

1 Abgrenzung und Einordnung

Im Bereich der analyseorientierten Informationssysteme gibt es nur wenige Begriffe, die seit den 90er Jahren häufiger und andauernder erwähnt und diskutiert wurden als der des Data Warehouse. Viele Zeitungsartikel, Forschungsbeiträge und Produktinformationen propagieren zwar die Notwendigkeit eines Data Warehouse, es geht aber selten eindeutig hervor, worin die Charakteristika und der Nutzen eines Data Warehouse liegen. Die Verwendung des Begriffes ist derart vielseitig, dass es notwendig ist, nicht nur die Eigenschaften eines Data Warehouse aufzuzeigen, sondern auch eine einheitliche Begriffsverwendung im Sprachgebrauch zu erreichen.¹

Die Vielseitigkeit des Data-Warehouse-Begriffes resultiert aus zwei grundlegenden Bereichen, die diesen Begriff geprägt haben: Auf der technischen Seite stehen die Grundlagen der Datenintegrationsmöglichkeiten und Datenbanksysteme, auf der Anwendungsseite finden sich die betriebswirtschaftlichen, wissenschaftlichen und technischen Anforderungen aus einer Nutzungsperspektive. Weiterhin ist der Einfluss der Marktanalysten, Beratungshäuser und Softwarefirmen nicht zu gering zu erachten. Es ist somit ein Muss, diese oft gegensätzlichen Gebiete in diesem Buch gleichermaßen zu betrachten.

Eine Lösung dieses Dualismus von Informatik und Betriebswirtschaft kann nur in der Kombination liegen. Das Data Warehouse wird deshalb als eine Datenbank gesehen, die aus der technischen Sicht Daten aus verschiedenen Datenquellen integriert und aus der betriebswirtschaftlichen Sicht dem Anwender diese Daten zu Analyse Zwecken zur Verfügung stellt.

Weiterhin soll an dieser Stelle auch angemerkt werden, dass der Begriff »Data Warehouse« zunehmend durch den Begriff »Business Intelligence« ergänzt, überlagert oder oft auch ersetzt wird. Business Intelligence ist als Erweiterung zum Data Warehousing zu sehen, da insbesondere die Anwendungen sowie die anwendungsseitigen Prozesse darunter verstanden werden. Das Data-Warehouse-

1. Leider muss auch nach zwei Auflagen dieses Buches festgehalten werden, dass sich an dieser Tatsache nur wenig geändert hat.

System dient weiterhin der Integration eines zentralen, analyseorientierten Datenbestandes.

In Kapitel 1.1 werden für das weitere Verständnis wichtige Definitionen und Abgrenzungen zu verwandten Bereichen gegeben. In Kapitel 1.2 wird die lange Historie des Themengebietes sowohl von Anwendungs- als auch von Datenbankseite skizziert. Nachfolgend wird der Begriff »Business Intelligence« aufgegriffen (Kap. 1.3), um sowohl aktuelle Tendenzen als auch die Fokussierung dieses Buches herauszuarbeiten. Im daran anschließenden Kapitel folgt ein Überblick über die Vielfältigkeit der möglichen Einsatzbereiche eines Data Warehouse. Ein spezielles Anwendungsbeispiel wird in Kapitel 1.5 vertieft. Dieses Szenario dient als Beispiel für das gesamte Buch. Kapitel 1.6 umreißt abschließend den Inhalt des Buches.

1.1 Begriffliche Einordnung

Zu den drei konventionellen Produktionsfaktoren Boden, Arbeit und Kapital wird immer häufiger die Information als vierte Säule hinzugenommen. Informationen basieren auf Daten, die entweder aus einem Unternehmen selbst stammen oder extern zugekauft werden. Die Tatsache, dass Daten eine besondere Bedeutung zukommt, ist aber nicht allein im betriebswirtschaftlichen Kontext zu finden, sondern gilt ebenso für statistische, wissenschaftliche oder technische Anwendungen.

Verschiedenen Informationssystemen ist gemein, dass Daten erfasst und verwaltet werden. Daten in einem Datenbanksystem zu erfassen, ist an sich nichts Neues. In jedem Unternehmen werden Personaldaten eingegeben oder Verkäufe durch Scannerkassen erfasst. Die Verarbeitung und Verwaltung der Daten geschieht in der Regel aber autonom unter Verantwortung der jeweiligen Abteilung. Interessant wird es erst, Daten aus autonomen Quellen zu vereinen. Dieser Vorgang ist besonders schwierig, wenn heterogene Daten unterschiedlichster Qualität, in verschiedenen Datenformaten, in heterogenen Datenmodellen und Datenbanksystemen gehalten werden.

Zur Vollständigkeit soll an dieser Stelle ein Exkurs zu verschiedenen Integrationsmöglichkeiten gegeben werden. Grundsätzlich wird eine *Backend-Integration* von einer *Frontend-Integration* unterschieden. Während bei der Frontend-Integration die Daten und Applikationen aus verschiedenen Systemen nur durch eine gemeinsame Oberfläche integriert werden (Stichwort: Portal, [Lint00b]), werden bei der Backend-Integration entweder die Daten physisch oder virtuell integriert oder Applikationen über eine gemeinsame Schnittstelle zusammengebracht (Stichwort: Enterprise Application Integration (EAI), [Lint00a]). Das Data Warehouse kann in dieser Klassifikation der Backend-Integration zugeordnet werden. Eine pauschale Bewertung kann leider nicht erfolgen, da die Vorteile jeder Integrationsart von dem jeweiligen Zweck und der Strategie abhängen.

Anzustreben ist grundsätzlich eine möglichst frühzeitige Integration der Daten und Anwendungen, was jedoch immer einen mitunter beträchtlichen Aufwand nach sich zieht.

Ein Data Warehouse ist aber nicht nur von diesem integrativen Aspekt geprägt, sondern zusätzlich vom analytischen Aspekt. Die Verwendung von Daten in operativen Anwendungen war lange Zeit geprägt von einer *transaktionalen* Verarbeitung mit vielen kurzen Lese- und Schreiboperationen. Im Gegensatz dazu steht beim Data Warehouse eine eher vergleichende oder auswertende *analytische* Verwendung der Daten im Vordergrund, bei der auf große Datenmengen lesend zugegriffen wird.

Einige Fragen müssen jetzt erlaubt sein: Was ist eigentlich ein Data Warehouse und was zeichnet es aus? Ist ein Data Warehouse eine integrierte Datenbank oder eine Datenbasis zu Analysezwecken? Wo liegen die Gemeinsamkeiten der Einsatzbereiche? Was bedeutet der Begriff Data-Warehouse-System? Die vielen Interpretationsmöglichkeiten machen es notwendig, einige Begriffe zu definieren.

1.1.1 Definitionen

Die Tatsache, dass dieses Themengebiet sowohl von der Anwendungsseite als auch der Informatikseite durch eigene Fachtermini geprägt ist, impliziert ein unterschiedliches Begriffsverständnis. Verschiedene Normungsgremien versuchen, diese Begriffe zu standardisieren. Diese Bestrebungen waren aber bislang wenig erfolgreich.

Eine der ersten Definitionen des Begriffes Data Warehouse wurde von Inmon geprägt:

»A data warehouse is a subject oriented, integrated, non-volatile, and time variant collection of data in support of management's decisions.« [Inmo96]

Ein Data Warehouse hat seiner Ansicht nach also vier Eigenschaften, die alle der Entscheidungsunterstützung dienen. Die Eigenschaften sollen hier kurz skizziert werden:

- *Fachorientierung* (engl. subject orientation): Der Zweck der Datenbasis liegt nicht mehr auf der Erfüllung einer Aufgabe wie eine Personaldatenverwaltung, sondern auf der Modellierung eines spezifischen Anwendungsziels.
- *Integrierte Datenbasis* (engl. integration): Die Datenverarbeitung findet auf integrierten Daten aus mehreren Datenbanken statt.
- *Nicht flüchtige Datenbasis* (engl. non-volatile): Die Datenbasis ist als stabil zu betrachten. Daten, die einmal in das Data Warehouse eingebracht wurden, werden nicht mehr entfernt oder geändert.

- *Historische Daten* (engl. time variance): Die Verarbeitung der Daten ist so angelegt, dass vor allem Vergleiche über die Zeit stattfinden. Es ist dazu unumgänglich, Daten über einen längeren Zeitraum zu halten.

Diese Definition ist einerseits nicht aussagekräftig genug, um sie in der Praxis oder der Theorie verwenden zu können, andererseits ist sie so einschränkend, dass viele Anwendungsgebiete und Ansätze herausfallen. Eine neue Definition ist notwendig, um dieses Manko zu überwinden:

»*Ein Data Warehouse ist eine physische Datenbank, die eine integrierte Sicht auf beliebige Daten zu Analysezwecken ermöglicht.*«

Aus der vermeintlich trivialen Forderung nach einer Integration entstehen Fragestellungen wie die Integration von Schemata und Daten aus unterschiedlichen Quellen. Diese Thematik ist zwar auch in *föderierten Datenbanksystemen* [Conr97] anzutreffen. Im Unterschied zu diesen besteht aber auch die Forderung nach der *physischen Integration* und dem *Analyseaspekt*, d. h., es wird zusätzlich ein adäquater Modellierungsansatz benötigt. Häufig wird diese Anforderung durch das *multidimensionale Datenmodell* [KRRT98] erreicht, das die Denkweise des Anwenders in *Dimensionen* und *Klassifikationshierarchien* widerspiegelt. Das multidimensionale Datenmodell stellt im Gegensatz zu anderen Datenmodellen besondere Strukturen und Auswertungsmöglichkeiten zur Verfügung, die schon bei der Modellierung einen Analysekontext schaffen. Eine in diesem Zusammenhang wichtige Anwendung ist das *Online Analytical Processing* (OLAP, [CoCS93]), das eine explorative, interaktive Datenanalyse auf der Grundlage des konzeptuellen multidimensionalen Datenmodells darstellt. Weiterhin fällt oft das Stichwort *Data Mining*, darunter versteht man eine Suche nach unbekanntenen Mustern oder Beziehungen im Datenbestand des Data Warehouse (Kap. 3.6.3).

Ein weiterer Unterschied eines Data Warehouse gegenüber anderen Datenbanken liegt darin, dass die Daten in der Regel *nicht modifiziert* werden. Daten, die einmal in das Data Warehouse übernommen wurden, dürfen nicht mehr verändert werden. Es können aber neue Daten in das Data Warehouse aufgenommen werden, ohne die bereits vorhandenen zu überschreiben.

Ein Data Warehouse kann all diese Eigenschaften selten alleine zur Verfügung stellen. Deshalb ist ein Data Warehouse in ein *Data-Warehouse-System* eingebettet. Es umfasst alle für die Integration und Analyse notwendigen Komponenten. Besonders hervorzuheben sind die Komponenten der *Datenbeschaffung*, die für die Integration der Daten notwendig sind, die Komponenten zur *Analyse* und die *Basisdatenbank*, die der Haltung eines homogenen Datenbestands in einer transaktionalen Umgebung dient. Diese Komponenten haben einen *statischen* Charakter.

Der *Data-Warehouse-Prozess*, auch *Data Warehousing* genannt, beschreibt den *dynamischen* Vorgang, angefangen beim *Datenbeschaffungsprozess* über das Speichern bis zur Analyse der Daten, d. h. den Fluss und die Verarbeitung der Daten aus den Datenquellen bis zum Analyseergebnis beim Anwender. Ein Data-Warehouse-System ist also mehr als die Summe seiner Komponenten, d. h., erst mit dem Data-Warehouse-Prozess kann es seine Aufgaben erfüllen. Für die exakten Definitionen sei auf das Glossar in Anhang B verwiesen.

1.1.2 Abgrenzung von transaktionalen Systemen

Die unterschiedlichen Anforderungen rechtfertigen es, von einem Dualismus² zwischen der *Datenintegration* und der *Analysefunktion* zu sprechen. Die Vereinigung von Daten aus diversen Datenquellen, wie sie z. B. in modernen Kaufhäusern vorhanden sind, macht es für eine effiziente Verarbeitung notwendig, sich nicht nur um die Integration der Daten zu kümmern, sondern sie auch genau in der Form bereitzustellen, die der Anwender fordert. Die Frage nach den Anwenderwünschen wird hier deutlich: Für eine Analyse sind nicht alle möglichen Daten notwendig, sondern nur genau die Daten seines Entscheidungsgebiets, für das er in adäquater Zeit Informationen benötigt. Der Anwender braucht also ein Informationssystem, das mehrere Datenquellen vereinigt und einen expliziten Bezug zum jeweiligen Anwendungsfall hat.

Im Gegensatz dazu stehen die transaktionalen Systeme, die oft auch als *Online Transactional Processing* (OLTP) bezeichnet werden. Es gibt Unterscheidungsmerkmale, die sich in die Bereiche *Anfragen*, *Daten* und *Anwender* klassifizieren lassen. Anzumerken ist, dass ein Data-Warehouse-System auch in einem operativen Produktionssystem eingesetzt werden kann; typischerweise ist es aber in die analytisch dispositive Kategorie einzuordnen.

Anfragen

Die Charakteristika von Anfragen haben einen großen Einfluss auf die Anfrageverarbeitung und Speicherform von Daten (Tab. 1–1). Der Fokus unterscheidet sich grundlegend darin, ob Daten in transaktionalen oder analyseorientierten Systemen verwaltet werden. Erstere lesen, modifizieren oder löschen Daten in kurzen und einfachen Transaktionen. Analyseorientierte Systeme hingegen gewinnen häufig Informationen aus den Daten, indem lange Lesetransaktionen in komplexen Anfragen erfolgen. Eine Anfrage im transaktionalen Betrieb betrifft meist nur wenige Datensätze in einem anfrageflexiblen Modell, d. h., keine Anfrage hat eine modellseitige Bevorzugung. Eine analytische Verarbeitung betrifft Millionen von Datensätzen und findet oft ad hoc nach Benutzerwünschen in einem analyseorientierten Modell statt.

2. im Sinne von Gegensätzlichkeit

Anfragen	transaktional	analytisch
Fokus	Lesen, Schreiben, Modifizieren, Löschen	Lesen, periodisches Hinzufügen
Transaktionsdauer und -typ	kurze Lese-/Schreibtransaktionen	lange Lesetransaktionen
Anfragestruktur	einfach strukturiert	komplex
Datenvolumen einer Anfrage	wenige Datensätze	viele Datensätze
Datenmodell	anfrageflexibles Datenmodell	analysebezogenes Datenmodell

Tab. 1-1 Gegenüberstellung der Anfragecharakteristika von transaktionalen und analytischen Anwendungen

Daten

Ein Beispiel für einen transaktionalen Betrieb findet sich in der Personalabteilung. Anwender verwalten Daten über die Angestellten in einer Datenbank. Die Daten werden aus einer Produktionsumgebung durch ein Datenbanksystem verwaltet, um darauf transaktionale ein-/ausgabeorientierte Anwendungen auszuführen. Die Daten eines Data Warehouse stammen ebenso aus der Produktionsumgebung, aber physisch aus einer oder mehreren operativen Datenbanken (Tab. 1–2). Daten in einem Data Warehouse sind also abgeleitete Daten. In der transaktionalen Verarbeitung stehen Aufgaben wie die Konsistenzsicherung einzelner Transaktionen der Sammlung von Daten im Data Warehouse gegenüber. Die Daten sind im transaktionalen Betrieb meist zeitaktuell, nicht abgeleitet, autonom, aus einer Datenquelle und dynamisch, d. h. von vielen Schreibvorgängen und ständigen Modifikationen betroffen. Die Analyseorientierung hingegen erfordert Daten, die konsolidiert, integriert, stabil und meist aggregiert sind. Durch diese Stabilität der Daten und die Vereinigung mehrerer Datenquellen wächst das Datenvolumen von Mega- oder Gigabyte in transaktionalen Anwendungen auf Datenvolumina bis in den Terabyte-Bereich bei analytischen Anwendungen an. Die Datenproblematik macht es auch notwendig, dass das Design der Datenbank sich dem Einsatzzweck anpasst, d. h., das Datenbankschema wandelt sich von der Anwendungsneutralität beim transaktionalen Bereich mit einer Anfrageflexibilität zu einer Auswertungsorientierung mit vorgedachten Analysepfaden. Weiterhin impliziert die Anwendung andere Zugriffsstrukturen: Im transaktionalen Fall finden weitgehend Einzeltupelzugriffe statt. Dagegen sind im analyseorientierten Fall häufig Suchläufe über den gesamten Datenbestand notwendig.

Daten	transaktional	analytisch
Datenquellen	meist eine	mehrere
Eigenschaften	nicht abgeleitet, zeitaktuell, autonom, dynamisch	abgeleitet, konsolidiert, historisiert, integriert, stabil
Datenvolumen	Megabyte – Gigabyte	Gigabyte – Terabyte
Zugriffe	Einzeltuplezugriff	Bereichsanfragen

Tab. 1-2 Gegenüberstellung der Datencharakteristika von transaktionalen und analytischen Anwendungen

Anwender

Der Anwender eines Data-Warehouse-Systems ist typischerweise nicht mehr der Sachbearbeiter, der Daten erfasst oder abfragt, sondern der Manager, Controller oder Analyst, der Daten in verdichteter Form zur Entscheidungsunterstützung benötigt (Tab. 1–3). Die Anzahl der Anwender reduziert sich damit von Tausenden auf wenige Hundert. Eine explizite Nebenläufigkeit mit einem ausgefeilten Transaktionskonzept wird durch den lesenden Zugriff nicht mehr benötigt. In beiden Anwendungsfällen wird eine kurze Antwortzeit erwartet; bei einer Anwendung, die sehr große Datenmengen in komplexen Anfragen verwendet, kann eine Forderung nach Antwortzeiten im Sekunden- bis Minutenbereich zu einem erheblichen Problem werden.

Anwender	transaktional	analytisch
Anwendertyp	Ein-/Ausgabe durch Sachbearbeiter	Auswertungen durch Manager, Controller, Analysten
Anwenderzahl	sehr viele	wenige (bis einige Hundert)
Antwortzeit	ms – s	s – min

Tab. 1-3 Gegenüberstellung der Anwendercharakteristika von transaktionalen und analytischen Anwendungen

1.2 Historie des Themenbereichs

Die Data-Warehouse-Idee ist nicht neu: Entscheidungsträgern in verschiedenen Funktionsbereichen und auf verschiedenen Hierarchieebenen eines Unternehmens sollen im Moment des Informationsbedarfs alle Informationen zur Verfügung stehen, die sie benötigen. Die Informationsbereitstellung soll zeitnah, fehlerfrei, flexibel, ergonomisch, effizient, effektiv und inspirativ erfolgen. Der Begriff *Managementinformationssystem* (MIS) wurde bereits Ende der sechziger Jahre geprägt und ist seither unter wechselnden Bezeichnungen Gegenstand intensiver Bemühungen in Wissenschaft und Praxis. Die Begriffe *Managementinforma-*

tionssystem (MIS), *Executive Information System* (EIS), *Führungsinformationssystem* (FIS), *Chefinformationssystem* (CIS), *Entscheidungsunterstützungssystem* (EUS), *Decision Support System* (DSS) usw. werden hier als weitgehend synonym aufgefasst. Jedes System dieser Kategorie bekommt sein individuelles Erscheinungsbild erst vor dem Hintergrund der spezifischen Anforderungen in Unternehmen. Die nuancierten Abgrenzungen, die teilweise angestellt werden, sind daher überflüssig. Sie tragen eher zur Verwirrung als zur Klärung bei, was durchaus auch die Intention mancher kreativer »Wortgestalter« sein mag [HiMo95a]. Fehlende Voraussetzungen, wie z. B.

- schnelle und flächendeckende Kommunikationstechnologien,
- grafische Benutzeroberflächen,
- ausreichende, kostengünstige und schnelle Datenspeicher,
- kostengünstige und leistungsfähige Prozessoren,
- große Datenbasen durch integrierte operative Systeme,

die die MIS-Ansätze der 60er, 70er und 80er Jahre scheitern ließen, sind heute erfüllt.

Mangelte es früher neben der technischen Infrastruktur vor allem daran, dass keine ausreichende, elektronisch verfügbare und skalierbare Datenbasis vorhanden war, geht es in den meisten Unternehmen heute vor allem um die Frage, das vorhandene Potenzial zu erschließen, möglichst bevor es einem der Wettbewerber gelingt. Das Data Warehousing der 90er Jahre zielte in erster Linie darauf ab, alle entscheidungsrelevanten Informationen verfügbar zu machen, die bereits an irgendeiner Stelle im Unternehmen elektronisch gespeichert sind. Nicht zuletzt durch die zunehmende Verbreitung der geschäftsprozessorientierten Transaktionssysteme (z. B. SAP ERP) ist in Unternehmen ein großes Volumen an potenziell entscheidungsrelevanten Informationen vorhanden. Idealerweise sollen in einem Data-Warehouse-System darüber hinaus auch Informationen aus externen Informationssystemen integriert werden.

Was sich im Laufe der MIS-Bemühungen der 70er Jahre zunächst als Utopie abzeichnete, nämlich das Konzept einer völligen Integration in eine EDV-Lösung mit Zugriff auf eine Datenbank, erhält durch den Fortschritt in der Informationstechnologie im Gewand des Data Warehousing eine Renaissance. Entscheidender Unterschied dabei ist allerdings, dass es sich beim Data Warehouse erstens um redundant gehaltene Daten handelt, die von den Quellsystemen losgelöst sind, und zweitens, dass es sich nur um einen Teil der Daten handelt, der dem jeweiligen Analysezweck dient. Auf dieser Datenbasis setzen dann funktionspezifische und benutzeradäquat gestaltete Managementinformationssysteme aller Facetten bzw. analytische Informationssysteme aus anderen betrieblichen Bereichen auf.

Neben der Einordnung zu Informationssystemen ist es ebenso schwierig, ein Data-Warehouse-System von anderen Ansätzen abzugrenzen. Die Integration

von Daten aus vielen heterogenen, autonomen und verteilten Quellsystemen in einem Data-Warehouse-System kann mit der Integration von Daten in dem Gebiet der *Mehrrechnerdatenbanksysteme* verglichen werden. Eine Unterscheidung eines Data Warehouse von anderen Datenbankansätzen kann nach der *Homogenität*, der *Kopplung* oder der *räumlichen Verteilung* der Datenbanksysteme getroffen werden [Rahm02]. Aus diesen Kriterien werden dann die unterschiedlichen Datenbankansätze der *parallelen*, *verteilten* oder *föderierten* Datenbanksysteme abgeleitet. Dem Data-Warehouse-Ansatz am nächsten kommt der föderierte Ansatz [Conr97], da dort ein neues konzeptuelles Schema gebildet wird, die Quellsysteme autonom bleiben und das originäre konzeptuelle Schema beibehalten wird. Die Unterschiede zum Data Warehouse liegen darin, dass das Data-Warehouse-Schema einem speziellen Analysezweck dient, die Daten redundant vorgehalten werden und kein schreibender Zugriff auf die Quellsysteme gefordert wird. Auf Ebene der physischen Verarbeitung der Daten liegen die Vorfahren von Data-Warehouse-Systemen im Bereich der *Statistical and Scientific Databases* (SSDB, [Mich91]), die ihren Schwerpunkt auf der Masendatenverarbeitung zur statistischen Analyse haben.

1.3 Einordnung und Abgrenzung von Business Intelligence

Ausgehend von der Diskussion der Historie kann und muss nach 20 Jahren des Begriffs »Data Warehouse« die Frage gestellt werden, ob nicht die Bezeichnung »Business Intelligence« geeigneter wäre.

Um darauf eine eindeutige Antwort zu geben, bedarf es auch hierbei einer einheitlichen und präzisen Definition. Stattdessen hat sich auch hier eine Vielfalt von Definitionen, Beschreibungen und Vorstellungen breit gemacht, die von unterschiedlichen Interessengruppen (Marktforscher, Wissenschaftler, Hersteller) geprägt wird. Auch in diesem Umfeld existieren bereits einige Konsolidierungsansätze wie z. B. in [KeMU07]. Dort wird definiert:

»Unter *Business Intelligence (BI)* wird ein integrierter, unternehmensspezifischer, IT-basierter Gesamtansatz zur betrieblichen Entscheidungsunterstützung verstanden.

- *BI-Werkzeuge dienen ausschließlich der Entwicklung von BI-Anwendungen.*
- *BI-Anwendungssysteme bilden Teilaspekte des BI-Gesamtansatzes ab.*« ([KeMU07])

Aus diesen und weiteren Ansätzen können zwei Tendenzen gegenüber dem Begriff »Data Warehouse« abgeleitet werden:

- Ausweitung des Themas Integration – neben Daten werden auch Strategien, Prozesse, Anwendungen und Technologien integriert.

- Ausweitung des Themas Analyse – neben der Analyse der Daten wird die Erzeugung von Wissen über Status, Potenziale und Perspektiven im Geschäftsumfeld gefordert.

Damit soll für dieses Buch als Abgrenzung festgehalten werden, dass das Data Warehouse mit seiner Definition, Architektur und den verwendeten Mechanismen weiter Bestand hat. Das Data Warehouse kann somit als Teil von Business Intelligence gesehen werden. Erst die Erweiterung der Integrationsmöglichkeiten und die Hinzunahme des Wissensmanagements führt zum Business Intelligence. Ein wesentlicher Aspekt des Business Intelligence ist Ausrichtung am betriebswirtschaftlichen Anwendungsfeld, indem es den analytischen Prozess, der Unternehmens- und Wettbewerbsdaten in handlungsgerechtes Wissen für die Unternehmenssteuerung überführt, darstellt. Dadurch kann Data Warehousing in die Business-Intelligence-Organisation eingebettet werden. In einigen Firmen wird dem bereits Rechnung getragen, indem Data Warehouse Competence Center zu BI Competence Center wurden.

1.4 Anwendungsbereiche

Die Triebfeder hinter dem Data-Warehouse-Gedanken ist die Analysierbarkeit der Daten: eine homogene und integrierte Datenbasis, geeignet aufbereitet, um effiziente und zielorientierte Analysen durchzuführen. Nahezu überall, wo Daten gespeichert werden, entsteht auch der Wunsch, diese auswerten zu können. Die Anwendungsgebiete sind entsprechend weit gestreut und reichen von betriebswirtschaftlichen Aufgabenstellungen über wissenschaftliche Analysen bis hin zu technischen Belangen. Kapitel 12 dieses Buches bietet einen Einblick in eine Auswahl von Data-Warehouse-Projekten und gibt somit einen detaillierteren Eindruck über Einsatzpotenziale von Data-Warehouse-Systemen.

Die häufigsten Einsatzfelder finden sich im betriebswirtschaftlichen Bereich. Nach einem kurzen Überblick sollen deshalb anhand dieses Bereichs die vier Anwendungsgebiete *Informationsbereitstellung*, *Analyse*, *Planung* und *Kampagnenmanagement* beschrieben werden. Grundlegendes Element in allen Einsatzbereichen ist immer die Informationslieferung. Mit den dargestellten Informationen können qualifizierte Anwender Analysen durchführen und weitergehende Erkenntnisse gewinnen. Planung schließt den betriebswirtschaftlichen Regelkreis aus Planung, Steuerung und Kontrolle. Planzahlen stellen neben den Istzahlen eine zweite wesentliche Kennzahlenkategorie im Data Warehouse dar. Kampagnenmanagement letztlich ist ein Beispiel einer Anwendung von Data-Warehouse-Technologie, die Informationsbereitstellung, Analyse und Planung integriert. Die skizzierten Anwendungsgebiete bauen mit steigender Komplexität und damit sinkenden Anwenderzahlen aufeinander auf.