über das Berichtswesen mit Entscheidungsvorlagen oder Empfehlungen.

Die VBI unterscheidet sich zum ID bzw. Reporting im Wesentlichen darin, dass sie interaktiver ist und tiefgreifende Analysen möglich macht. In Dashboards kann der Nutzer drillen, filtern und die Darstellung der Daten seinen Wünschen anpassen. Zudem kommen auch komplexere Diagramme als die herkömmlichen Balken-, Säulen- und Kreisdiagramme zum Einsatz. Als Beispiel seien hier Treemaps, Netzdiagramme oder interaktive Blasendiagramme genannt, die tiefergehende Analysen ermöglichen als die gewohnten Darstellungsformen.

Die letzte Disziplin im VBA-Modell ist VA, ein Ansatz, der sich dem visuellen Umgang mit Big Data widmet. VA setzt dort an, wo herkömmliche Diagramme nicht mehr

ausreichen, um große Datenmengen abzubilden. Dabei führen in Unternehmen Spezialisten, sogenannte Data Scientists, mit Spezialwerkzeugen Analysen durch. Der Unterschied zum ID und VBI, wo die Daten (vorstrukturiert) vorgegeben sind, ist, dass VA interaktiv und explorativ ist. Das heißt, dass Data Scientists am Anfang ihrer Untersuchung noch gar nicht wissen, was sie am Ende finden werden. VA erfordert gut ausgebildete Data Scientists, denn um solche Analysen durchführen zu können, sollten sie sich sowohl in komplexen Visualisierungstechniken als auch im Umgang mit Werkzeugen und Programmierschnittstellen auskennen [KoPW13].

15.3 Information Design und Reporting

Der Controller in (Groß-) Unternehmen steht in der Regel nicht mehr vor der Herausforderung, Daten für seine Berichte zu erhalten. Vielmehr stellt er sich die Frage, wie er sie auf geeignete Weise darstellt, um beim Management möglichst objektive Entscheidungen herbeizuführen. Dazu hat er drei Möglichkeiten:

- Visuell (z. B. Diagramme, Schaubilder oder Tab.)
- Textuell (z. B. E-Mails oder Texte)
- Verbal (z. B. Gespräche oder Präsentationen)

Im Folgenden sollen vor allem für die visuelle Aufbereitung von Daten Empfehlungen gegeben werden. Viele der gezeigten Beispiele sind mit klassischen Tabellenkalkulations- und Präsentationsprogrammen erstellt.

Um aufzuzeigen, wie gutes ID funktioniert, lohnt es sich, im ersten Schritt Beispiele für schlechtes ID anzuschauen. Diese machen schnell deutlich, welche Fallen in puncto Design auf diejenigen lauern, die solche Visualisierungen erstellen. Alle dargestellten Empfehlungen gelten auch für den darauffolgenden Abschnitt „Visual Business Intelligence und Dashboarding“.

15.3.1 Dekoration

Eine Dekoration von Daten in Diagrammen wird meist dann verwendet, wenn das Dargestellte Gefahr läuft, allzu banal und langweilig zu sein. Dadurch verspüren die Ersteller den Drang, das Dargestellte zu schmücken bzw. „aufzuhübschen“, wie Abb. 15.2 verdeutlicht.

Alle gezeigten Beispiele in Abb. 15.2 sind Geschäftsberichten entnommen, die in den letzten Jahren zu den Top 10 der besten Geschäftsberichte in Deutschland zählten. Das erste Beispiel zeigt ein Säulendiagramm mit aufeinandergestapelten Abbildungen von Münzen als Säulen. Wie nicht schwer zu erraten, stammt dieses Beispiel aus dem Geschäftsbericht einer Bank. Anstatt ein klassisches Säulendiagramm zu wählen, entschied sich der Ersteller für eine Darstellung, die jedoch denkbar schlecht geeignet ist, die

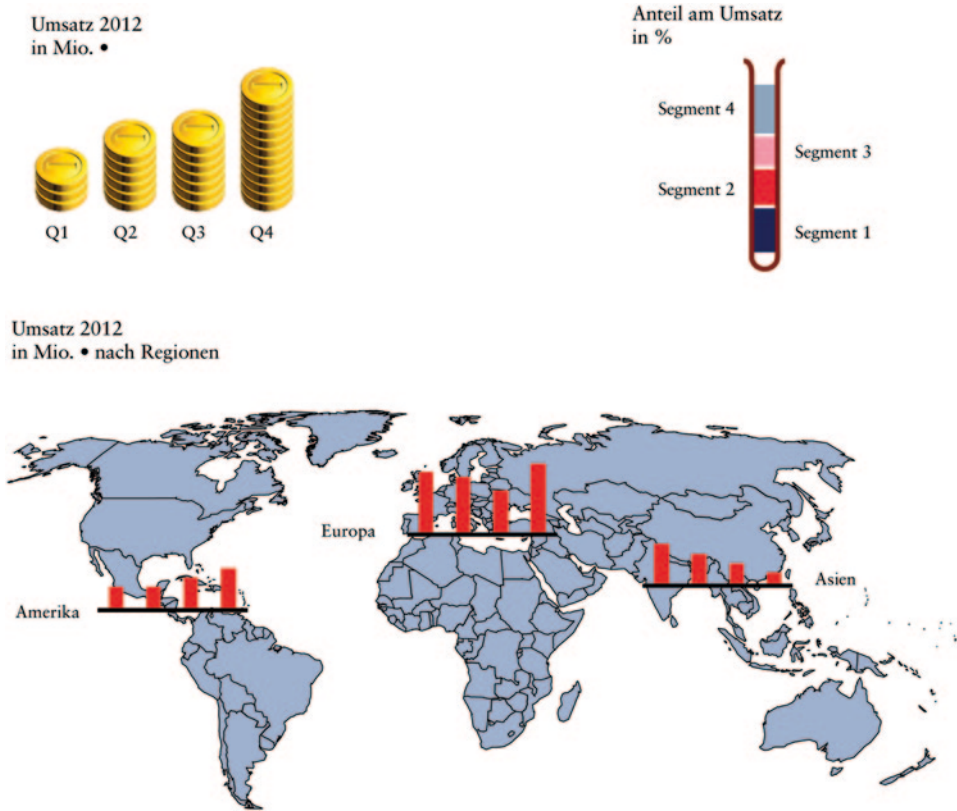


Abb. 15.2 Nachbau von dekorierten Diagrammen aus Geschäftsberichten [KoPW13]

unterschiedlichen Höhen der Säulen genau zu erkennen. Hier wurde der Dekoration der Vorzug vor der Lesbarkeit gegeben.

Ähnlich verhält es sich mit dem zweiten Beispiel: Ein Pharmakonzern entschied sich, seine Segmente in einem Reagenzglas darzustellen. Auch diese Dekoration geht klar zu Lasten der Lesbarkeit und lenkt den Empfänger vom Wesentlichen ab, nämlich dem Verhältnis der Segmente zueinander.

Das dritte Beispiel zeigt die in Präsentationen häufig zu findende Landkarte. Der Umsatz pro Quartal wird für jede Region als Säulendiagramm dargestellt. Der Ersteller hat in diesem Beispiel zur Dekoration eine Weltkarte in den Hintergrund gesetzt. Die Umsätze sind den Regionen Amerika, Europa und Asien zugeordnet, doch es ist vollkommen unnötig, zu zeigen, wo auf der Weltkarte diese Regionen liegen – das weiß der Empfänger, der sich nur den Umsatz ansehen möchte, ohnehin. Das eigentliche Ziel der Visualisierung, die Unterschiede pro Quartal und Region darzustellen, wird hier durch die Landkarte zusätzlich erschwert. Durch eine Darstellung aller Diagramme auf Höhe ihrer X-Achsen könnte der Empfänger die Quartalsumsätze der Regionen wesentlich besser miteinander vergleichen.

15.3.2 3-D-Diagramme

In vielen Berichten werden 3-D-Diagramme verwendet, da diese den Empfänger angeblich mehr ansprechen. Der Einsatz von 3-D sollte jedoch vermieden werden, denn diese Form der Darstellung führt dazu, dass der Leser die einzelnen Werte nicht so gut zuordnen kann. Die VBA-Studien 2012 und 2013 haben diesen Umstand eingehend untersucht (Kohlhammer et al. 2012; Kohlhammer et al. 2013a). Die Teilnehmer erhielten dabei die Aufgabe, in einem 3-D-Diagramm und in einem 2-D-Diagramm den gleichen Wert zu schätzen. Beim 3-D-Diagramm wurde diese Aufgabe 7-mal schlechter bewältigt als beim 2-D-Diagramm. Zudem wirkt 3-D mitunter zu verspielt; man stelle sich nur einmal vor, dieser Artikel wäre in einer 3-D-Schrift verfasst – die Skepsis der Leser gegenüber dem Inhalt des Artikels wäre vorprogrammiert.

15.3.3 Skalierung

Viele Ersteller verschwenden keinen Gedanken an die Skalierung ihrer Diagramme. Oftmals wird automatisch diejenige Skalierung verwendet, die das Tabellenkalkulations- oder Präsentationsprogramm dynamisch vorschlägt. Die Skalierung orientiert sich in diesem Fall am höchsten Wert und beschneidet bei Bedarf die Y-Achse so, dass sie nicht bei null beginnt (Bosbach und Korff 2011).

Die große Bedeutung einer einheitlichen Skalierung zeigt Abb. 15.3. Das obere Beispiel vermittelt dem Betrachter einen falschen Eindruck von den in den beiden Diagrammen dargestellten Werten: Die 85 im linken Diagramm ist höher als die 625 im rechten Diagramm. Im unteren Beispiel wird eine einheitliche Skalierung verwendet; die 85 steht jetzt – wie auch alle anderen Werte – zur 625 im richtigen Verhältnis. Erst dadurch wird es möglich, beide Diagramme visuell miteinander zu vergleichen. Als Faustregel sollte gelten, dass auf einer Berichts- oder Präsentationsseite stets einheitlich skaliert wird. Dies ist besonders dann wichtig, wenn Entscheider nur wenig Zeit haben und auf Basis der Diagramme schnell eine Entscheidung treffen wollen oder müssen. Eine uneinheitliche Skalierung birgt stets die Gefahr, dass ein falscher Eindruck von den Verhältnissen entsteht und dadurch falsche Schlüsse gezogen werden.

15.3.4 Einsatz von Farben, schlanke Visualisierung und hohe Informationsdichte

Gutes ID befolgt im Wesentlichen drei Grundregeln [KoPW13]:

- Intelligenter Einsatz von Farben (keine Willkür!)
- Klare, präzise Darstellungen ohne Ablenkung (keine Dekoration!)
- Viele Informationen (keine Banalitäten!)

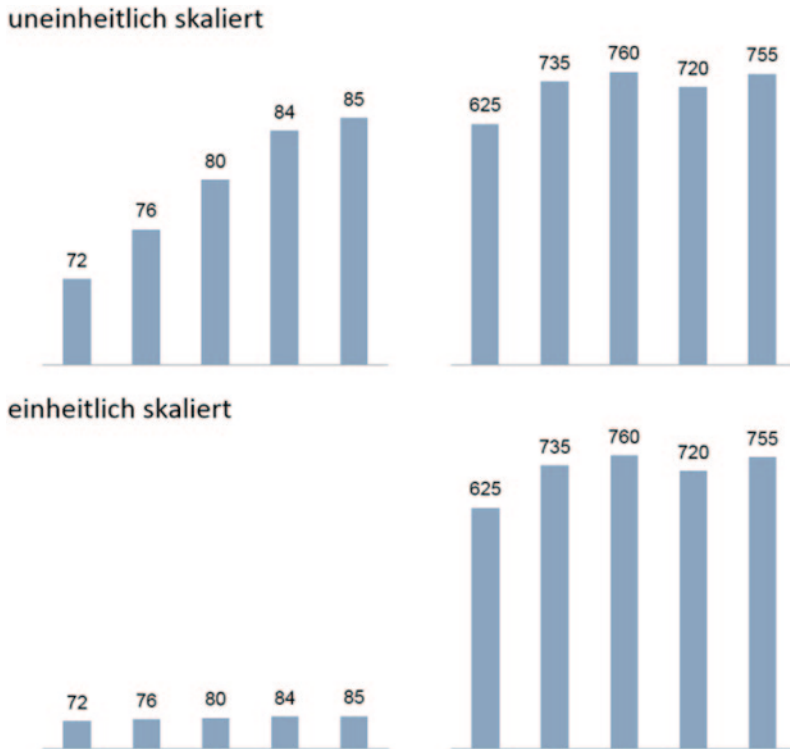


Abb. 15.3 Einheitliche und uneinheitliche Skalierung bei kleinen und großen Werten [KoPW13]

Das Vorher-Nachher-Beispiel in Abb. 15.4 aus der Marketing-Broschüre des TDWI zeigt beispielhaft, wie Informationen mithilfe von gutem ID dargestellt werden können.

Beide abgebildeten Seiten zeigen Teilnehmerstatistiken der TDWI Konferenz in München 2013. Auf der linken Seite werden die Daten mittels Kreisdiagrammen, einem doppelten Säulendiagramm und einer Vielzahl von Farben dargestellt. Der Nachteil an Kreisdiagrammen ist, dass der Leser nicht auf einen Blick sieht, welcher Wert zu welcher Farbe gehört. Er muss sich den Wert merken und die Farbe anschließend in der Legende suchen. Zudem werden die Farben nicht einheitlich eingesetzt. So wird die Farbe Rot mehrfach verwendet (öffentliche Hand, USA, Jemen, Frankreich, Top-Management, Teilnehmer und Gäste). In Kreisdiagrammen sind vor allem kleine Werte schwer zu lesen und zuzuordnen.

Auf der rechten Seite wird ein einheitliches Farbkonzept verwendet. Jede Farbe hat eine feste Bedeutung: Hellblau für Vorjahre, Dunkelblau für das aktuelle Jahr und schraffiert für die Prognose. Zeitreihen werden horizontal und Verhältnisse vertikal dargestellt. Die kleinen Bildchen sind verschwunden und wurden durch zusätzliche Informationen ersetzt. Überflüssige bzw. redundante Darstellungen wie die Y-Achse oder die Hilfslinien wurden entfernt. Im rechten Beispiel werden die drei oben erwähnten Grundregeln eingehalten. Der TDWI verwendet seit 2013 in seiner Broschüre eine Darstellung mit gutem ID.



Abb. 15.4 Vorher-Nachher-Beispiel aus einer Marketing-Broschüre des TDWI (Proff et al. 2013)

15.4 Visual Business Intelligence und Dashboarding

Die im vorangegangenen Abschnitt thematisierten Regeln für gutes ID gelten fast durchgängig auch für VBI bzw. für das Gestalten guter Dashboards. Hier liegen weitere Herausforderungen, nun allerdings in der Interaktivität der Darstellungen sowie den Scroll- und Drill-Down-Möglichkeiten. Eine gute Lösung hat die Otto Group in ihr Konditionenmanagement implementiert: klare Regeln, Verzicht auf Scrollbalken und die Verwendung einheitlicher Darstellungen bei Drill-Downs.

Das Dashboard zeigt den Umsatz sowohl des Konzerns als auch der dazu gehörigen Einzelgesellschaften der Otto Group pro Saison. Analog zu den Regeln, die für Reports gelten, werden auch in Dashboards bei zeitlichen Verläufen vergangene Werte grau und die aktuelle Periode schwarz dargestellt; dies erleichtert dem Leser die Orientierung. So auch in diesem Dashboard, das sich auf den Herbst-Winter-Zeitraum (HW) bezieht: Die zwei grauen Säulen stellen die Vorsaisons dar, die schwarze Säule die aktuelle Saison. Zudem sind sowohl die absoluten als auch die prozentualen Abweichungen zur Vorsaison über den Säulendiagrammen pro Einzelgesellschaft und Konzern dargestellt. Positive Werte werden grün dargestellt, negative rot. Die absoluten Abweichungen zur Vorsaison ($\Delta V S$) werden als Säulendiagramme und die prozentualen Abweichungen zur Vorsaison ($\Delta V S \%$) als Nadel-diagramme abgebildet. So kann ein Nutzer des Dashboards schon anhand des Diagrammtyps visuell erkennen, ob es sich um Prozentangaben oder um absolute Werte handelt.

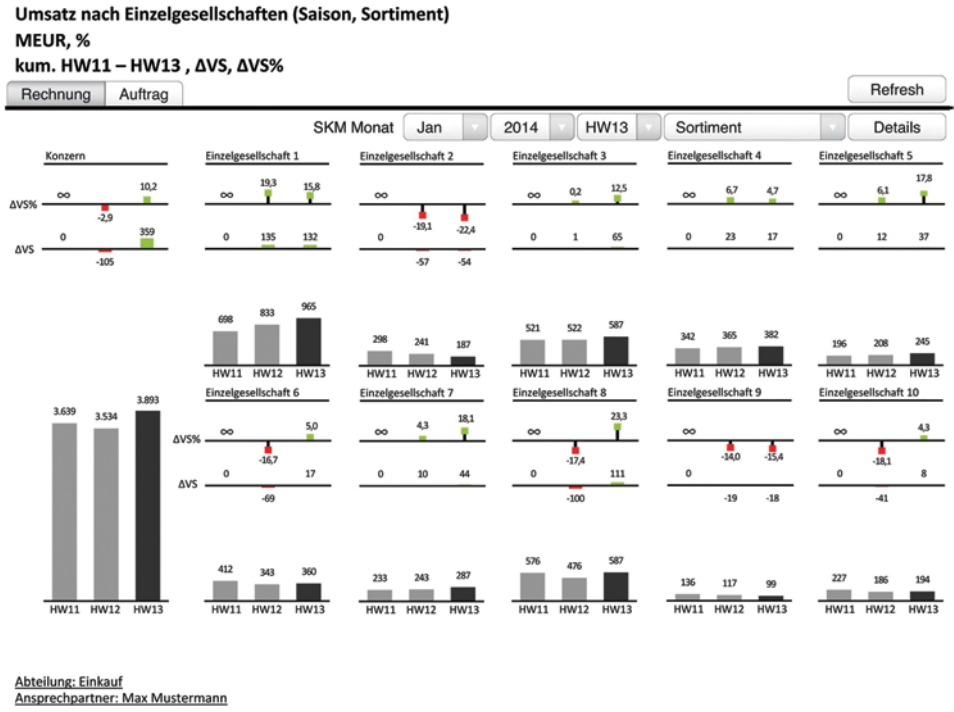


Abb. 15.5 Umsatzdashboard der Otto Group im Konditionenmanagement (Siegel et al. 2014)

Das Umsatzdashboard verwendet eine einheitliche Skalierung. Das heißt, dass alle gezeigten Säulen im gleichen Verhältnis zueinander stehen. Der Mehrwert dieser Vorgehensweise wird dadurch deutlich, dass die Säulen des Konzerns (gesamt) gegenüber den Einzelgesellschaften richtig skaliert sind. Würde man die einzelnen Säulen der Einzelgesellschaften aufeinanderstapeln, so entstünde eine Säule in Höhe derjenigen des Konzerns. Dadurch kann der Nutzer die Relationen sowohl der Einzelgesellschaften untereinander als auch zum (Gesamt-) Konzern sofort visuell nachvollziehen (Abb. 15.5).

Neben der grafischen Aufbereitung wird auch ein einheitliches Titelkonzept verwendet. Die erste Zeile beschreibt den Inhalt des Reports, dahinter steht in Klammern, welche Filtermöglichkeiten dem Nutzer zur Verfügung stehen. In der zweiten Zeile werden die dargestellten Einheiten angegeben. Zudem ist im Notationskonzept die Schreibweise geregelt: Zum Beispiel darf für die Währung Euro nur der Term „EUR“ genutzt werden. Andere Variationen wie „Euro“ oder „€“ sind nicht gestattet. Je nachdem, ob es sich bspw. um Tausend oder Millionen Euro handelt, wird vor „EUR“ ein „T“ bzw. ein „M“ geschrieben. In dem gezeigten Beispiel sind die angegebenen Werte Millionenbeträge, also wird die Währung im Titelkonzept mit „MEUR“ angegeben. In der dritten Zeile des Titelkonzepts wird der abgebildete Zeitraum gezeigt, der sich über die gebotenen Filtermöglichkeiten dynamisch anpasst.

Darüber hinaus bietet das Dashboard verschiedene Filtermöglichkeiten. Zum einen wird zwischen dem beauftragten und dem den Lieferanten tatsächlich in Rechnung gestellt-

Umsatz nach Produktgruppen (Monat, Standort, Abweichung absolut / relativ)
TEUR, %
Mai 2013, ΔVJ



Abb. 15.6 Umsatz-Treemap eines Baumarktes, eigene Darstellung nach [KoPW13]

ten Umsatzvolumen unterschieden. Während die Rechnungssicht damit eine Rückschau ermöglicht, bietet die Auftragssicht eine Vorschau, da sie auch diejenigen Beauftragungen beinhaltet, die noch nicht zur Auslieferung gekommen sind. Weiterhin kann der Nutzer auch die Zeiträume pro Monat und Jahr wählen oder aus verschiedenen Sortimentgruppen (Textil, Schuhe/Accessoires, Hartwaren/Möbel, Elektro) wählen. Die Werte in der Abbildung entsprechen den Werten aller Sortimente. Wählt der Nutzer mittels des Filters ein Sortiment aus, so sieht er den Umsatz pro Einzelgesellschaft und Konzern für drei Saisons mit prozentualen wie auch absoluten Abweichungen in einem Monat nach Auftrags- oder Rechnungssicht für das ausgewählte Sortiment (Siegel et al. 2014).

Neben diesen „klassischen“ Dashboards, die sich an den Regeln von Prof. Dr. Hichert orientieren, wird an dieser Stelle eine weitere Möglichkeit der Visualisierung für Dashboards vorgestellt. Bereits 1990 erfand Ben Sheidermann die Treemap, doch hat sie sich in Unternehmen bis heute noch nicht richtig durchgesetzt. Studien sprechen davon, dass nur etwa 2 % aller Unternehmen in den USA Treemaps nutzen (Eckerson und Hammond 2011). Treemaps eignen sich sehr gut dafür, typische Geschäftsfragen zu beantworten, wie im folgenden Beispiel gezeigt.

Abbildung 15.6 zeigt den gesamten Umsatz eines Baumarktes im Monat Mai 2013. Dieser ist in verschiedene Kategorien (in diesem Beispiel Produktgruppen) wie Garten, Sanitär oder Baustoffe unterteilt. Jede Kategorie ist von einem weißen Rechteck umrahmt,

dessen Fläche den Anteil des Umsatzes am Gesamtumsatz repräsentiert. Es lässt sich beispielsweise ablesen, dass die Kategorie Garten den größten Anteil am Umsatz hat. Jede Kategorie kann sich aus Unterkategorien zusammensetzen (in diesem Beispiel Produktarten), die wiederum als Rechteck dargestellt sind [KoPW13].

Mit Hilfe der Flächen lassen sich die Umsätze der Kategorien oder Unterkategorien gut miteinander vergleichen – und das sogar hierarchieübergreifend. Beispielsweise hat die Unterkategorie Pflanzen den größten Anteil am Umsatz in der Kategorie Garten. Genauso kann der Anwender sofort erkennen, dass die Unterkategorie Pflanzen einen vergleichbaren Umsatz erzielt, wie die Kategorie Holz.

Neben den Umsatzanteilen enthält die Treemap durch ihre Farbgebung eine zweite Informationsdimension. Die Farbgebung unterliegt einem Notationskonzept: Grün signalisiert in unserem Beispiel ein Umsatzwachstum im Vergleich zum Vorjahresmonat, Rot hingegen einen Umsatzrückgang. Die verschiedenen farblichen Abstufungen deuten auf die Höhe der Abweichung hin; je dunkler die Farbe, desto größer die Abweichung. Da es keine Normvorgaben für die Farbgebung gibt, ist es eine Herausforderung, die Schwellenwerte so zu definieren, dass die dargestellten Informationen die richtige Botschaft transportieren. Als sinnvoll erweist es sich, für diese Aufgabe einen Spezialisten des Fachbereiches zurate zu ziehen. Im Beispiel ist zu erkennen, dass sich in der Kategorie Garten nur Dünger erfolgreich entwickelt hat, während sich die anderen Unterkategorien gegenüber dem Vorjahresmonat deutlich schlechter verkauft haben.

Es ist offensichtlich, wie einfach man mithilfe der Treemap eine Winner-Loser-Analyse abbilden kann und auf welche Unterkategorien sich weiterführende Analysen konzentrieren sollten. In unserem Beispiel fokussieren wir uns auf die Kategorie Garten, da diese Kategorie den größten Umsatzanteil im Unternehmen hat und mit dem dunkelrot eingefärbten Segment Pflanzen auch noch jene Unterkategorie beinhaltet, die im gesamten Unternehmen der größte Umsatzträger ist. Für weitergehende Analysen könnte man zum Beispiel eine Drill-Down-Möglichkeit anbieten, die sich ebenfalls nach den Regeln des ID richtet. Abbildung 15.7 zeigt den Drill-Down auf ein Dashboard, das die Unterkategorie Pflanzen für den Monat Mai 2013 zeigt.

Das Dashboard untergliedert die Unterkategorie in ihre einzelnen Produkte und zeigt mit den absoluten Umsätzen sowie den absoluten und prozentualen Abweichungen zum Vorjahresmonat mehr Details als die Treemap. Die Farbnotation verwendet hier Dunkelgrau zur Darstellung des Monatsumsatzes 2013, während Hellgrau für die absoluten Umsätze des Vorjahresmonats verwendet wird. Mit Hilfe der detaillierten Darstellung lässt sich der weitere Entscheidungsprozess operationalisieren. Man kann direkt erkennen, dass nur die Produkte Blumenzwiebeln, Stauden und Kräuter eine positive Umsatzentwicklung zu verzeichnen haben.

Treemaps können somit im Vorfeld einer Analyse einen guten Überblick geben und lassen sich über Drill-Downs gut mit Detailanalysen kombinieren.

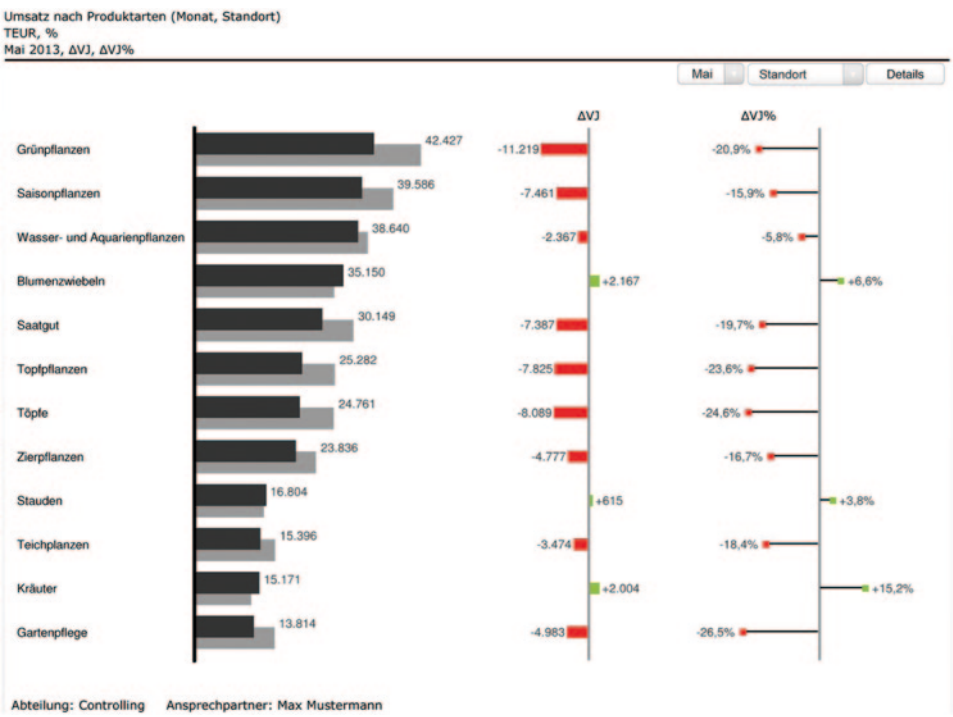


Abb. 15.7 Drill-Down-Ansicht einer Umsatz-Treemap eines Baumarktes, eigene Darstellung nach [KoPW13]

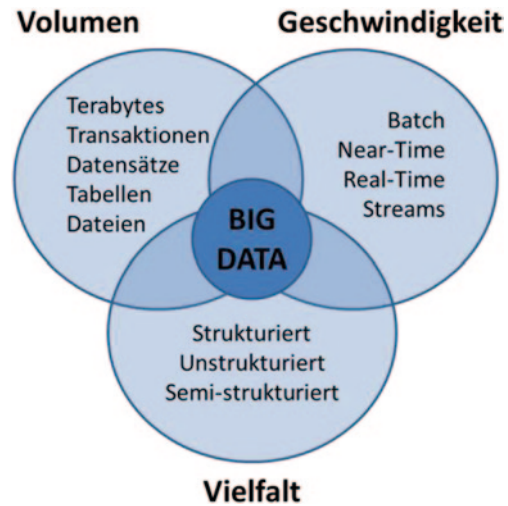
15.5 Visual Analytics und Big Data

Nach den Bereichen ID und VBI wendet sich der Beitrag nun VA zu. VA ist ein Ansatz zur explorativen Datenanalyse und ist ein idealer Weg, um massive Datenmengen (Big Data) auf interaktive Weise zu analysieren. Nach einer Einführung in Big Data und VA wird vor allem auf die Nutzung innerhalb von Unternehmen eingegangen.

15.5.1 Big Data

Big Data ist heute in vielen Unternehmensbereichen ein kritisches Thema. Schätzungen für das weltweite Datenvolumen gehen heute von weit über 2,5 Mrd. Terabyte aus. Viele Jahre schon steigt das Volumen stärker als die Verarbeitungskapazität – bei einem hohen Anteil an unstrukturierten Daten. Die notwendige Vorverarbeitung der Daten verschärft das Problem zudem immens. Auch wenn Herausforderungen mit großen Datenmengen

Abb. 15.8 Die drei Big-Data-Aspekte Volumen, Geschwindigkeit und Vielfalt [KoPW13], angepasst aus (Russom 2011)



aus vielen Branchen kommen, können die Gründe für das Wachstum der Daten auf drei gemeinsame technologische Entwicklungen zurückgeführt werden: eine schnellere Datenerzeugung durch zunehmende Digitalisierung, eine breite Durchdringung effektiver Datenverwaltung in Unternehmen und eine fast grenzenlose Datenkapazität (Kohlhammer et al. 2010).

Big Data ist allerdings nicht auf bestimmte Datenvolumen beschränkt, sondern übertrifft normale Datenmengen an Volumen, Geschwindigkeit und Vielfalt (vgl. Abb. 15.8 nach (Russom 2011)); andere Definitionen ergänzen diese drei Dimensionen noch durch weitere). Ein wesentlicher Aspekt von Big Data ist zum Beispiel die Fähigkeit von Unternehmen, neue Aspekte der Analyse in die Unternehmensprozesse zu integrieren. Davenport und Prusak nennen dies die „Fähigkeit zur analytischen Datenauswertung“ (Davenport und Prusak 1998). Denn Daten an sich sind noch keine relevanten Informationen, sie werden dazu erst durch Zusammenhänge, aufschlussreiche Muster und Erkenntnisse. Diese wiederum müssen in der Menge aller Informationen effektiv zugänglich sein und dargestellt werden, um überhaupt für die Entscheidungsfindung nützlich zu sein. Die Gründe für das Scheitern von Big-Data-Projekten sind häufig nicht in der Hardware- oder EDV-Technik zu suchen, sondern eher darin, dass man sich zu wenig auf die Nutzer der Daten und die Informationsempfänger konzentriert.

Natürlich sind die aktuellen Lösungen für Big Data, die klar definierte Aufgaben auf Basis massiver Datenmengen bearbeiten können, ein großer Fortschritt und eine grundlegende Voraussetzung für Business Analytics. Ohne eine Betrachtung der Nutzer und Entscheidungsträger, die am Ende von Big Data profitieren sollen, gehen die Lösungen jedoch noch nicht weit genug. Verlässt man sich komplett auf automatische Verarbeitung und entsprechende Modelle, geht viel an Kontrolle über Entscheidungen verloren – das hat die letzte Finanzkrise eindrucksvoll gezeigt. Modelle zu erstellen ist ein notwendiger Schritt in der Automatisierung von komplexen Datenanalyseschritten. Eine genaue Ab-

schätzung der Schwächen und Grenzen der Modelle ist für eine sinnvolle Verwendung in Business Analytics jedoch essenziell. Werden Modelle zu einer Black Box, die nur noch wenige Experten verstehen, besteht die Gefahr, dass sie im Unternehmen auf unreflektierte Weise eingesetzt werden.

Der Nutzen von Big Data Analytics in datenintensiven Unternehmen ist dabei bereits erwiesen. Neue Analysenmöglichkeiten haben dort zu einer Veränderung der Geschäftsprozesse geführt. Die Prozesse und Entscheidungen stehen dabei bereits im Mittelpunkt der Big-Data-Technologien. So können Firmen wie eBay, Amazon oder Netflix ihre Prozesse vollständig an ihren Geschäftsdaten ausrichten und dadurch extrem schnell auf neue Kundeninteressen und Marktveränderungen reagieren. Im Jahr 2012 nannte – laut BITKOM-Leitfaden (Urbanski und Weber 2012) – allerdings erst eine relative Minderheit als wichtigsten Parameter bei Big-Data-Projekten den Nutzen für Entscheidungsträger. Das Thema VA setzt genau an der Stelle an, an der die Investitionen in Big Data zu einer immensen Datenverarbeitungskapazität geführt haben, aber noch nicht zu gewinnbringenden neuen Möglichkeiten beim Entscheider.

15.5.2 Visual Analytics

Direkt anknüpfend an die Einführung in Big Data, wird von folgender Definition der VA ausgegangen [KoPW13]:

Visual Analytics kombiniert automatische Analysetechniken mit interaktiven Visualisierungen für ein schnelles Verstehen, Schlussfolgern und Entscheiden auf Basis extrem großer, dynamischer und heterogener Datenmengen (Big Data).

VA greift auf zwei verschiedene Herangehensweisen zurück: Auf der einen Seite werden statistische Methoden und automatische Algorithmen verwendet, also die Verarbeitungskapazität heutiger IT genutzt. Auf der anderen Seite nutzt VA die interaktive Visualisierung, um menschliches Wissen und Erfahrung in einen neuen Prozess einzubringen. Dieser Prozess verbindet Mensch und Maschine – im Dienste einer besseren Arbeitsteilung bei der Nutzung von Big Data. Diese Arbeitsteilung hat Vorteile gegenüber einer separaten Lösung mit automatischen Methoden oder mit Visualisierungen. Die Modelle und Algorithmen können nur auf Wissen basieren, das bei der Implementierung bereits vorhanden war. Verändert sich die Situation auf dynamische und unvorhersehbare Weise, greift das Modell nicht mehr; so sind Kreditkartenbetrüger den Betrugsfiltern beispielsweise immer einen Schritt voraus. Visualisierungen wiederum haben ein Skalierbarkeitsproblem, d. h. sie funktionieren nur mit einer gewissen Menge an Daten und Attributen (oder Dimensionen). Erst die Kombination mit automatischer Vorverarbeitung erhöht diese Grenze. Betrachtet man diese Arbeitsteilung etwas genauer und sehen wir uns in Abb. 15.9 das Referenzmodell des VA-Prozesses an, wie es von (Keim et al. 2010) dargestellt wird.

Die Schritte von den Rohdaten hin zum Wissen werden durch farbige Ovale dargestellt, die Übergänge zwischen den Schritten durch Pfeile. Die beiden getrennten Wege – die

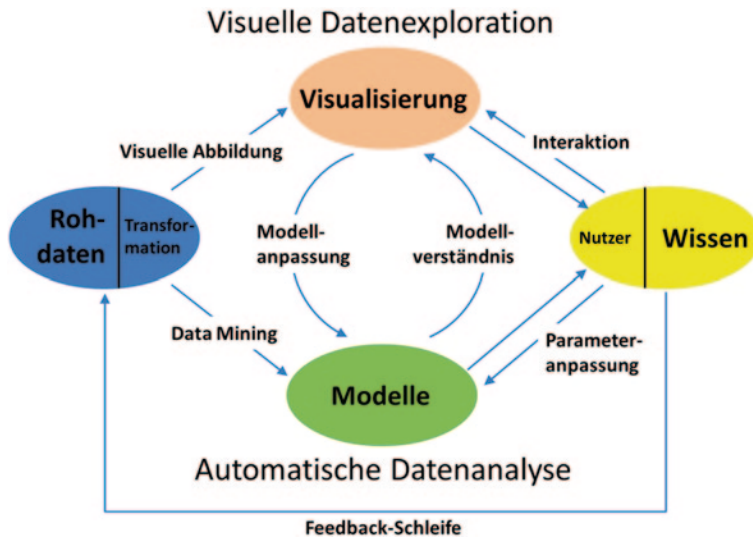


Abb. 15.9 Der VA-Prozess verbindet zwei Wege zwischen Daten und Wissen: visuelle Datenexploration und automatische Datenanalyse (Kohlhammer et al. 2013b), angepasst aus (Keim et al. 2010)

visuelle Datenexploration und die automatische Datenanalyse – werden dabei durch die Verbindungen zwischen Visualisierung und Modellen ergänzt. Wie bereits in Abb. 15.1 gezeigt, ist VA explorativ. Ein Data Scientist kann dabei in verschiedenen Schritten einen der beiden Wege gehen oder eine Kombination aus beiden Wegen verwenden. Nach Anwendung einer automatischen Methode kann der Data Scientist sich die Ergebnisse oder Modellparameter visualisieren lassen (Modellverständnis) und sehen, wie er sein Modell adaptieren muss (Modellanpassung). Oder er lässt sich die Daten erst einmal aggregiert anzeigen, um so erste Muster in den Daten zu entdecken und die geeignete automatische Methode auszuwählen. Die neuesten VA-Ansätze verbinden beide Wege in einer interaktiven Technik. So bildet die Benutzerinteraktion mit Zwischenschritten die Eingabe für die automatische Methode, und zwar während diese ausgeführt wird.

15.5.3 Aktueller Einsatz in Unternehmen

VA unterstützt vor allem zwei Ziele in der Datenanalyse: zum einen die Detektion bereits erwarteter Aspekte, die mithilfe von VA (also auch mittels Nutzung automatischer Verfahren) schneller und leichter zu entdecken sind, und zum anderen die Exploration von Daten und das Entdecken neuer Zusammenhänge. Gerade diese Exploration ist eine Aufgabe, in der der Mensch eine zentrale Rolle spielt. Beide Ziele werden auch in Unternehmen verfolgt, wenn Rohdaten nach unternehmensrelevanten Informationen durchsucht oder schlichtweg auf ihre Relevanz untersucht werden sollen. Statistische Methoden und Datenvorverarbeitung sind dabei unabdingbare Hilfsmittel, die aber ohne die richtigen

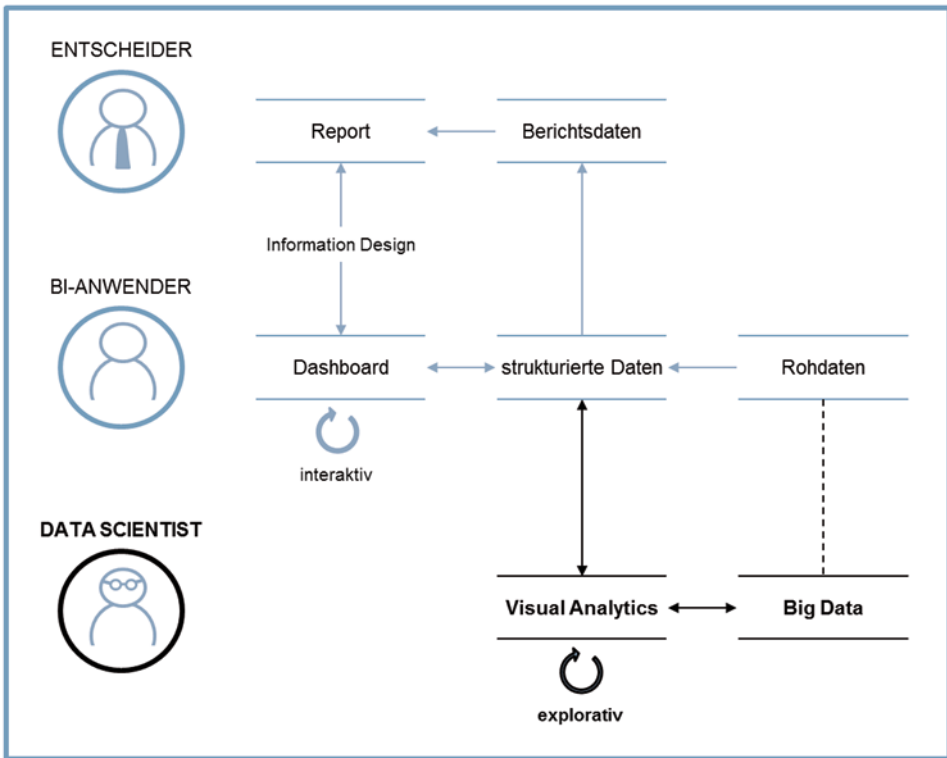


Abb. 15.10 Der Data Scientist arbeitet mit massiven Datenmengen und kann VA-Techniken für die explorative Analyse einsetzen [KoPW13]

Nutzerschnittstellen und interaktiven Visualisierungen meist nur ineffektiv zu nutzen sind. Auch deshalb sucht man heute händeringend nach den bereits erwähnten Data Scientists, denen der Umgang mit Hadoop oder R keine Probleme bereitet. Gleichzeitig sollen diese Analysen aber auf Markt- und Kundenerfahrungen basieren, was voraussetzt, dass verschiedene Rollen im Unternehmen eng zusammenarbeiten.

Abbildung 15.10 zeigt die Rolle des Data Scientist im Rahmen des VBA-Modells. Nach einer Vorverarbeitung der Rohdaten stehen dem Data Scientist massive, potenziell dynamische Datenmengen zur Verfügung. Die VA-Technologien in diesem Modell greifen also auf Big Data wie auf einen Cache zu, der beispielsweise mittels Hadoop auf Rohdaten basiert. Wie am Anfang dieses Beitrags ausgeführt, können die Daten dabei in mehreren verteilten Datenquellen und in verschiedenen Formaten vorliegen. Der Data Scientist arbeitet dabei nicht isoliert. Vielmehr werden Analyseergebnisse, die längerfristig relevant sind, in die strukturierten Daten des Unternehmens überführt. Diese stehen so für BI-Anwender und VBI-Tools zur Verfügung.

Diese Vorgehensweise ist in heutigen Unternehmen noch nicht üblich, da die Visualisierung in allen Ebenen der Entscheidungsfindung keinen definierten Stellenwert hat. Will man VBA im Unternehmen einführen, bewegt man sich sinnvollerweise entlang

der Kapitel von [KPW], um über die Einführung eines Notationskonzeptes und die Gestaltung der interaktiven Visualisierungen und Dashboards hin zu einer Visualisierungsunterstützung für Data Scientists zu gelangen. Dies schafft auch Vertrauen in die Visualisierungen der Daten, und ein solches Vertrauen ist nötig, um Entscheidungen darauf zu stützen. Der interessierte Leser sei hier auf [CCS12] verwiesen, die sich insbesondere den Herausforderungen für die Unternehmenskultur bei der Einführung neuer VA-Techniken widmen.

Dieselben Autoren identifizieren aber auch einige ganz generelle Herausforderungen bei der Einführung von VA-Technologien, auf die wir hier kurz eingehen wollen. Sicherlich ergeben sich bei der Einführung jedweder Software ganz grundsätzlich bestimmte Probleme oder Aspekte, die zu beachten sind.

Chinchor et al. haben sechs VA-spezifische Herangehensweisen – vier technologische und zwei kulturelle – identifiziert, die ein Unternehmen bei der Einführung von VA-Technologien unterstützen können:

- VA-Werkzeuge akribisch testen und evaluieren, um kritische Mitarbeiter zu überzeugen
- Frühzeitig Datensicherheit und Schutz personenbezogener Daten ansprechen (unternehmensintern wie -extern)
- Interoperabilität der VA-Technologien mit existierenden Datenbeständen und Analysesoftware unterstützen
- VA-Werkzeuge in existierende Arbeitsprozesse integrieren bzw. Übergang zu neuen VA-unterstützten Arbeitsprozessen schrittweise vollziehen
- Mitarbeiter, die durch neue VA-Ansätze am meisten Zeit einsparen oder anderweitig profitieren, identifizieren, involvieren und überzeugen
- Analysten rollenspezifisch schulen (BI-Anwender vs. Data Scientist)

Für Datenanalyseprobleme, die in vielen Unternehmen vorkommen, finden sich in aktuellen BI-Werkzeugen bereits einige VA-Technologien. So bieten beispielsweise SAS, TIBCO und Tableau (um nur drei zu nennen) eine entsprechende Software, die relativ schnell auf den Unternehmensdaten eingesetzt werden kann. Die sechs oben genannten Herangehensweisen finden jedoch auch bei der Einführung kommerzieller Lösungen ihre sinnvolle Anwendung. Für spezielle Analyseherausforderungen sind auch spezielle Visualisierungen gefragt, die gerade bei den interaktiven und visuellen Möglichkeiten über die Standard-Software hinausgehen. Im Folgenden werden wir noch auf ein solches Beispiel eingehen, in dem eine sogenannte Designstudie zum Einsatz kommt; dabei werden zu einer bestimmten Aufgabe eines Analysten und zu einer oder mehreren Datenmengen adäquate Visualisierungs- oder VA-Technologien identifiziert. In vielen Fällen werden hierbei auch gänzlich neue Visualisierungsmethoden entwickelt.

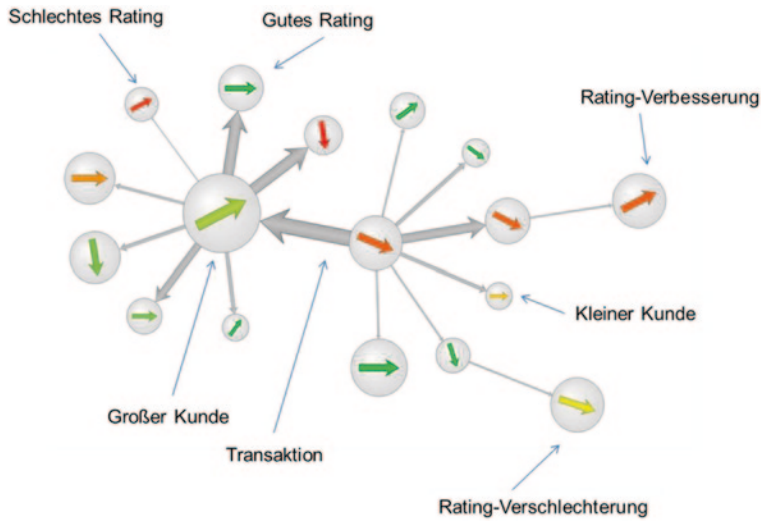


Abb. 15.11 Die Darstellung eines Risikoportfolios über ein visuell-interaktives Netzwerk [KoPW13]

15.5.4 Anwendungsbeispiel

Das gewählte Beispiel stammt aus dem Bereich Risikomanagement. Dabei werden über ein Risikomodell in die Vorausberechnung eines Risikos möglichst viele relevante Daten und Informationen einbezogen. Die Nachvollziehbarkeit der Qualität des Modells und der Vorausberechnung sind dabei selbst für Experten nicht ganz einfach nachvollziehbar oder kommunizierbar, und dies verhindert natürlich auch ein einfaches Verständnis auf Entscheider-Ebene. Für eine deutsche Großbank wurden daher Möglichkeiten gesucht, die Nachvollziehbarkeit und die Kommunikation des Risikomodells zu verbessern. Gleichzeitig sollten aber noch weitere Daten in die Modellberechnung mit einfließen. Die Grundidee war, die isolierte Betrachtung einzelner Kunden durch eine breitere Betrachtung von Kundennetzwerken zu ergänzen. Dabei war neben den Transaktionen untereinander auch die Zugehörigkeit zu gleichen Regionen und Branchen interessant.

Abbildung 15.11 zeigt die damals gewählte Visualisierung eines Netzwerks, die interaktiv verändert und erweitert werden kann. Die einzelnen Kunden sind dabei als Kreise dargestellt, deren Radius proportional zur Größe der Firmen ist, und die Pfeile stehen für individuelle Ratings der Kunden. Dickere Verbindungskanten zeigen ein höheres Transaktionsvolumen zwischen den Kunden an. Der visuelle Eindruck eines solchen Risikoportfolios geht über einzelne Betrachtungen bestimmter Kunden hinaus und zeigt nicht nur das direkte Umgebungsnetzwerk, sondern vermittelt auch einen Gesamteindruck des Kundenstamms. Auch sogenannte Cluster von Kunden, in denen mehrere Kunden von einem Kunden abhängig zu sein scheinen, werden hier sofort ersichtlich.

15.6 Zusammenfassung

Die Digitalisierung in der Geschäftswelt führt mit ihren exponentiell wachsenden Datenmengen dazu, dass die Ansprüche der Adressatengruppen in Hinblick auf die Verarbeitung und die effiziente Aufbereitung entscheidungsrelevanter Daten zunehmen. Wichtige Informationen in Rohdaten können mithilfe von VA identifiziert und mittels VBI und ID effizient in die richtigen Entscheidungen überführt werden. Die entscheidungsorientierte Verwendung der Daten und die Bedürfnisse des Nutzers werden bei zukünftigen Investitionen in BI-Systeme und bei der Überarbeitung von Entscheidungsprozessen ins Zentrum der Überlegungen rücken.

Viele Unternehmen haben die Bedeutung der sinnvollen und adressatengerechten Anwendung von Visualisierungen bereits erkannt. Die Herausforderung besteht jedoch darin, die relevanten Daten verständlich und intuitiv zu visualisieren und Berichten die notwendige Objektivität zu verleihen, damit Entscheider in ihren Aufgaben bestmöglich unterstützt werden.

Im vorgestellten Beispiel aus dem Konditionenmanagement der Otto Group können Anwender das Dashboard ihren Bedürfnissen gemäß filtern und interaktiv zwischen zwei Ansichten (Vor- und Rückschau) wechseln. Dank einheitlicher Skalierung, einem durchdachten Einsatz von Farben und dem konsequenten Verzicht auf Dekorationen sind die Kriterien der objektiven Darstellung erfüllt. Die daneben vorgestellte Treemap geht über das klassische Dashboard hinaus – sie verwendet ein mehrdimensionales Konzept, das im Vorfeld tiefgreifender Analysen zum Einsatz kommen kann. So lassen sich problematische Geschäftsfelder effizient identifizieren.

Beide Beispiele verdeutlichen, wo VBA in einem Unternehmen Analyse-, Entscheidungs- und Kontrollaufgaben optimieren kann. Sie zeigen auch, wie eng verzahnt die drei im VBA-Modell vorgestellten Nutzerrollen (Entscheider, BI-Anwender, Data Scientist) im Unternehmen zusammenarbeiten müssen. Damit diese drei Gruppen nicht isoliert arbeiten, empfiehlt sich die Einführung eines Notationskonzepts, das für alle Adressatengruppen gilt. Es erhöht die Akzeptanz für VBA, schafft Vertrauen in die Aussagefähigkeit von Daten und erlaubt damit einen funktions- und hierarchieübergreifenden Einsatz von Visualisierungen [KoPW13].

In Zukunft wird die Nachfrage nach einheitlichen Normen für die Geschäftskommunikation auf internationaler Ebene steigen. Die von Prof. Dr. Hichert veröffentlichten International Business Accounting Standards (IBCS) sind aus diesem Grund bereits komplett auf Englisch verfügbar. Alle Stakeholder müssen sich mit dieser Entwicklung auseinandersetzen; dies gilt insbesondere für die Softwarehersteller, die auf das veränderte Anforderungsprofil reagieren müssen und passende Werkzeuge für ID, VBI und VA bereitstellen sollten. Und tatsächlich bieten einige Hersteller bereits Softwarelösungen für visuelles Arbeiten an, die die drei Nutzerrollen des VBA-Modells unterstützen.

Angewandte und Grundlagenforschung führen zu neuen Erkenntnissen über das Arbeiten mit großen Datenmengen. Gleichzeitig benötigen die Unternehmen praxisrelevantes Wissen über die richtige Anwendung von VBA-Methoden. Das alles

führt für Softwarehersteller, Unternehmen und Dienstleister zu einem dynamischen und innovativen Marktumfeld. Doch eines ist klar: Diejenigen Unternehmen, die diese Entwicklungen verfolgen und VBA-Methoden einsetzen, werden im Informationszeitalter die Nase vorn haben.

Literatur

- Beyer, M.A., Laney, D.: The Importance of Big Data: A Definition. Gartner, USA (2012)
- Bosbach, G., Korff, J.J.: Lügen mit Zahlen. Heyne, München (2011)
- Chinchor, N., Cook, K.A., Scholtz, J.: Building adoption of visual analytics software. In: Dill, J., Earnshaw, R., Kasik, D., Vince, J., Wong, P.C. (Hrsg.) *Expanding the Frontiers of Visual Analytics and Visualization*, S. 509–529. Springer, London (2012)
- Davenport, T.H., Prusak, L.: *Working Knowledge*. Harvard Business Press, Boston (1998)
- Eckerson, W., Hammond, M.: *Visual Reporting and Analysis – Seeing is Knowing*. TDWI, Renton (2011)
- Keim, D.A., Kohlhammer, J., Ellis, G., Mansmann, F. (Hrsg.): *Mastering the Information Age – Solving Problems with Visual Analytics*. Eurographics Association, Goslar (2010)
- Kohlhammer, J., May, T., Davey, J., Ruppert, T.: Visual Analytics – Verbindung von Analyseverfahren und Visualisierungstechniken. *IM – Die. Fachz. Inf. Manage. Consult.* **25**, 10–17 (2010)
- Kohlhammer, J., Proff, D.U., Stahl, K-U., Wiener, A.: *Visual Business Analytics Studie 2012*. BLUEFORTE, Hamburg (2012)
- Kohlhammer, J., Proff, D.U., Stahl, K-U., Wiener, A.: *Visual Business Analytics Studie 2013*. BLUEFORTE, Hamburg (2013a)
- Kohlhammer, J., Proff, D.U., Wiener, A.: *Visual Business Analytics – Effektiver Zugang zu Daten und Informationen*. dpunkt Verlag GmbH, Heidelberg (2013b)
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H.: *Big Data: the Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, New York City (2011)
- Proff, D., Stahl, K-U., Wiener, A.: *Die TDWI Konferenz München – Eine Erfolgsgeschichte seit 10 Jahren*. TDWI Marketing- und Medienleistungen, Troisdorf (2013)
- Russom, P.: *Big Data Analytics*. TDWI, Renton (2011)
- Siegel, U., Wittmaack, C., Schröder, G., Wiener, A.: *Anwenderbericht: Konditionenmanagement in der Unternehmenssteuerung, Visualisierung bei der Otto Group*. In: Gluchowski, P. (Hrsg.) *BI Spektrum*, Ausgabe 02/2014, S. 22–25. Fokus Visual BI, Troisdorf (2014)
- Urbanski, J., Weber, M.: *Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte*. BITKOM, Berlin (2012)

Markus Linden und Frank Navrade

Inhaltsverzeichnis

16.1 Einführung 326

16.2 Evolutionsstufen in der Unternehmenssteuerung 327

 16.2.1 Finanzorientierte Steuerungskonzepte 327

 16.2.2 Ausgewogene Steuerungskonzepte 329

 16.2.3 Gegenüberstellung aktueller Steuerungskonzepte 331

16.3 Erfolgskritische Komponenten eines Geschäftsmodells 333

 16.3.1 Abdeckungsgrad der Geschäftsmodellkomponenten 333

 16.3.2 Steuerungsdomänen eines Geschäftsmodells 336

 16.3.3 Zuordnung steuerungsrelevanter Kennzahlen 338

16.4 Planung und Steuerung mit analytischen Informationen 340

 16.4.1 Herausforderung in der Informationsbedarfsanalyse 340

 16.4.2 Identifikation von Kennzahlen-Dimensionen-Kombinationen 341

 16.4.3 Informationssystem zur Steuerung des Geschäftsmodells 345

16.5 Fazit und Ausblick 347

Literatur 348

M. Linden (✉)
SAP Deutschland SE & Co. KG, Ratingen, Deutschland
E-Mail: Markus.Linden@uni-due.de

F. Navrade
PwC cundus AG, Duisburg, Deutschland
E-Mail: frank.navrade@cundus.de

Zusammenfassung

Ein Unternehmen lebt und stirbt mit seinem Geschäftsmodell. Es ist die Existenzberechtigung eines Unternehmens, die beschreibt, wie die innerorganisationale Wertschöpfung generiert wird. Vor diesem Hintergrund sollte das Management danach bestrebt sein, das Unternehmen anhand des zugrunde liegenden Geschäftsmodells zu planen, zu steuern und zu kontrollieren. In den letzten Jahrzehnten sind jedoch Steuerungskonzepte entwickelt und eingesetzt worden, die lediglich die finanzielle Perspektive eines Unternehmens fokussieren. Da die Finanzkennzahlen eines Unternehmens jedoch immer nur das Ergebnis und nicht die Ursache des wirtschaftlichen Handelns abbilden, ergänzten Kaplan und Norton in ihrem Konzept der Balanced Scorecard nichtfinanzielle Perspektiven. Dadurch sollte eine ausgewogene Steuerung des Unternehmens ermöglicht werden, die jedoch zentrale Steuerungsparameter nicht berücksichtigt. In ihrem Ursprung integriert die BSC lediglich die prozessualen Aspekte eines Unternehmens, sodass wesentliche erfolgskritische Faktoren vernachlässigt werden. Genau diese zentralen Erfolgsfaktoren werden jedoch durch die Komponenten eines Geschäftsmodells widerspiegelt, weshalb eine danach ausgerichtete Planung und Steuerung des Unternehmens heutzutage eine logische Konsequenz darstellen sollte.

16.1 Einführung

In den vergangenen Jahren ist die IT-basierte Unternehmenssteuerung immer stärker in den Fokus des Managements gerückt (Silvi et al. 2010, S. 3). Gründe hierfür sind rückläufige Leistungskennzahlen aber auch externe Impulse, wie die vergangene Wirtschafts- und Finanzkrise. Um einen genaueren Einblick in das Unternehmen zu gewinnen und es erfolgreich steuern zu können, reicht es allerdings nicht, lediglich die Möglichkeiten aktueller Business-Intelligence(BI)-Technologien auszureizen. Vielmehr gilt es, ein belastbares fachliches Fundament zu legen, das im Zusammenwirken mit modernen IT-Werkzeugen einen echten Mehrwert in der Planung und Steuerung generiert.

Im Kern wird dieses fachliche Fundament durch die Zusammenführung der strategischen und operativen Unternehmensebenen dargestellt (Gluchowski et al. 2009, S. 8). Als Bindeglied können dabei sogenannte Geschäftsmodellkomponenten dienen, die mithin den Steuerungscharakter gängiger Konzepte schärfen (Chamoni und Linden 2011, S. 279; Silvi et al. 2010, S. 3). Bislang werden zur Überbrückung der hierarchischen Strukturen Konzepte, wie die Balanced Scorecard (BSC), eingesetzt (Kaplan und Norton 1997, S. 2). Mit Blick auf ihre Praxisrelevanz wird jedoch deutlich, dass sie sich nicht flächendeckend gegen die vorherrschenden finanziellen Steuerungskonzepte durchsetzen konnte. Dies liegt mitunter an politischen Gründen, aber auch daran, dass nicht alle erfolgskritischen Aspekte reflektiert werden (Bange et al. 2008, S. 10). An dieser Schwachstelle knüpften Möller, Drees und Schläpke an, indem sie das im Entrepreneurship verankerte Konzept des

Geschäftsmodellmanagements erstmals auf steuerungsrelevante Fragestellungen übertragen (Möller et al. 2011, S. 213; Silvi et al. 2010, S. 13).

Als konzeptionelle Grundlage für diese Fragestellungen kann das Rahmenwerk von Osterwalder und Pigneur dienen, da sie in ihrem wissenschaftlichen Diskurs *die* erfolgskritischen Kernkomponenten für Gründungsunternehmen identifizierten und deren Zusammenwirken veranschaulichten (Osterwalder und Pigneur 2010, S. 15). Die dabei entdeckten Komponenten stellen für das hier fokussierte Konzept den Ausgangspunkt dar. Angelehnt daran, wird im weiteren Verlauf des Beitrags angenommen, dass die identifizierten Kernkomponenten eines Gründungsunternehmens ebenfalls bei etablierten Unternehmen als erfolgskritisch anzusehen sind. Zwar existieren zahlreiche Unterschiede zwischen diesen beiden Organisationsformen, wie beispielsweise hinsichtlich ihrer Struktur, ihrer prozessualen Abläufe sowie ihrer grundsätzlichen Komplexität. Und dennoch basieren beide Konstrukte auf einem inhärenten Geschäftsmodell. Dies gilt auch, sofern das zugrundeliegende Geschäftsmodell nicht formal ausgearbeitet ist.

Mit Blick auf die Leistungsmessung im Unternehmen und der dabei erforderlichen IT-Unterstützung sind diese zentralen Komponenten eines Geschäftsmodells aufzugreifen und in einem Steuerungskonzept zu manifestieren. Sofern diese Bestrebungen erfolgreich in der Unternehmenspraxis eingesetzt werden, kann zukünftig eine effektivere Planung und Steuerung auf Basis des Geschäftsmodells erfolgen. Der wesentliche Nutzen eines dafür entwickelten Informationssystems stellt die umfassende Betrachtungsweise der Faktoren dar, die auf die Unternehmensperformance einwirken. Dabei stehen neben den erfolgskritischen Komponenten vor allem Ursache-/Wirkungszusammenhänge im Mittelpunkt, die als Resultat des wirtschaftlichen Handelns in dem finanziellen Ergebnis des Unternehmens münden. Bevor dieser geschäftsmodellorientierte Lösungsansatz näher beschrieben wird, sind im folgenden Abschnitt zunächst der Ursprung sowie die Weiterentwicklungen einschlägiger Konzepte der Unternehmenssteuerung dargestellt.

16.2 Evolutionsstufen in der Unternehmenssteuerung

16.2.1 Finanzorientierte Steuerungskonzepte

Seit Beginn des wissenschaftlichen Diskurses in der Unternehmenssteuerung bilden die finanzorientierten Steuerungskonzepte den Ausgangspunkt späterer Weiterentwicklungen. Finanzorientierte Ansätze konzentrieren sich im Wesentlichen auf die monetären Indikatoren einer Wirtschaftsorganisation. Vor diesem Hintergrund werden die rein finanziellen Größen, beispielsweise in Form von Ergebnisrechnungen, in den Vordergrund steuernder Aktivitäten gestellt. Die Finanzergebnisse werden dabei meist in unterschiedlichen Kennzahlensystemen gruppiert und konsolidiert, sodass die sogenannten Top-Level-Größen als Zielgrößen geplant werden können (Gladen 2008, S. 236). Als eine der bekanntesten Top-Level-Kennzahlen gilt der Return on Investment (ROI), der sich aus der Kapitalumschlagshäufigkeit und der Umsatzrentabilität errechnen lässt. Das Kennzahlensystem,

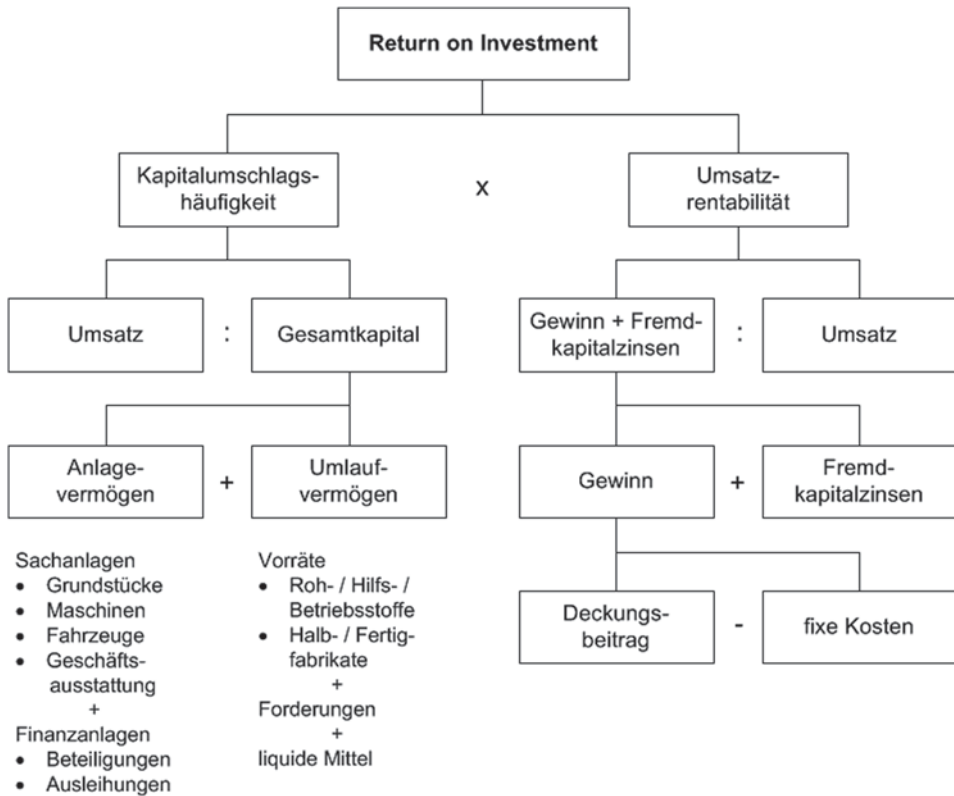


Abb. 16.1 DuPont-Schema. (Preißner 2010, S. 221)

das die Transformation solch hierarchisch angepasster Größen abbildet, wird als Du-Pont-Schema bezeichnet (Küpper 2008, S. 397). Anhand der hierbei definierten Berechnungsvorschriften sollen Gesamtzusammenhänge aufgedeckt werden, die anhand der Betrachtung isolierter Kennzahlen nicht möglich wären (Preißler 2008, S. 17). In der nachfolgenden Abb. 16.1 werden die hierarchischen Strukturen und die Zusammenhänge zwischen den einzelnen Größen verdeutlicht.

Mit Blick auf Abb. 16.1 wird deutlich, dass das DuPont-Schema im Wesentlichen auf eine partielle Analyse von korrespondierenden Kennzahlen abzielt, die über mehrere Verdichtungsebenen in einem sachlogischen Zusammenhang stehen. Durch ein solch strukturiertes Rechensystem wird eine Transparenz hergestellt, die Aufschluss über die operativen Treiber einer übergeordneten Spitzenkennzahl gibt. Zudem kann das DuPont-Schema durch die einheitliche Definition monetärer Größen als unternehmensübergreifender Standard eingesetzt werden. Doch trotz dieser positiven Charakteristika sieht sich das DuPont-Schema mit kritischen Stimmen konfrontiert. Dabei werden hauptsächlich die stark isolierte Betrachtung lediglich einer Spitzenkennzahl sowie die ausschließliche Konzentration auf finanzielle Steuerungsparameter als Schwachstellen angesehen. Vor diesem Hintergrund besitzt das DuPont-Schema eine geringe Relevanz in der Unterneh-

menssteuerung, auch wenn der ROI als eine zentrale Kennzahl in den Ergebnisberichten von Unternehmen regelmäßig zu finden ist (Grüning 2002, S. 64). Neben dem DuPont-Schema existieren weitere finanzorientierte Steuerungskonzepte, wie beispielsweise der Shareholder-Value-Ansatz, der Geschäftswerttreiberbaum etc., auf die an dieser Stelle wegen ihrer vergleichbaren Struktur nicht näher eingegangen wird (Otley 2007, S. 14).

Aufgrund der dargelegten Schwächen von rein monetär ausgerichteten Steuerungskonzepten erfolgte in den 1990er Jahren eine Weiterentwicklung in diesem Bereich. Ihr bekanntester Vertreter, die BSC, wird im folgenden Abschnitt anhand ihrer Steuerungsperspektiven näher beschrieben.

16.2.2 Ausgewogene Steuerungskonzepte

Die BSC hat gegenüber traditionellen Systemen durch die Integration von nicht finanziellen Steuerungsperspektiven rasch einen hohen Bekanntheitsgrad erlangt (Reichmann 2011, S. 73). Neben der Tatsache ein durchgängiges Managementkonzept konzipiert zu haben, sind Kaplan und Norton anfangs auch für die Praxistauglichkeit der BSC gewürdigt worden (Jossé 2005, S. 3). Im Vordergrund dieser Erfolgsgeschichte steht dabei zum einen die Ergänzung der Finanzperspektive um sogenannte weiche Steuerungsparameter, die in die Perspektiven *Innovationen*, *Prozesse* und *Kunden* eingeteilt worden sind (Kaplan und Norton 1996, S. 7). Zum anderen wird durch die korrespondierenden Strategy Maps eine Durchgängigkeit von der Strategie bis zur Umsetzung erzeugt, die Ursache-/Wirkungszusammenhänge innerhalb der Wertschöpfungsarchitektur eines Unternehmens abzubilden vermag (Jossé 2005, S. 3). Die damit geschaffene Transparenz ist zu jener Zeit ein Novum gewesen, die begleitet durch weitere ausgewogene Steuerungskonzepte, wie beispielsweise die Performance Pyramid und das EFQM-Modell, eine neue Evolutionsstufe in der Unternehmenssteuerung begründete (Lynch und Cross 1995; Jossé 2005, S. 287).

Auch wenn die BSC innovative Aspekte beinhaltet, so liegt ihr Ursprung dennoch in der traditionellen Prozesslehre. Der Charakter der ergänzten Perspektiven geht nämlich auf den klassischen Dreiklang *Input – Throughput – Output* zurück. Dabei stellt die Perspektive *Innovation* den Ausgangspunkt dar, der durch die innerorganisationalen Inputfaktoren Lernen und Entwickeln widergespiegelt wird. In diesem Kontext spielt neben der Entwicklung neuer Produkte und Dienstleistungen insbesondere die Förderung der personellen Ressourcen eine entscheidende Rolle für ein zukunftsfähiges Unternehmen. Die zur Wertschöpfung erforderlichen betrieblichen Abläufe werden ferner in der Perspektive *Prozesse* detailliert beschrieben, sodass eine konkrete Operationalisierung der strategischen Zielvorgaben in der BSC vorgenommen werden kann. Darüber hinaus reflektiert die Perspektive *Kunde* die in der Prozesslehre geforderte Orientierung an den Bedarfen der Empfänger eines Prozessergebnisses. Das Resultat der vorangegangenen BSC-Perspektiven drückt sich schließlich in den Finanzen eines Unternehmens aus, sodass dieser Steuerungsaspekt ebenfalls seinen berechtigten Platz findet (Kaplan und Norton 1996, S. 9).

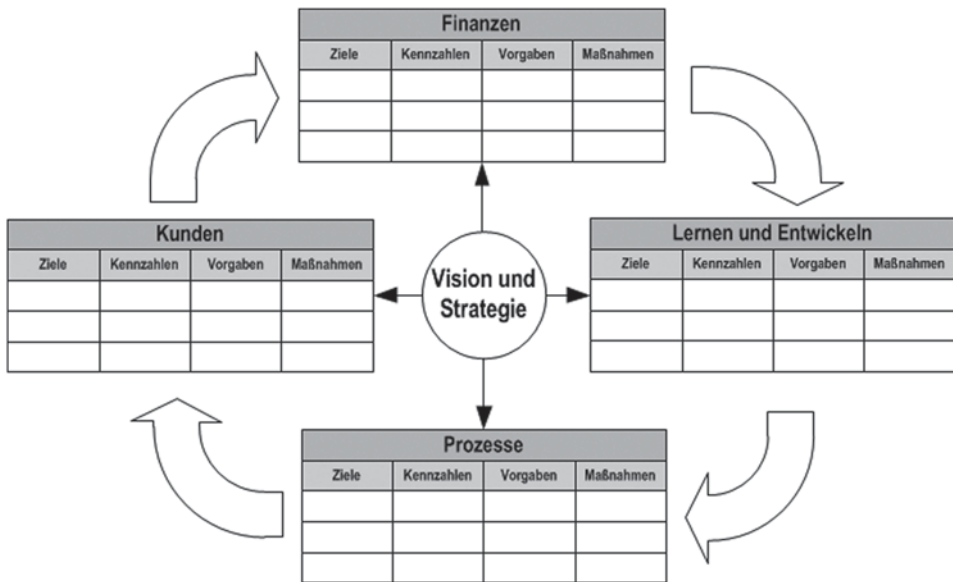


Abb. 16.2 Perspektiven der Balanced Scorecard. (Kaplan und Norton 1996, S. 9)

Durch das dabei erzeugte Zusammenspiel unterschiedlicher Parameter können die Werttreiber und deren Auswirkungen gesteuert werden (Kaplan und Norton 1996, S. 25). Vor diesem Hintergrund ermöglicht die BSC eine Kaskadierung der Steuerungskennzahlen, sodass notwendige Zielerreichungskontrollen bis in die operativen Abläufe eines Unternehmens transparent gemacht werden (Gladen 2008, S. 178).

Mit Blick auf die beschriebenen Vorteile gegenüber rein finanzorientierten Konzepten wird der Anspruch der BSC deutlich, eine ganzheitliche Unternehmenssteuerung zu etablieren (Jossé 2005, S. 18). In Abb. 16.2 werden die dafür relevanten Steuerungsperspektiven illustriert und in den darauffolgenden Ausführungen beschrieben.

Die Perspektive *Innovation* wird im vorangestellten Beispiel durch sogenannte Lern- und Entwicklungsziele repräsentiert, die anhand entsprechender Kennzahlen, Vorgaben und Maßnahmen definiert und im Rahmen der Steuerung kontrolliert werden. Der notwendige Bedarf eines Unternehmens beinhaltet diesbezüglich im Wesentlichen den Ausbau personeller und materieller Kapazitäten. Insbesondere die Entwicklung der Mitarbeiter soll als eine der Kernressourcen neben einem innovativen Technologieeinsatz die prozessualen Abläufe effizienter gestalten, um einen Wettbewerbsvorteil generieren zu können. Damit die Bedeutung der Prozessstrukturen auch innerhalb der Unternehmenssteuerung zur Geltung kommt, werden in der Perspektive *Prozesse* die bestehenden Abläufe an den übergeordneten Zielen ausgerichtet und entlang der Unternehmenshierarchie durchgängig miteinander verknüpft. Als Beispiel können hierfür die prozessualen Ziele in Form von Kosten, Qualität und Zeit dienen, die sich auf die Parameter der Kundenperspektive auswirken. Darüber hinaus werden neue Prozesse innerhalb der Scorecard definiert und entsprechend ihres Granularitätsgrades in einen Gesamtzusammenhang

gebracht. Die Steuerungsfaktoren der Perspektive *Kunde* werden ferner anhand der Kundenzufriedenheit und -profitabilität und weiterer repräsentativer Kennzahlen dargestellt, sodass eine Indikation hinsichtlich aktueller und zukünftiger Zielmärkte formuliert werden kann. In diesem Zusammenhang können auch klassische Kundenanforderungen abgebildet werden, die sich beispielsweise durch die Kennzahl *Lieferzeit* abbilden lassen. Diese und weitere Kennzahlen besitzen einen maßgeblichen Einfluss auf die Perspektive *Finanzen*, indem beispielsweise nicht wettbewerbsfähige Lieferzeiten durch ein entsprechendes Kaufverhalten abgestraft werden. Auf diese Weise werden zum einen zentrale Ursache-/ Wirkungsmechanismen zwischen den Perspektiven der BSC und zum anderen die unterschiedlichen Aggregationsebenen entlang der Unternehmenshierarchie abgebildet (Kaplan und Norton 1996, S. 47).

Aus den vorherigen Ausführungen ist festzuhalten, dass die BSC über mehrere Leistungsebenen hinweg eine kontinuierliche Zielumsetzung forciert. Dabei nutzt sie als ein ausgewogenes Steuerungskonzept sowohl monetäre als auch nichtmonetäre Indikatoren, die über die beschriebenen Perspektiven und Kennzahlen in einen sachlogischen Zusammenhang gestellt werden. Der verfolgte Gestaltungsansatz ist dabei grundsätzlich individuell an die jeweilige Organisation anzupassen. Aufgrund ihrer einfachen Struktur und der Berücksichtigung prozessualer Leistungstreiber eines Unternehmens hat sich die BSC schnell verbreitet und besitzt als Steuerungssystem eine hohe Praxisrelevanz (Grüning 2002, S. 64). Neben diesen Argumenten liegt die starke Verbreitung der BSC in Wissenschaft und Praxis jedoch maßgeblich darin begründet, dass bis dato die traditionellen Steuerungskonzepte lediglich die finanziellen Kennzahlen mit einem kurzfristigen Horizont berücksichtigten (Kaplan und Norton 2001, S. 22).

Trotz ihres anfänglichen Erfolgs wird die BSC in der aktuellen Unternehmenslandschaft meist nur rudimentär eingesetzt, sodass sie ihre volle Wirkungskraft als strategisches Steuerungswerkzeug bislang nicht entfalten konnte. Ein Grund könnte hierfür die einseitige Fokussierung von prozessualen Input-, Throughput- und Outputfaktoren sein, die die erfolgskritischen Komponenten in der Unternehmenssteuerung vernachlässigen. Diese und weitere Schwachstellen etablierter Steuerungskonzepte werden im folgenden Abschnitt näher erläutert.

16.2.3 Gegenüberstellung aktueller Steuerungskonzepte

Die bereits angedeuteten Schwachstellen aktueller Steuerungskonzepte können anhand verschiedener Kriterien strukturiert gegenübergestellt werden. Im Zuge eines solchen Vergleichs wird deutlich, dass die BSC das bislang ausgereifteste Konzept darstellt. Dennoch beinhaltet sie lediglich in Hinblick auf die prozessuale Organisation einen ausgewogenen Steuerungscharakter. Weitere relevante Kernelemente eines Unternehmens, die über Erfolg und Misserfolg entscheiden, fehlen hingegen in ihrer Betrachtung. In Abb. 16.3 wird dieser Aspekt bei der Gegenüberstellung zwischen dem DuPont-Schema und der BSC durch das Kriterium der Ausgewogenheit vergegenwärtigt.

Ausprägung	DuPont-System	Shareholder Value	EFQM	Balanced Scorecard
Aufgabe	Partielle Analyse	Bewertung	Ganzheitliche Bewertung	Kontinuierliche Zielumsetzung
Hierarchieebenen	Verdichtungsebenen	Verdichtungsebenen	Organisations- und Leistungsebenen	Organisations- und Leistungsebenen
Kennzahlentyp	Monetär	Monetär	Monetär und nichtmonetär	Monetär und nichtmonetär
Kennzahlenbeziehung	Definitionslogisch (Rechensystem)	Definitionslogisch (Rechensystem)	Sachlogisch (Ordnungssystem)	Sachlogisch (Ordnungssystem)
Gestaltungsansatz	Standardisiert	Standardisiert	Standardisiert	Organisations-spezifisch
Ausgewogenheit	Konzentration auf monetäre Größen	Konzentration auf monetäre Größen	Ausgewogene Sichten	Ausgewogene Kennzahlen und Perspektiven
Klassifizierung	Analysesystem	Bewertungssystem	Bewertungssystem	Steuerungssystem
Praxisrelevanz	Gering	Hoch	Hoch	Hoch




Abb. 16.3 Vergleich aktueller Steuerungskonzepte. (Grüning 2002, S. 64)

Die hieraus resultierende Fragestellung, ob und inwiefern die vier BSC-Perspektiven zur ausgewogenen Steuerung eines Unternehmens genügen, stellen Kaplan und Norton in einer ihrer kritischen Würdigungen selbst. Ergebnis dieser Reflektion ist, dass die externen Einflussfaktoren in der BSC nicht ausreichend berücksichtigt werden. In der Tat finden die schon zu Beginn der BSC-Ära etablierten wettbewerbsstrategischen Konzepte keine angemessene Beachtung innerhalb der überwiegend intern fokussierten Perspektiven (Baum et al. 2007, S. 370). Vor diesem Hintergrund plädieren Kaplan und Norton in einer ihrer Veröffentlichungen für eine nachrangige Anpassung oder gar eine Erweiterung der internen BSC-Perspektiven, sofern die externen Faktoren maßgeblich auf sie einwirken. Als Beispiel für solche als relevant geltende und damit zu berücksichtigende externe Einflüsse können beispielsweise vorherrschende Branchenspezifika dienen (Kaplan und Norton 1996, S. 34).

Auf einen weiteren zentralen Aspekt für eine ausgewogene Steuerung ist seitens Kaplan und Norton jedoch nicht konkret eingegangen worden. Hierbei steht die Fragestellung im Vordergrund, ob und inwiefern ein Unternehmen überhaupt anhand der vier betrachteten Perspektiven ganzheitlich gesteuert werden kann. Allein der zuvor angedeutete prozessuale Ursprungsgedanke der BSC weist bereits darauf hin, dass durch sie keine ganzheitliche Planung, Steuerung und Kontrolle ermöglicht wird. Der entscheidende Grund hierfür ist, dass ein Unternehmen mehr steuerungsrelevante Faktoren besitzt als jene, die in der Prozessorganisation verankert sind. Zur Identifikation genau dieser bislang vernachlässigten erfolgskritischen Faktoren ist das zugrundeliegende Geschäftsmodell eines Unternehmens zu betrachten. Grundsätzlich expliziert ein Geschäftsmodell die zentralen, inhärenten Wirkungsmechanismen eines Unternehmens, die im Ergebnis durch die Rentabilität der Geschäftsaktivitäten zum Ausdruck gebracht werden. Im Zuge dessen beschreibt es, wie die einzelnen Kernelemente eines Unternehmens ein Nutzen- bzw.

Wertangebot generieren, das an den Bedarfen der potenziellen Nachfrage ausgerichtet ist und an die entsprechenden Zielgruppen übermittelt wird. Die nachfolgende Definition verdeutlicht die zentralen Aspekte des Zusammenwirkens essentieller Geschäftsmodellkomponenten (Osterwalder und Pigneur 2010, S. 14):

A business model describes the rationale of an organization creates, delivers, and captures value.

Basierend auf diesem Definitionsansatz kann konstatiert werden, dass jedes Unternehmen über ein Geschäftsmodell verfügt, selbst dann, wenn sich das Management nicht über die Existenz der einzelnen Geschäftsmodellelemente bewusst ist. Das Geschäftsmodell kann dabei als Bindeglied zwischen der Unternehmensstrategie und den Geschäftsprozessen fungieren, sodass eine Konkretisierung der meist generisch geprägten Strategien sowie eine Explikation essentieller Handlungsfelder ermöglicht werden. Andererseits können durch das Aufdecken von erfolgskritischen Faktoren und der damit einhergehenden Transparenz wesentliche Stellschrauben in der Unternehmenssteuerung identifiziert werden (Möller et al. 2011, S. 222).

Dieser deutlich breitere Wirkungsradius von Geschäftsmodellen bescheinigt einmal mehr, dass im Vergleich die BSC keine umfassende Unternehmenssteuerung ermöglicht. Auch die Ergänzung der Finanzperspektive um qualitative Perspektiven ändert an dieser Tatsache nichts, da auch weiterhin zentrale Steuerungselemente vernachlässigt werden. Durch den gedanklichen Transfer der zentralen Komponenten eines Geschäftsmodells auf das Controlling werden genau diese Elemente sichtbar. Welche dieser erfolgskritischen Elemente ein Geschäftsmodell zur Unternehmenssteuerung beinhaltet, wird im nachfolgenden Abschnitt thematisiert.

16.3 Erfolgskritische Komponenten eines Geschäftsmodells

16.3.1 Abdeckungsgrad der Geschäftsmodellkomponenten

Ein Geschäftsmodell kann grundsätzlich in die Komponenten Value Proposition, Customer Segments, Channels, Customer Relationships, Key Resources, Key Activities, Key Partners, Cost Structure und Revenue Streams unterteilt werden (Osterwalder und Pigneur 2010, S. 18). Zentraler Baustein der sogenannten Business Model Canvas, die in Abb. 16.4 dargestellt wird, ist die Value Proposition. Hierbei handelt es sich um das Wertangebot eines Unternehmens, das sich an den Kundenbedürfnissen ausrichtet und konkrete Problemlösungen in Form von Produkten oder Dienstleistungen aufzeigt. In den Customer Segments werden die forcierten Zielgruppen abgebildet, für deren Bedürfnisbefriedigung die zuvor beschriebenen Wertangebote generiert werden. Dabei kann es sich um eine oder mehrere Zielgruppen handeln, je nachdem, wie umfangreich das Produkt- oder Dienstleistungsportfolio ausgestaltet ist. Die Channels des Geschäftsmodells geben Auskunft darüber, welche Kanäle das erzeugte Wertangebot an die jeweilige Zielgruppe übertragen.

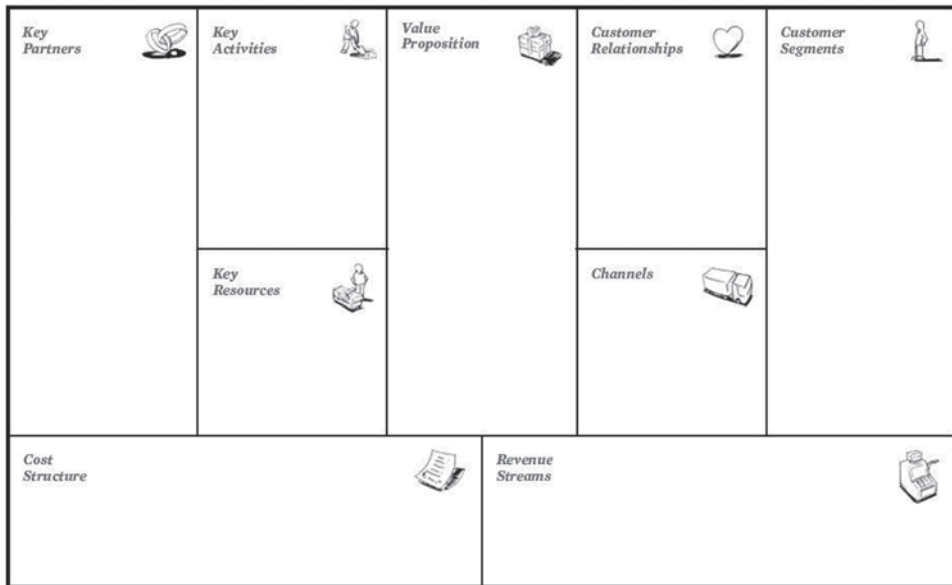


Abb. 16.4 Erfolgskritische Komponenten eines Geschäftsmodells. (Osterwalder und Pigneur 2010, S. 18)

Die hier fokussierten Kanäle müssen nicht zwingend physischer Natur sein, wie es bei Standorten der Fall wäre. Insbesondere bei innovativen Geschäftsmodellen werden neben den klassischen Distributionskanälen auch Kommunikationsmodelle eingesetzt, die auf sozialen Medien und Netzwerken basieren. In einem direkten Zusammenhang stehen dabei die Customer Relationships als weiterer Baustein des Geschäftsmodells. In diesem Aktionsfeld wird festgelegt, mit welchen Instrumenten die Kundenbeziehung über den gesamten Lebenszyklus initiiert, gepflegt und möglicherweise beendet wird. Hierbei gilt es, zielgruppenspezifische Maßnahmen zu identifizieren und umzusetzen. Die Key Resources zeigen ferner die wesentlichen Kapazitäten eines Unternehmens auf, die notwendig für die Erstellung des Wertangebots sind. Beispiele für solche Ressourcen unterscheiden sich stark je nach Ausrichtung des Geschäftsmodells. Grundsätzlich lassen sich jedoch mit finanziellen, (im-)materiellen und personellen Ressourcen drei wesentliche Arten voneinander abgrenzen. Die Key Resources und die Key Activities bedingen sich in ihrer Ausgestaltung gegenseitig. Der Grund hierfür liegt in dem Einsatz der zuvor definierten Ressourcen zur Ausführung von betrieblichen Aktivitäten, die wertschöpfenden oder auch unterstützenden Charakter annehmen können. Unabhängig von ihrem direkten oder indirekten Wertschöpfungsbeitrag ist, dass es sich um zentrale Aktivitäten zur Erstellung und Vermarktung des Wertangebots handelt und nicht zwangsläufig die klassischen Funktionsbereiche abgebildet werden. In dem Baustein der Key Partners werden die Partnerstrukturen des Unternehmens festgelegt. Dabei werden von typischen Lieferantenstrukturen über strategische Allianzen bis hin zu konkreten Outsourcingobjekten alle

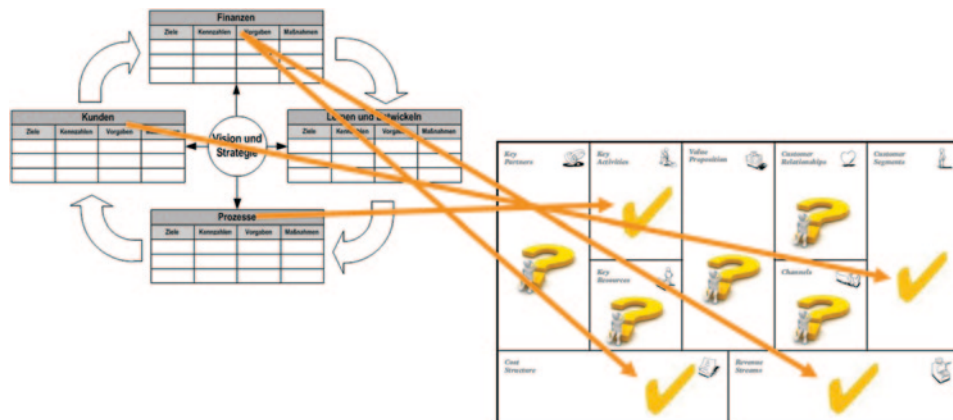


Abb. 16.5 Abdeckungsgrad erfolgskritischer Komponenten

relevanten Netzwerke dargestellt, die die Tragfähigkeit des Geschäftsmodells unterstützen (Osterwalder und Pigneur 2010, S. 17).

Das Ergebnis dieser Geschäftsmodellkomponenten spiegelt sich abschließend in der Cost Structure und den Revenue Streams wider. Innerhalb der Cost Structure werden die wesentlichen Kostenverursacher beschrieben, die durch den Ressourceneinsatz und die Prozessabläufe innerhalb des Geschäftsmodells erzeugt werden. Die Revenue Streams zeigen hingegen die Einnahmequellen auf, die aus den Transaktionen von Wertangeboten und den jeweiligen Kundensegmenten resultieren. Diese neun Bausteine der Business Model Canvas stellen die Kernkomponenten dar, die die wesentlichen Elemente eines Unternehmens beschreiben. Jede einzelne dieser Komponenten illustriert dabei einen erfolgskritischen Faktor für Unternehmen, den es mithin in der Steuerung zu berücksichtigen gilt (Osterwalder und Pigneur 2010, S. 17 ff.). Diese Annahme basiert auf der Relevanz von Geschäftsmodellen als Gradmesser für unternehmerischen Erfolg. Mit anderen Worten wird sich ein Unternehmen nur dann langfristig am Markt durchsetzen können, sofern das Zusammenspiel der inhärenten Komponenten eines Geschäftsmodells reibungslos funktioniert. Wenn sich das Geschäftsmodell hingegen aufgrund der mangelhaften Ausgestaltung der Komponenten nicht zu etablieren vermag, wird das Unternehmen über kurz oder lang nicht überlebensfähig sein. Demnach stellt ein funktionierendes Geschäftsmodell die zentrale Voraussetzung für die Existenz eines Unternehmens dar, die durch die innerorganisationale Wertschöpfung sowie deren Kosten- und Erlösstruktur abgebildet wird.

Mit Blick auf die zuvor formulierte These, dass die Komponenten eines Geschäftsmodells als Steuerungsfaktoren zu nutzen sind, werden die in der BSC fehlenden Parameter transparent. Abbildung 16.5 verdeutlicht den Abdeckungsgrad zwischen den BSC-Perspektiven und den Geschäftsmodellkomponenten.

Die im vorherigen Abschnitt beschriebenen Perspektiven *Innovation*, *Prozesse*, *Kunde* und *Finanzen* der BSC decken erkennbar nicht alle erfolgskritischen und somit steuerungsrelevanten Faktoren eines Unternehmens ab. In Hinblick auf die

Geschäftsmodellkomponenten wird deutlich, dass die Finanzperspektive in Form der Cost Structure bzw. Revenue Streams, die Kundenperspektive durch die Customer Segments und die Prozessperspektive anhand der Key Activities widergespiegelt werden. Allerdings legt ein solcher Abgleich auch weitere steuerungsrelevante Faktoren offen, die im Rahmen der BSC nicht berücksichtigt werden. Dabei handelt es sich um die Komponenten *Value Proposition*, *Channels*, *Customer Relationships*, *Key Resources* und *Key Partners*. Durch den erweiterten Abdeckungsgrad erfolgskritischer Faktoren zeigt sich die Relevanz des Geschäftsmodells für die Unternehmenssteuerung. Wie die hierzu erforderlichen Steuerungsdomänen innerhalb des Geschäftsmodells ausgestaltet werden können, wird im nachfolgenden Abschnitt beschrieben.

16.3.2 Steuerungsdomänen eines Geschäftsmodells

Grundsätzlich kann zur Ermittlung der einzelnen Geschäftsmodellelemente ein generisches Vorgehen empfohlen werden. Bei einem solchen Ansatz erfolgt zunächst ein Abgleich der neun Bausteine mit den ökonomischen und kulturellen Aspekten ebenso wie mit der vorherrschenden Wettbewerbssituation des Unternehmens. Dadurch werden erste Indizien ermittelt, die eine Grundkonstellation der Organisation offenlegen. Als Beispiel für eine solche Grundkonstellation können die kostengetriebene sowie die wertgetriebene Ausrichtung des Unternehmens in ihrem Markt und Wettbewerbsumfeld dienen (Osterwalder und Pigneur 2010, S. 16). Auch im Hinblick auf die Unternehmenssteuerung werden hierdurch bereits erste wichtige Erkenntnisse erlangt, da sich der Umfang relevanter Kennzahlen an solchen Grundkonstellationen ausrichtet. Außer den Steuerungsaspekten kann durch die Betrachtung der ökonomischen Rahmenbedingungen das Wertangebot mit der Bedürfnisstruktur der Kundensegmente in Einklang gebracht werden.

Gemäß der stark kostengetriebenen Logistikbranche werden in Abb. 16.6 als Value Propositions exemplarisch die Verlässlichkeit sowie die Kosteneffizienz der Transporte in den Vordergrund gestellt. Zudem spielt in der hier betrachteten Lebensmittellogistik ebenso die Sicherheit der Kühlkette eine wichtige Rolle im Rahmen des Wertangebots. Vor diesem Hintergrund setzen sich auch die Kundengruppen aus Lebensmittelherstellern zusammen, die je nach Wertangebots- und Zielgruppendefinition des Unternehmens weiter aufgegliedert werden können. Die verschiedenen Kanäle und Beziehungen zum Kunden sind im vorliegenden Fall vornehmlich durch individuelle Netzwerke sowie Vertragsgestaltungen geprägt. Zur Erstellung des Wertangebots eines Logistikdienstleisters sind ferner entsprechende Ressourcen und Aktivitäten erforderlich, die sich vor allem durch eine länderübergreifende IT und eine darauf basierende Kapazitätsplanung auszeichnen. Gerade in einer wettbewerbsintensiven und kostengetriebenen Industrie können diese Elemente einen entscheidenden Konkurrenzvorteil herbeiführen. Sofern die betrachteten Elemente demnach extern effizienter bezogen werden können, sollten Überlegungen angestellt werden, ob bestimmte Ressourcen und Aktivitäten an ein Partnernetzwerk auszulagern sind. Insbesondere IT-Infrastrukturthemen stellen hierfür typische Beispiele



Abb. 16.6 Ausgewählte Steuerungsdomänen der Geschäftsmodellkomponenten

dar, die anhand der Elemente *Rechenzentrum* bzw. *Softwarehaus* in Abb. 16.6 aufgezeigt werden. Solche Auslagerungsbestrebungen unternehmensinterner IT auf externe Partner wirken sich entsprechend positiv auf die Kostenstruktur des Logistikers aus. Hingegen schwierig beeinflussbare Kostentreiber sind vor allem die Energiekosten, die auf die Sicherstellung der Kühlkette sowie die Gebäude- und Lagerinfrastruktur zurückzuführen sind. Bei den Einnahmequellen können schließlich neben den klassischen Transport- und Logistikdienstleistungen auch sogenannte Value Added Services einen Beitrag zur Rentabilität des Unternehmens leisten.

Die zuvor beschriebenen Elemente der Geschäftsmodellkomponenten eines Lebensmittellogistikers werden im Folgenden als Steuerungsdomänen bezeichnet. Diese begriffliche Abgrenzung ist dem gedanklichen Transfer der Business Model Canvas geschuldet, die grundsätzlich zur Identifikation strategischer Stoßrichtungen entwickelt worden ist und nun auf den Bereich der Unternehmenssteuerung übertragen wird. Das Rahmenwerk für die betrachteten Steuerungsdomänen wird durch die neun Bausteine des Geschäftsmodells dargestellt. Der Kerngedanke dieser Überlegung ist die bereits aufgezeigte Relevanz des Geschäftsmodells für die wirtschaftliche Tragfähigkeit eines Unternehmens. Mithin ist das Zusammenwirken der einzelnen Geschäftsmodellkomponenten ein wesentlicher Indikator für die aktuellen und zukünftigen Geschäftserfolge von Organisationen. Auf Basis dieser Fakten kann gefolgert werden, dass nicht nur für die Gestaltung strategischer Überlegungen, sondern auch zu Planungs- und Steuerungszwecken innerhalb deren Umsetzung die Bausteine eines Geschäftsmodells herangezogen werden sollten. Die neun Bausteine der Business Model Canvas geben hierfür einen konsistenten Rahmen vor, der durch die Konkretisierung erfolgsrelevanter Faktoren die wesentlichen Steuerungsdomänen offenlegt.

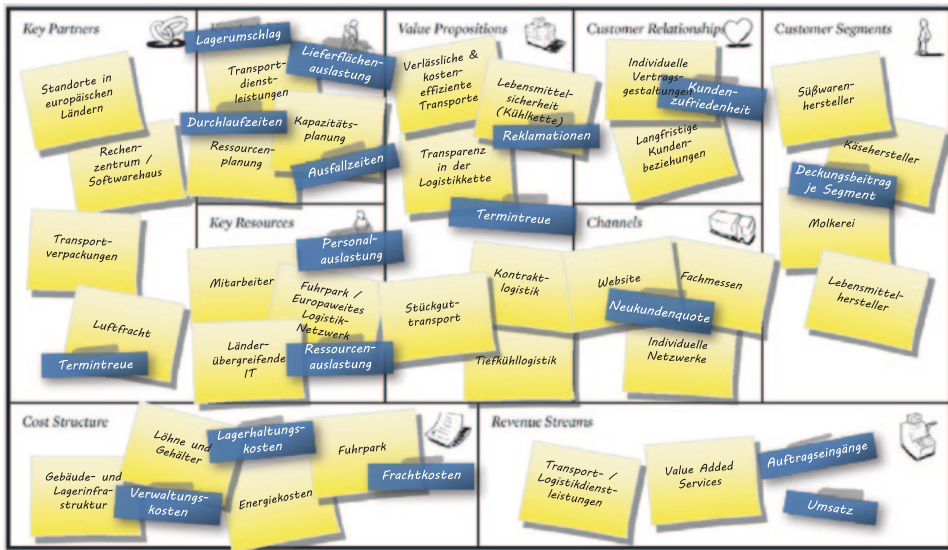


Abb. 16.7 Identifikation steuerungsrelevanter Kennzahlen

Zur Erhebung der vorangestellten Steuerungsdomänen sind grundsätzlich drei Schritte erforderlich. In einem ersten Schritt werden dabei für alle Geschäftsmodellkomponenten zentrale Steuerungsfragen eines bestimmten Anwendungsfalls identifiziert. Im zweiten Schritt werden die formulierten Steuerungsfragen in die zentralen Domänen übertragen, die es im Rahmen einer effektiven Unternehmenssteuerung zu betrachten gilt. In einem dritten Schritt wird abschließend der Wirkungsgrad der identifizierten Steuerungsdomänen auf den Erfolg des Unternehmens beurteilt, sodass lediglich die entscheidenden Elemente im betrachteten Rahmenwerk verbleiben. Mit der Identifikation und Konsolidierung der relevanten Steuerungsdomänen ist das konzeptionelle Vorgehen allerdings noch nicht abgeschlossen. Damit im weiteren Verlauf ein Entwurf eines groben Steuerungskonzepts entwickelt werden kann, müssen für die einzelnen Domänen zudem passende Kennzahlen identifiziert werden.

16.3.3 Zuordnung steuerungsrelevanter Kennzahlen

Die Identifikation und Zuordnung steuerungsrelevanter Kennzahlen erfolgt grundsätzlich entlang der bereits ermittelten Steuerungsdomänen. Die dabei zugrundeliegende Leitfrage orientiert sich an dem Bedarf an analytischen bzw. dispositiven Informationen, die für eine effektive Planung und Steuerung des Geschäftsmodells notwendig sind. Abbildung 16.7 zeigt eine Übersicht ausgewählter Kennzahlen und deren Zuordnung zu den identifizierten Steuerungsdomänen eines Logistikdienstleisters.

Angelehnt an das Vorgehen zur Ermittlung der Steuerungsdomänen wird bei der Identifikation relevanter Kennzahlen die Steuerungsfrage erneut aufgegriffen. Durch die Beantwortung der domänenspezifischen Steuerungsfragen ergeben sich unweigerlich Kennzahlen, die den Kriterien objektiver Informationsbedarfe genügen. Im vorliegenden Beispiel des Logistikdienstleisters können im Rahmen der Value Propositions Kennzahlen, wie beispielsweise Reklamationen und Termintreue, steuerungsrelevante Informationen in Hinblick auf die Lebensmittelsicherheit und die Transparenz in der Logistikkette darstellen. Auf Seiten der Customer Segments können unter anderem die Deckungsbeiträge je Segment wertvolle Hinweise über die wirtschaftliche Kapazität einzelner Kundengruppen geben. Die Ergebnisse solcher Analysen können ferner in einer Einordnung münden, die eine Verschiebung der Schlüsselkunden zur Folge hat. Insbesondere durch die Gegebenheiten in der Logistikbranche, in der komplexe Sendungsstrukturen vorherrschen, können Informationen zur Profitabilität von einzelnen Kundengruppen oder gar einzelner Kunden einen entscheidenden Wettbewerbsvorteil darstellen. Des Weiteren kann im Zuge der Channels des betrachteten Geschäftsmodells die ermittelte Neukundenquote dazu dienen, die aktuellen Vertriebskanäle zu klassifizieren und in erforderlichen Fällen Umschichtungsmaßnahmen zu effektiveren Kanälen einzuleiten. Die Customer Relationships können anhand der Kundenzufriedenheit bewertet werden, die gerade bei den hier fokussierten langfristigen Kundenbeziehungen wertvolle Hinweise in Hinblick auf die zukünftige Konditionenpolitik sowie den damit einhergehenden Vertragsausgestaltung geben kann. Im Rahmen der Key Resources können als steuerungsrelevante Kennzahlen vor allem Informationen angeführt werden, die auf die Auslastung abstellen. Vor dem Hintergrund der hier identifizierten Mitarbeiter und Fuhrparks, die als essentielle Ressourcen des Logistikdienstleister gelten, stellen die Personalauslastung sowie die Ressourcenauslastung des Fuhrparks typische Kennzahlen zur Steuerung des vorliegenden Geschäftsmodells dar. Diese auslastungsgetriebene Kennzahlen spiegeln sich auch in den Key Activities wider, indem die Transportdienstleistungen neben weiteren Kennzahlen anhand der Lieferflächenauslastung beurteilt werden. Außerdem spielen in diesem Zusammenhang der Lagerumschlag, die Durchlauf- sowie Ausfallzeiten eine wichtige Rolle in der Planung und Steuerung. Zur Bewertung der Lieferanten im Rahmen der Key Partners kann unter anderem die Termintreue ein zentrales Auswahlkriterium darstellen, das im Zuge einer ganzheitlichen Unternehmenssteuerung durchaus Berücksichtigung finden sollte. Die typischen Kostenblöcke eines Lebensmittellogistikers werden durch die Fracht-, Lagerhaltungs-, und Verwaltungskosten innerhalb der Cost Structure des Geschäftsmodells abgebildet. Auf Seiten der korrespondierenden Revenue Streams gehören die klassischen Kennzahlen, wie beispielsweise der Umsatz, zu den zentralen finanziellen Steuerungsfaktoren. Ferner können zu Planungszwecken die Auftragseingänge erste Indikatoren für die zukünftige Geschäftsentwicklung sein, die entsprechend in das Steuerungskonzept zu integrieren sind.

Nachdem potenzielle Steuerungsdomänen sowie eine Auswahl relevanter Kennzahlen identifiziert worden sind, gilt es im Weiteren die analytischen Anforderungen an eine Kennzahl herauszuarbeiten. Zur Abdeckung solcher Anforderungen werden Business-

Intelligence-Systeme eingesetzt, deren Inhalte auf verschiedenen Ansätzen zur Informationsbedarfsanalyse basieren. Welche Herausforderungen bei der Abbildung multidimensionaler Analyseanforderungen entstehen können, wird im nachfolgenden Kapitel veranschaulicht.

16.4 Planung und Steuerung mit analytischen Informationen

16.4.1 Herausforderung in der Informationsbedarfsanalyse

In der Planung und Steuerung von Unternehmen wird grundsätzlich auf interne und externe Informationen zurückgegriffen, deren Qualität einen maßgeblichen Einfluss auf die Entscheidungen des Managements besitzt. Als technologisches Fundament für solche steuerungsrelevanten Informationen werden Business-Intelligence-Systeme eingesetzt, um die zugrunde gelegten Daten aus unterschiedlichen Quellsystemen konsistent zusammenzuführen. Auf Basis dieser Daten können ad hoc Analysen durchgeführt und Berichte generiert werden. Die Informationen, die sich aus diesen Analysen und Berichten für das Management ergeben, dienen der Erfüllung des Informationsbedarfes in Entscheidungssituationen. Sowohl beim initialen Aufbau als auch beim Redesign eines Business-Intelligence-Systems steht die Ermittlung des Informationsbedarfs im Mittelpunkt der Anforderungsanalysephase (Jung 2006, S. 60). Ziel ist es, ein möglichst präzises Bild der steuerungsrelevanten Informationen sowie der verfügbaren Quelldaten zu erhalten. In der Praxis stellt diese Phase regelmäßig eine große Herausforderung dar, weil unter anderem der objektive Informationsbedarf innovativer Unternehmenssteuerungskonzepte sowie der subjektive Informationsbedarf des Managements mit der Datenverfügbarkeit abgeglichen werden müssen (Koreimann 1976, S. 65; Strauch 2002, S. 70; Frie und Strauch 2001, S. 241 ff.). Weitere typische Herausforderungen, die sich bei einer solchen Anforderungsanalyse ergeben, werden im Folgenden näher erläutert.

Wie bereits zu Beginn dieses Beitrags aufgezeigt worden ist, sind Steuerungs-, Controlling- oder Kennzahlenkonzepte häufig sehr generisch gestaltet und gehen dementsprechend nicht bis auf das letzte, spezifische Detail ein. Dennoch gewährleisten diese Konzepte und die damit verbundene Ermittlung des objektiven Informationsbedarfs in der Regel die unternehmensweite strategische Abstimmung zwischen allen Unternehmensbereichen. Hier besteht die Aufgabe, das Design dieser Konzepte intentionskonform zu verfeinern und die Besonderheiten aller Unternehmensbereiche zu berücksichtigen. Als weitere Herausforderung ist die Artikulationsfähigkeit und der damit einhergehende Spezifikationsgrad anzumerken. Trotz seines breiten Erfahrungsschatzes in der Unternehmenssteuerung ist das Management häufig nicht in der Lage, seinen subjektiven Informationsbedarf präzise zu formulieren (Meyer und Strauch 2000, S. 98; Konzelmann 2008, S. 274). Zusätzlich erhöhen multidimensional auswertbare Kennzahlen die Komplexität des zu ermittelnden Informationsbedarfs und erschweren damit dessen Spezifikation. Zudem tragen alternative Hierarchien zur Strukturierung der Informationen sowie nicht-

triviale Berechnungsformeln der Kennzahlen zu dieser steigenden Komplexität bei. Darüber hinaus besteht zwischen der Wertart *Ist* und den Wertarten *Plan* und *Forecast* ein nicht zu vernachlässigender Unterschied in der verfügbaren und geforderten Granularität. Diese Umstände erweisen sich als hinderlich für eine schnelle und präzise Spezifikation des Informationsbedarfs. Schließlich stellt die unternehmensweit uneinheitlich genutzten Bezeichnungen für dieselben Informationen (Synonyme) bzw. identisch genutzte Bezeichnungen für unterschiedliche Informationen (Homonyme) eine zentrale Herausforderung in der Informationsbedarfsanalyse dar. Insbesondere in diesem Zusammenhang kann es zu zunächst unerkannten Missverständnissen beim Design eines in sich schlüssigen und einheitlichen Modells kommen, die möglicherweise weitreichende Folgen für eine in sich konsistente Unternehmenssteuerung mit sich bringen (Goeken 2005, S. 175; Stroh et al. 2011, S. 44).

Zur Vermeidung der zuvor beschriebenen Unstimmigkeiten und Probleme kann die sogenannte Kennzahlen-Dimensionsmatrix (KDM) dienen. Im folgenden Abschnitt wird dargestellt, wie die Kennzahlen-Dimensionsmatrix aufgebaut ist und wie sie die Übergänge vom Ist-Zustand zum Soll-Zustand sowie vom Grobkonzept zum Feinkonzept im Rahmen der Informationsbedarfsmodellierung unterstützt.

16.4.2 Identifikation von Kennzahlen-Dimensionen-Kombinationen

Zur Identifikation von relevanten Kennzahlen-Dimensionen-Kombinationen kann die zuvor erwähnte KDM genutzt werden. Ziel der KDM ist es, ein Informationskernmodell im Soll-Zustand auf einem mittleren Detaillierungsgrad zu generieren, das alle wesentlichen Informationen für IT-basierte Berichts- und Analyseanwendungen enthält. Hierbei ist zu berücksichtigen, dass die KDM im vorliegenden Fall auf die Anwendung von geschäftsmodellrelevanten Domänen ausgelegt ist und sich damit insbesondere zur Erstellung von Modellen im Bereich der Unternehmenssteuerung eignet.

Im Rahmen der KDM werden Kennzahlen, Dimensionen und Dimensionsebenen sowie Wertarten definiert, die aus den fachlichen Steuerungsanforderungen abgeleitet werden und zugleich als solide Ausgangsbasis für eine detailreichere Modellierung des Informationsbedarfs dienen. Mit Hilfe eines derartigen Modells kristallisiert sich die Menge der wesentlichen Informationen und deren Struktur heraus, die bislang im Unternehmen genutzt worden sind. Ferner wird die Geschäftsmodellrelevanz dieser Informationen bestimmt, sodass der Objektivitätsgrad der Steuerungsinformationen transparent wird. Schließlich wird die Verfügbarkeit dieser Daten innerhalb der aktuellen Systemlandschaft abgebildet, sodass neben der nachfrageorientierten auch der angebotsorientierten Informationsbedarfsermittlung Rechnung getragen wird (Navrade 2011, S. 165 ff.). Um diese Aspekte zu verdeutlichen, werden in Abb. 16.8 einige ausgewählte Kennzahlen, wie beispielsweise der Umsatz, aus dem zuvor beschriebenen Geschäftsmodell eines Logistikdienstleisters in die KDM übertragen.

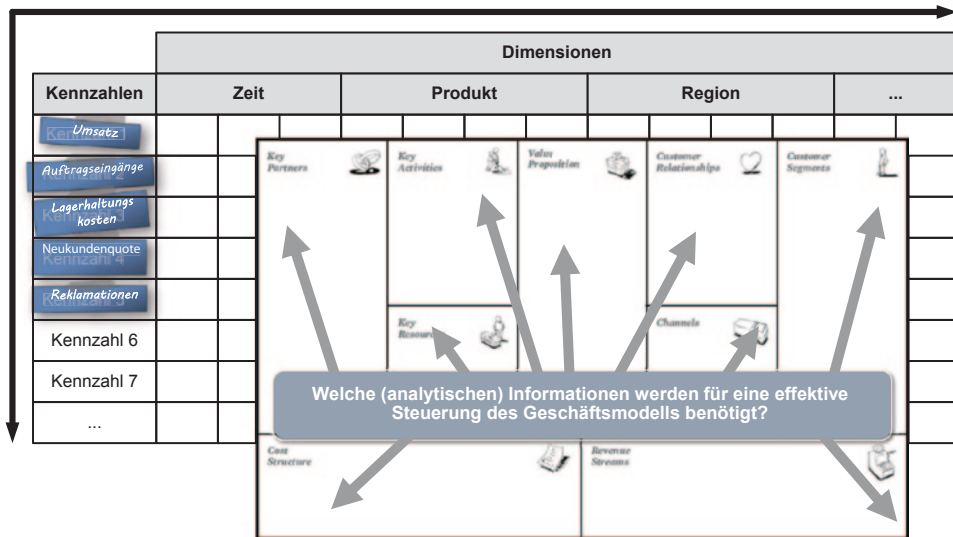


Abb. 16.8 Ermittlung steuerungsrelevanter Auswertungsdimensionen

Das Vorgehen zum Aufbau einer solchen KDM kann grundsätzlich in fünf Schritte gegliedert werden. Im ersten Schritt wird die Anwendungsbreite und -tiefe festgelegt. Für die Anwendungsbreite muss zunächst definiert werden, welche Komponenten des Geschäftsmodells zu erfassen sind. Die Anwendungstiefe bestimmt ferner, welche Management-Level Berücksichtigung finden. Im Zuge dessen besitzen die Anwendungsbreite und -tiefe einen direkten und zugleich wesentlichen Einfluss auf die Komplexität der KDM.

Im zweiten Schritt wird die KDM sukzessive aufgespannt. Der Spaltenaufbau der KDM gliedert sich dabei wie folgt. In der ersten Spalte wird die Bezeichnung einer Kennzahl eingetragen, die aus dem zugrundeliegenden Berichtsdokument stammt. Die zweite Spalte enthält eine Berichts-ID, damit die Rückverfolgung zu den Originaldokumenten des betrachteten Logistikdienstleisters gewährleistet ist. In den weiteren Spalten der KDM werden die relevanten Dimensionen sowie Dimensionsebenen der einzelnen Kennzahlen gemäß Originaldokument identifiziert (Strauch 2002, S. 190). Im hier zugrunde gelegten Beispiel wird zwischen den Wertarten *Ist*, *Plan* und *Forecast* unterschieden. Ein *IP* innerhalb einer Zelle bedeutet somit, dass die Kennzahl für die zugeordnete Dimensionsebene in den Wertarten *Ist* und *Plan* verwendet wird. Im Laufe der Erfassung der Kennzahlen aus den Berichtsdokumenten und deren wertartenspezifischer Zuordnung zu den Dimensionsebenen wird die KDM sukzessive um zusätzliche Spalten und Zeilen erweitert. Die KDM wächst dabei in der Regel mit jedem weiteren analysierten Dokument. Bereits während der Erfassung kann konstatiert werden, dass einige Kennzahlen und Dimensionen häufig in leicht unterschiedlicher Bezeichnung mehrfach vorkommen.

Das Konsolidieren der KDM sowie das Zusammenführen der Kennzahlen und Dimensionen auf einen gemeinsamen redundanzfreien Kern werden in einem dritten Schritt durchgeführt. Nach der sukzessiven Erweiterung wird die KDM nun auf möglichst wenige Zeilen und Spalten reduziert. Dabei sind homonyme und synonyme Bezeichnungen in Kennzahlen, Dimensionen bzw. Dimensionsebenen zu identifizieren. Kennzahlen, die in den Dokumenten mit einer festen Dimensionalität versehen waren, werden davon befreit. Als Beispiele können der Umsatz *Nordeuropa* und Umsatzerlöse *Südeuropa* dienen, die aus jeweils unterschiedlichen Berichten in der KDM erfasst worden sind und sich zunächst als zwei unterschiedliche Kennzahlen darstellen. Hinter beiden Kennzahlen steckt jedoch die identische Messgröße *Umsatz*. *Nordeuropa* und *Südeuropa* stellen demnach lediglich eine Differenzierung dieser Messgröße entlang einer regionalen Dimension dar. In diesem Beispiel werden die beiden Kennzahlen *Umsatz Nordeuropa* und *Umsatz Südeuropa* auf die Kennzahl *Umsatz* reduziert und die Zuordnung zur Dimension bzw. Dimensionsebene *Region* vorgenommen. Existiert eine Dimension bzw. Dimensionsebene bereits in einer Spalte der KDM, wird die entsprechende Zuordnung über die Matrix vorgenommen. Fehlt eine identifizierte Dimension bzw. Dimensionsebene in den Spalten der KDM, wird eine neue Spalte *Dimension* bzw. *Dimensionsebene* erstellt und der Kennzahl zugeordnet. Nach Beendigung dieses aufwändigen Konsolidierungsprozesses wird schließlich der gemeinsame Informationskern (Ist-Zustand) aller analysierten Dokumente offengelegt und die Problematik von Synonymen und Homonymen erkannt.

Im vierten Schritt wird der Ist-Zustand der KDM mit dem Soll-Zustand verglichen. Dazu ist ein Abgleich mit einem strategieorientierten Konzept notwendig (Meyer und Strauch 2000, S. 93; Prakash und Gosain 2008, S. 49). Im vorliegenden Fall erfolgt dieser Schritt anhand eines Abgleichs der Kennzahlen und deren Dimensionalität mit den zuvor beschriebenen Geschäftsmodellkomponenten. Die dabei identifizierten Elemente, die sich im Geschäftsmodell wiederfinden, werden innerhalb der KDM als relevant markiert. Zugleich muss die KDM um die Kennzahlen und Dimensionen erweitert werden, die bislang nicht enthalten sind, aber gemäß Konzept als strategisch relevant erachtet werden (Bidgood und Jelley 1991, S. 38; Navrade 2011, S. 167 ff.). Abbildung 16.9 zeigt exemplarisch sowohl markierte als auch neue Elemente innerhalb der KDM.

Der theoretische Idealfall, dass alle Informationen der KDM als geschäftsmodellrelevant markiert werden können, wird in der Praxis kaum eintreffen. Vielmehr zeigt sich häufig eine nur sehr punktuelle Abdeckung der geschäftsmodellrelevanten Kennzahlen innerhalb der KDM. Dafür kann es mehrere Gründe geben. In den analysierten Dokumenten kann sich beispielsweise über längere Zeit hinweg der subjektive Informationsbedarf manifestiert haben. Kennzahlen und deren Analyseperspektiven sind dabei mitunter in die Betrachtung aufgenommen worden, da sie zu einem gewissen Zeitpunkt als hilfreich erachtet worden sind oder besondere Ereignisse genau diese Betrachtung erforderlich gemacht haben. Die Dokumente sind dann speziell für diesen Zweck zum entsprechenden Zeitpunkt erweitert worden, ohne sie jedoch später wieder zu reduzieren. Darüber hinaus werden je nach Festlegung der Anwendungstiefe auch Kennzahlen oder Dimensionen mit stark operativem Charakter in der KDM erfasst. Solche Kennzahlen fallen typischerweise

		Dimensionen												
Kennzahlen	Berichts-ID	Zeit			Produkt				Region				...	
		Jahr	Monat	Tag	Ebene 1	Ebene 2	Ebene 3	Ebene 4	Ebene 1	Ebene 2	Ebene 3	Ebene 4	Ebene 1	Ebene 2
Kennzahl 1	4			I	I	I			I	I	I	I		
Kennzahl 2	4	IP	I	I	I	IPF	I	I		IP	I			I
Kennzahl 3	4	I	I	I			I	I	I	I	I		I	I
Kennzahl 7		I	I				I	I				I		

Legende

I = Ist-Wert

F = Forecast-Wert

P = Plan-Wert

= Geschäftsmodell-Relevanz

= Neue Elemente

Abb. 16.9 Analyse der Informationswerte mit Geschäftsmodellrelevanz. (Navrade 2011, S. 171)

durch das geschäftsmodellorientierte Raster. Ein weiterer Grund für die häufig nur punktuelle Abdeckung der KDM ist der Umstand, dass Kennzahlen durch logische Beziehungen zu Kennzahlensystemen verknüpft werden können. Geschäftsmodellrelevanz besitzen dabei allerdings meist nur die oberen Kennzahlen eines derartigen Systems, während die im Steuerungsprozess beeinflussbaren Kennzahlen eher im hierarchisch unteren Bereich des Systems zu finden sind. In einem solchen Fall ist zu prüfen, welche dieser Kennzahlen explizit in das Steuerungskonzept Einzug gefunden haben und welche Kennzahlen mittels der KDM bereits implizit erfasst worden sind.

Im fünften und letzten Schritt wird die KDM um nicht mehr relevante Kennzahlen und Dimensionen bereinigt. Zusammen mit den erfolgten Erweiterungen um bislang fehlende Kennzahlen und Dimensionen wird der Soll-Zustand dargestellt. In diesem letzten Schritt wird zudem geprüft, in welchem Grad die erforderlichen Daten zur Berechnung der Kennzahlen in der aktuellen Systemlandschaft verfügbar sind (Goeken 2005, S. 169).

Wie aus dem dargelegten Vorgehen ersichtlich wird, kann mit Hilfe der KDM in mehreren Schritten ein Informationskernmodell erarbeitet werden, das den konsolidierten Informationsbedarf gemäß definierter Anwendungsbreite und -tiefe auf einem mittleren Detaillierungsgrad repräsentiert. Dieses Informationskernmodell bietet sich neben der rein fachlichen Informationsbedarfsermittlung auch als solide Ausgangsbasis für ein Feinkonzept an, das beispielsweise mit der ADAPT-Notation modelliert werden kann (Goeken 2005, S. 170; Navrade 2011, S. 167 ff.). Im folgenden Abschnitt wird nun mit Blick auf die zuvor abgeleiteten fachlichen Anforderungen sowie die steuerungsrelevanten Kennzahlen eines Logistikdienstleisters exemplarisch die Planung und Steuerung erfolgskritischer Komponenten anhand der Customer Segments aufgezeigt.

16.4.3 Informationssystem zur Steuerung des Geschäftsmodells

Mit Blick auf das zuvor dargelegte Geschäftsmodell eines Logistikdienstleisters und der Strukturierung des Analyse-raums anhand der KDM, kann nun ein erster konzeptioneller Entwurf für die Entwicklung eines Steuerungssystems erarbeitet werden. Dabei werden die vorgestellten neun Komponenten eines Geschäftsmodells herangezogen und deren Objekte anhand quantitativer und qualitativer Merkmale unterschieden. Die im Vorfeld bereits identifizierten und entsprechend zugeordneten Kennzahlen stellen die sogenannten quantitativen Merkmale dar. Neben diesen und auch weiteren Kennzahlen sowie deren Anreicherung mit erforderlichen Auswertungsdimensionen (Kennzahlen-Dimensionen-Kombinationen) können effektive Planungs- und Steuerungsaktivitäten systemseitig umgesetzt werden. Qualitative Merkmale spiegeln hingegen die deskriptiven Aspekte in der Unternehmenssteuerung wider, die ebenfalls innerhalb der Entwicklung eines Informationssystems Berücksichtigung finden sollten. Im Rahmen der qualitativen Merkmale kann es sich beispielsweise um die Beschreibung von Produktkategorien oder auch die Kommentierungen von Partnerbeziehungen handeln. In diesem Kontext stellen zudem ordinale Bewertungsschemata relevante Möglichkeiten dar, Objekte der fokussierten Geschäftsmodellkomponenten qualitativ zu bewerten und somit einen Mehrwert innerhalb der Planung und Steuerung zu generieren. Damit die vorangegangenen Ausführungen greifbarer werden, stellt Abb. 16.10 die Geschäftsmodellkomponente *Customer Segments* mit den zuvor bereits erarbeiteten Objekten sowie qualitativen Merkmalen dar.

Die Objekte, wie beispielsweise die Lebensmittelhersteller, werden dabei in unterschiedlichen Rubriken abgebildet und daran anlehnend innerhalb des betrachteten Informationssystems als Reiter dargestellt. Mit diesen Reitern können die einzelnen Kunden des abgestellten Geschäftsmodells erfasst (Reiter: *Kunde erfassen*) und nach Relevanz mit Hilfe unterschiedlicher Kriterien, wie Umsatzanteile etc., priorisiert werden (Reiter: *Kunde priorisieren*). Ferner können die Interessen der Kunden erfasst (Reiter: *Interessen erfassen*) sowie kundenspezifische Fundamentalziele formuliert werden (Reiter: *Fundamentalziele formulieren*).

Bei Auswahl des Reiters *Fundamentalziele formulieren* innerhalb der Stakeholder-Kategorie *Kunde* stehen unterschiedliche Bewertungsmöglichkeiten zur Auswahl, die in einem direkten Zusammenhang zu der definierten Geschäftsmodellkomponente *Value Propositions* stehen. Die zuvor aufgezeigten Value Propositions umfassen unter anderem die Lebensmittelsicherheit, die gegenständlich durch die Temperatur (Kühlkette) innerhalb der Logistikdienstleistungen zum Ausdruck gebracht wird. Diese stakeholder-spezifischen Nutzenversprechen können im dargestellten Informationssystem innerhalb der Kategorien *Interessen* Berücksichtigung finden, indem sie aufgelistet und mit entsprechenden Priorisierungen versehen werden. Neben dieser einfachen Auflistung der Value Propositions in Form von spezifischen Interessen können weitere Details zu diesen Interessen textuell erfasst werden. Auf diese Weise können die meist abstrakten Value Propositions

Abb. 16.10 Planung und Steuerung der Kundensegmente des Geschäftsmodells

konkretisiert und deren Priorisierung nachvollziehbar begründet werden. Im vorliegenden Beispiel sind zur Illustration die hypothetisch vereinbarten Temperaturspannen hinterlegt, die zu Steuerungszwecken abgerufen werden können. Korrespondierend mit der Kategorie *Interessen* werden innerhalb der Kategorie *Fundamentalziele* übereinstimmende Ziele definiert, die für die jeweiligen Customer Segments sowie für den langfristigen Erfolg des zugrundeliegenden Geschäftsmodells von zentraler Bedeutung sind. Im vorliegenden Fall ist die hohe Zuverlässigkeit (Kühlkette) einer der Garanten für die Erfüllung der identifizierten Kundenanforderungen sowie für dauerhafte und stabile Kundenbeziehungen. Zugleich kann durch eine hohe Zuverlässigkeit bei der Einhaltung der Kühlkette die Zahlung festgelegter Pönalen vermieden werden, die möglicherweise einschneidende monetäre Auswirkungen auf die Cost Structure des Logistikdienstleisters mit sich bringen könnten. Ebenso wie die zuvor beschriebenen Interessen können auch die Fundamentalziele anhand einer Ordinalskala vor dem Hintergrund ihrer Geschäftsmodellrelevanz eingestuft werden. Analog zur Illustration der hier fokussierten Customer Segments anhand qualitativer Merkmale sind die weiteren acht Komponenten eines prototypischen Geschäftsmodells zu integrieren. Dabei sollten für jede einzelne Komponente sowohl die zuvor exemplarisch dargestellten Kennzahlen-Dimensionen-Kombinationen (quantitative Merkmale) als auch die deskriptiven Steuerungsparameter (qualitative Merkmale) bei der Entwicklung des Informationssystems berücksichtigt werden.

Durch die hier dargelegten Zusammenhänge zwischen den einzelnen Geschäftsmodellkomponenten wird deutlich, dass die Komponenten zwar isoliert voneinander erhoben

werden können, sich ihre Wirkungskraft für eine effektive Planung und Steuerung des Geschäftsmodells allerdings erst durch ihr inhärentes Zusammenspiel entfalten kann.

16.5 Fazit und Ausblick

Den Ausgangspunkt für den geschäftsmodellbasierten Lösungsansatz bilden die aktuell immer noch unzureichenden Konzepte zur Planung und Steuerung von Unternehmen. Die Hauptursache für die Schwächen derzeitig eingesetzter Konzepte liegt im Wesentlichen in der Vernachlässigung erfolgskritischer Steuerungskomponenten. Weder das DuPont-Schema, das sich durch seinen Fokus auf rein finanzwirtschaftliche Kennzahlen auszeichnet, noch die BSC, die als Inbegriff für ausgewogene Steuerungskonzepte gilt, konnten bislang dem Anspruch eines ganzheitlichen und effektiven Ansatzes gerecht werden.

Erst durch die Integration der hier proklamierten erfolgskritischen Steuerungskomponenten, die durch das individuelle Geschäftsmodell eines Unternehmens vorgegeben werden, kann die Grundvoraussetzung für eine effektive Planung und Steuerung geschaffen werden. Hintergrund für diese These ist die Tatsache, dass der Erfolg ebenso wie auch der Misserfolg einer Unternehmung durch das Zusammenspiel der Komponenten eines Geschäftsmodells determiniert wird. Demzufolge gilt es diese Komponenten als zentrale Stellschrauben innerhalb der IT-gestützten Unternehmenssteuerung zu berücksichtigen. Zudem stellt die Betrachtung von insgesamt neun Geschäftsmodellkomponenten ein solides und zugleich breites Fundament für die Unternehmenssteuerung dar. Der Nachweis hierfür wird nicht zuletzt durch den aufgezeigten schwachen Abdeckungsgrad der BSC-Perspektiven innerhalb der Geschäftsmodellkomponenten erbracht. Ferner kann das Geschäftsmodell als wirksames Bindeglied zwischen der Strategie- und Prozessebene eines Unternehmens fungieren, da es einerseits die strategische Ausrichtung konkretisiert und andererseits neben den prozessorientierten Key Activities weitere relevante Steuerungskomponenten aufweist (Möller et al. 2011, S. 222). Vor dem Hintergrund dieser Erkenntnisse können die in Abb. 16.11 dargestellten Evolutionsstufen hergeleitet werden, die mit der geschäftsmodellbasierten Unternehmenssteuerung ihren aktuellen Entwicklungsstand widerspiegeln.

Auch wenn diese dritte Evolutionsstufe heutzutage einen konzeptionellen De-facto-Standard darstellen sollte, so sind auch zukünftig in der geschäftsmodellbasierten Unternehmenssteuerung und ihrem beschriebenen Vorgehen zur Identifikation von Kennzahlen-Dimensionen-Kombinationen gewiss weitere Verbesserungen erforderlich. Denkbar wäre es beispielsweise, bedeutende Steuerungsmethoden angrenzender Bereiche, wie den DMAIC-Zyklus (Define – Measure – Analyze – Improve – Control), in das beschriebene Konzept zu integrieren und damit einen weiteren Beitrag für eine effektive Geschäftsmodellsteuerung zu leisten. Außerdem könnten durch die zunehmenden technologischen Möglichkeiten und den Einsatz analytischer Methoden weitere Potenziale in der Unternehmenssteuerung gehoben oder gar neue Geschäftsmodelle generiert werden (Petry 2013, S. 66).

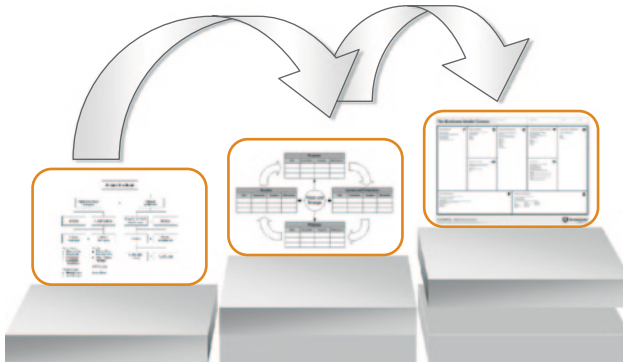


Abb. 16.11 Evolutionsstufen in der Unternehmenssteuerung

Literatur

- Bange, C., Dahnken, O., Friedrich, D.: Planung und Budgetierung in europäischen Unternehmen – Konzepte, Lösungen und Potenziale von Performance Management. BARC GmbH., Würzburg (2008)
- Baum, H-G., Coenenberg, A.G., Günther, T.: Strategisches Controlling, 4. Aufl. Schäffer-Poeschel, Stuttgart (2007)
- Bidgood, T., Jelley, B.: Modeling corporate information needs: fresh approaches to the information architecture. *J. Strategic Inf. Syst.* **1**(1), 38–42 Elsevier, (1991)
- Chamoni, P., Linden, M.: Business Intelligence – Entscheidungsunterstützung auf Basis analytischer Informationssysteme. *WiSt – Wirtschaftswissenschaftliches Studium*, 40. Jahrgang, Heft 6, S. 276–282. C.H.BECK oHG., München (2011)
- Frie, T., Strauch, B.: Die Informationsbedarfsanalyse im Data Warehousing – ein methodischer Ansatz am Beispiel der Balanced Scorecard. In: Britzelmaier, B., Geberl, S., Weinmann, S. (Hrsg.) *Informationsmanagement – Herausforderungen und Perspektiven*, 3. Liechtensteinisches Wirtschaftsinformatik-Symposium an der FH Liechtenstein. S. 241–253. Teubner, Stuttgart (2001)
- Gladen, W.: Performance Measurement – Controlling mit Kennzahlen, 4. Aufl. Gabler, Wiesbaden (2008)
- Gluchowski, P., Kemper, H-G., Seufert, A.: Was ist neu an Operational BI? Innovative Prozess-Steuerung. *BI-Spektrum* **4**(1), 8–11. SIGS DATACOM GmbH., Troisdorf (2009)
- Goeken, M.: Anforderungsmanagement bei der Entwicklung von Data Warehouse-Systemen – Ein sichtenspezifischer Ansatz. In: Schelp, J., Winter, R. (Hrsg.) *Auf dem Weg zur Integration Factory: Proceedings der DW2004– Data Warehousing und EAI*. S. 167–186. Physica, Heidelberg (2005)
- Grüning, M.: Performance-Measurement-Systeme – Messung und Steuerung von Unternehmensleistung, 1. Aufl. Springer, Wiesbaden (2002)
- Jossé, G.: Balanced Scorecard – Ziele und Strategien messbar umsetzen, 1. Aufl. Deutscher Taschenbuch Verlag, München (2005)
- Jung, R.: Architekturen zur Datenintegration – Gestaltungsempfehlungen auf der Basis fachkonzeptueller Anforderungen, zugl. Habilitationsschrift Univ. St. Gallen 2005. Deutscher Universitätsverlag, Wiesbaden (2006)
- Kaplan, R.S., Norton, D.P.: Balanced scorecard – translating strategy into action, 1. Aufl. Harvard Business Review Press, Boston (1996)

- Kaplan, R.S., Norton, D.P.: *Balanced Scorecard – Strategien erfolgreich umsetzen*, 1. Aufl. Schäffer-Poeschel, Stuttgart (1997)
- Kaplan, R.S., Norton, D.P.: *The strategy-focused organization – how balanced scorecard companies thrive in the new business environment*, 1. Aufl. Harvard Business Review Press, Boston (2001)
- Konzelmann, R.: Vorgehensmodell zur Erstellung eines Enterprise Data Warehouse. In: Töpfer, J., Winter, R. (Hrsg.) *Active Enterprise Intelligence – Unternehmensweite Informationslogistik als Basis einer wertorientierten Unternehmenssteuerung*, S. 273–311. Springer, Berlin (2008)
- Koreimann, D.S.: *Methoden der Informationsbedarfsanalyse*. De Gruyter, Berlin (1976)
- Küpper, H-U.: *Controlling – Konzeption-Aufgaben-Instrumente*, 5. Aufl. Schäffer-Poeschel, Stuttgart (2008)
- Lynch, R.L., Cross, K.F.: *Measure up! yardsticks for continuous improvement*, 2. Aufl. John Wiley & Sons, Cambridge (1995)
- Meyer, M.; Strauch, B.: Organisationskonzepte im Data Warehousing. In: Jung, R., Winter, R. (Hrsg.) *Data-Warehousing-Strategie: Erfahrungen, Methoden, Visionen*, S. 79–100. Springer, Berlin (2000)
- Möller, K., Drees, A., Schläfke, M.: Performance Management zur Steuerung. In: Bieger, T., Zu Knyphausen-Aufseß, D., Kryz, C. (Hrsg.) *Innovative Geschäftsmodelle – Konzeptionelle Grundlagen, Gestaltungsfelder und unternehmerische Praxis*, 1. Aufl., S. 213–228. Springer, Berlin (2011)
- Navrade, F.: Informationsbedarfsanalyse mittels Kennzahlen-Dimensionsmatrix. In: Felden, C., Krebs, S., Stock, S. (Hrsg.) *Perspektiven der Business Intelligence: Festschrift anlässlich des 60. Geburtstages von Prof. Dr. Peter Chamoni*, 1. Aufl., S. 165–174. Hanser, Paderborn (2011)
- Osterwalder, A., Pigneur, Y.: *Business model generation*, 1. Aufl. John Wiley & Sons, New Jersey (2010)
- Otley, D.: Accounting performance measurement: a review of its purposes and practices. In: Neely, A. (Hrsg.) *Business Performance Measurement – Unifying theories and integrating practice*, 2. Aufl., S. 11–35. Cambridge University Press, Cambridge (2007)
- Petry, M.: Schwerpunkt Interview: „Nach unserem Verständnis ist Big Data die Vision von der Entwicklung neuer Geschäftsmodelle durch die Sammlung und Analyse großer Datenmengen.“ – Prof. Dr. Hans Ulrich Buhl mit dem CIO der Hilti AG, Dr. Martin Petry. *Wirtschaftsinformatik Manag.* **5**(2), 66–69. Springer Gabler, Wiesbaden (2013)
- Prakash, N., Gosain, A.: An approach to engineering the requirements of data warehouses. *Requirements Eng* **13**(1), 49–72 Springer, (2008)
- Preißler, P.R.: *Betriebswirtschaftliche Kennzahlen: Formeln – Aussagekraft – Sollwerte – Ermittlungsintervalle*, 1. Aufl. Oldenbourg Wissenschaftsverlag, München (2008)
- Preißner, A.: *Praxiswissen Controlling: Grundlagen -Werkzeuge – Anwendungen*, 6. Aufl. Hanser, München (2010)
- Reichmann, T.: *Controlling mit Kennzahlen: Die systemgestützte Controlling-Konzeption mit Analyse- und Reportinginstrumenten*, 8. Aufl. Vahlen, München (2011)
- Silvi, R., Moeller, K., Schlaefke, M.: *Performance management analytics – The next extension in managerial accounting*. SSRN eLibrary, Rochester (2010)
- Strauch, B.: *Entwicklung einer Methode für die Informationsbedarfsanalyse im Data Warehousing; zugl. Diss. Univ. St. Gallen 2001. ECONBIZ*, St. Gallen (2002)
- Stroh, F., Winter, R., Wortmann, F.: Methodenunterstützung der Informationsbedarfsanalyse analytischer Informationssysteme – Stand der Forschung, Anforderungen aus der Praxis und Erweiterungspotenziale. *Wirtschaftsinformatik* **53**(1), 37–48 Gabler, (2011)

Stichwortverzeichnis

A

ACID, 206, 208, 209, 210, 212, 213
Advanced Analytics, 61, 227
Aggregation, 9, 100, 129, 131, 132, 139, 141, 149, 152, 153, 156, 159, 171, 216, 231, 250
Agile BI, 35, 226, 232, 233, 237
Analytische Datenbanken, 110
Analytische Datenbanken, 108
Anchor-Modeling, 175
Anreicherung, 9, 129, 131, 132, 141, 143, 154, 156, 159, 171, 293, 297, 345
Antwortzeitverhalten, 141
Appliances, 193, 201
Architektur
 Data-Mart-Bus, 155
 Hub-and-Spoke, 154
Availability, 210

B

Balanced Scorecard, 38, 43, 44, 113, 114, 326, 330
Bereinigung, 107, 131, 133, 134, 136, 137, 152, 159, 171, 183, 254
Berichtswesen, 45, 50, 113, 123, 124, 125, 305, 306
BICC (BI-Competency-Center), 24, 34, 41, 45, 46, 48, 286, 298
BI-Competency-Center siehe BICC (BI-Competency-Center), 45
Big Data, V, 3, 5, 7, 9, 16, 25, 29, 30, 34, 48, 49, 52, 56, 57, 58, 59, 60, 61, 62, 63, 64, 76, 77, 100, 101, 104, 107, 109, 124, 125, 143, 144, 198, 201, 205, 206, 212, 213, 214, 215, 219, 220, 221, 222, 226, 227, 285, 286, 289, 291, 298, 304, 306, 315, 316, 317, 319

BI-Governance, 23, 24, 28, 41, 53
BI siehe auch Business Intelligence, 23
BI-Strategie, 34, 35, 36, 37, 38, 39, 40, 42, 43, 44, 46, 48, 49, 50, 51, 52
Vorgehensmodelle, 36
Business Engineering, 69, 70, 76, 77, 78, 86
Business Intelligence, V, VII, IX, 8, 9, 11, 13, 15, 16, 17, 18, 20, 21, 23, 25, 29, 33, 34, 36, 39, 40, 41, 43, 46, 47, 49, 52, 56, 57, 71, 101, 106, 107, 111, 112, 124, 126, 162, 188, 193, 198, 220, 226, 236, 243, 260, 269, 278, 283, 285, 286, 304, 306, 307
Business Intelligence Competency Center siehe BICC (BI-Competency-Center), 34
Business-Intelligence-Strategie, 9, 33
Business Model Canvas, 333, 335, 337

C

Capabilities, 28, 69
Chunks, 216, 217
CIF (Corporate Information Factory), 156
CIM (Computer Integrated Manufacturing), 264, 265, 266
Cloud BI, 34
Cloud Computing, 21, 143, 144, 208, 275
Competence Center, 82, 83
Compliance, 89, 160
conformed dimension, 151
conformed facts, 151
Core Data Warehouse, 152
Corporate Information Factory, 156
corporate memory, 196
CPPS (cyberphysisches Produktionssystem), 276, 278, 279

CPPS(cyberphysisches Produktionssystem),
261

Cross Media Storage Manager, 157
cyberphysisches Produktionssystem, 261
Cyber-Physische Systeme, 271

D

Dashboards, 71, 91, 112, 305, 306, 311, 313,
320
Data-Mart-Busarchitektur, 155
Data Marts, 49, 52, 62, 109, 119, 130, 149, 150,
151, 153, 154, 155, 156, 157, 158, 160, 162,
164, 171, 172, 230
abhängige, 151
unabhängige, 150
Data Mining, XI, XIII, XIV, 3, 5, 7, 11, 20, 25,
28, 58, 75, 99, 105, 111, 119, 123, 124, 125,
220, 227, 278, 286
Data Scientist, 29, 125, 305, 307, 318, 319,
320, 322
Data Vault, 172
Methode, 172
Data Warehouse, XI, XIII, XIV, 3, 7, 8, 9, 10,
11, 15, 35, 47, 48, 58, 62, 63, 64, 100, 105,
108, 109, 111, 118, 129, 130, 132, 134, 136,
137, 139, 140, 141, 143, 144, 148, 151, 152,
153, 154, 155, 156, 157, 158, 159, 160, 161,
162, 163, 164, 166, 167, 168, 170, 171, 172,
173, 179, 181, 182, 183, 184, 189, 192, 193,
194, 195, 196, 198, 200, 219, 220, 286, 293,
294, 295, 296, 297, 298
Datenbanken
analytische, 108, 110
multidimensionale, 109
relationale, 15, 108, 208
Datenintegration, 105
Werkzeuge, 108
Datenmanagement, 98, 100, 101, 103, 104,
105, 109, 111, 125, 144, 226, 228, 237, 268,
279
Daten, polystrukturierte, 106
Datenqualität, 21, 33, 35, 36, 42, 48, 58, 99,
106, 133, 134, 161, 199, 228, 288, 293
Datenschutz, 58, 229
Datenspeicherung, 108
Dimensional Data Warehouse, 155
dimension tables, 139
DuPont-Schema, 328
DW 2.0, 157

E

EFQM-Modell, 329
Enterprise Architecture Management, 79
Enterprise Bus Architecture, 156
Entscheidungsunterstützungssystem, 18, 198
ERP-Systeme, 79, 130, 243, 246, 247, 248,
249, 252, 255, 266, 267
ETL (Extraktions-, Transformations- und Lade-
prozesse), 21, 83, 106, 107, 167, 170, 171,
174, 181, 182, 183, 195, 230, 231, 253
Eventual Consistency, 205, 208, 209, 210, 212,
213, 221
Extraktion, 25, 99, 106, 131, 133, 159, 161,
167, 270, 292, 296

F

fact tables, 139
Filterung, 9, 129, 131, 133, 137, 139, 143, 159,
171, 183
Führungsinformationssystem, 6, 75

G

Geschäftsmodell, 59, 60, 326, 327, 332, 333,
335, 341, 343, 345, 347
Governance, 42, 64, 76, 80, 81, 87, 118, 144,
160, 228, 229, 286
Granularität, 154

H

Hadoop, 10, 34, 61, 106, 109, 143, 205, 206,
208, 211, 213, 218, 220, 221, 319
Harmonisierung, 9, 129, 131, 137, 139, 143,
149, 152, 153, 159, 160, 167, 171, 173, 180,
181, 183, 192, 198
HDFS (Hadoop File System), 109, 208, 211,
218
Hub-and-Spoke, 62, 154
Hub-Tabelle, 172

I

Industrie 4.0, 11, 260, 261, 262, 270, 271, 273,
275, 276, 278, 279, 283, 284, 285, 291
Informations-Infrastruktur, 9, 68, 72, 74, 76,
78, 80, 81, 84, 85, 86, 88, 89, 91, 92, 93
Informationslogistik, 4, 8, 15, 39, 71, 73, 77,
79, 81, 87, 197

In-Memory, 10, 61, 62, 106, 109, 142, 144,
189, 192, 195, 199, 200, 215, 286, 288
Internet der Dinge, 271, 273, 274, 276, 278,
279, 284

K

Kennzahlensystem, 43, 327
Key-Value-Datenbanken, 62
Kosten, 34, 46, 47, 48, 49, 52, 64, 74, 84, 85,
86, 89, 93, 111, 141, 190, 192, 247, 254,
258, 264, 293, 330, 335

L

Ladezeitstempel, 172
Ladungsträgermanagement, 293
Link-Tabelle, 176

M

Many-Many-Beziehung, 176
Map-Reduce, 10, 205, 207, 208, 213, 214, 216,
217, 218, 220, 221, 222
MES, 283, 297
MES (Manufacturing Ecution System), 291,
295, 297, 298
MES (Manufacturing Execution System), 260,
267, 269, 270
MES(Manufacturing Execution System), 260,
267
Metadaten, 18, 62, 64, 99, 109, 196, 229, 230,
231
Metadatenverwaltung, 144
Mobile BI, 34, 35, 227, 236, 237
multidimensionale Analyse, 75
multidimensionale Datenbanken, 109
Mustererkennung, 99, 277, 293

N

Near-line Storage, 157
NewSQL, 10, 205, 206, 207, 208, 212, 213,
215, 221
NoSQL, 10, 62, 106, 109, 205, 206, 207, 208,
210, 211, 212, 213, 215, 221, 289
Nutzen, 40, 42, 46, 48, 52, 58, 74, 80, 89, 92,
156, 160, 219, 230, 317, 327, 332

O

ODS (Operational Data Store), 62, 195
OLAP (On-Line Analytical Processing), XIV,
11, 28, 49, 75, 109, 110, 120, 189, 191, 192,
195, 201, 220, 227, 244, 246, 247, 248, 249,
250, 251, 252, 253, 254, 255, 257, 258, 286,
293
On-Line Analytical Processing, XIII, XIV, 3,
5, 7, 109
On-Line Transaction Processing, 4
OODA-Loop, 14, 25
OpBI, 260, 261, 269, 270, 278, 279
Open Source, 61, 105, 108, 117, 125, 206, 211
Operational BI, 227, 269, 286, 295, 297
Operational Data Store, 62
operatives System, 6, 75, 99, 133, 137, 138,
142, 159, 180

P

Performance Management, 8, 36, 100
PIT-Tabelle, 175
Planung, 7, 8, 11, 15, 39, 50, 72, 74, 77, 80, 84,
98, 99, 111, 123, 124, 125, 199, 219, 242,
243, 247, 249, 250, 252, 253, 254, 255, 261,
264, 267, 268, 284, 286, 288, 326, 327, 332,
338, 339, 340, 344, 345, 346, 347
Point-In-Time-Tabelle, 175
polystrukturierte Daten, 106, 109, 305
Predictive Analytics, 100, 105, 125, 226, 227,
237
Produktionslogistik, 293
Produktionssystem, cyberphysisches, 261
Prognose, 26, 227, 310
Propagation Layer, 162

Q

Quellsysteme, 136, 137, 173, 182, 192, 196
Quellsystemen, 91, 136, 143, 148, 150, 161,
167, 180, 183, 192, 193, 196, 340

R

Radio Frequency Identification, 273, 288
Real-Time, 61, 63, 106, 148, 159, 171, 269
Referenzmodell, 317
relationale Datenbanken, 15, 108, 208

Reporting, 7, 35, 43, 52, 57, 58, 62, 64, 71, 83,
91, 100, 105, 111, 113, 114, 117, 118, 124,
125, 142, 143, 149, 158, 162, 163, 167, 171,
182, 183, 189, 192, 193, 197, 198, 230, 286,
297, 298, 305, 306, 307
Reporting Layer, 162
RFID, 273, 283, 288, 289, 291, 292, 293, 294,
295, 298
Roadmap, 36, 39, 42, 52

S

Sandboxing, 64, 235
Satellite-Tabelle, 172, 173
Scrum, 232, 233
Self-Service-BI, 34, 35, 48, 62, 64, 104, 117,
149, 183, 226, 233, 234, 235, 237
Silo, 150
Simulation, 123, 124, 244, 249, 261, 268, 270,
286, 288
Single Point of Truth, 161
Situation Awareness, 13, 16, 25, 27, 28
Smart Factory, 275
SPOT, 161
Staging-Bereich, 161
Staging-Modell, 177
Stove Pipe, 150
Streaming, 61, 63, 106
Supply Chain Management, 287
Surrogate Identifier, 173
Surrogate-Key, 173

T

Transformation
Aggregation, 159
Anreicherung, 159
Filterung, 159
Harmonisierung, 159

U

Use Case, 64

V

Variety, 56, 57, 214
VBA (Visual Business Analytics), 305, 306,
309, 319, 322
Velocity, 56, 57, 214
Verdichtung, 20, 25, 131, 139, 159, 214, 250,
252, 254
Visual BI, 235
Visual Business Analytics, 11, 305
Visual Business Intelligence, 311
Visualisierung, 58, 61, 118, 120, 235, 268, 269,
277, 278, 305, 308, 309, 313, 317, 318, 319,
321
Volume, 56, 57, 214

W

Workflow, 246, 247