

Universität Ulm

Abteilung Angewandte Informationsverarbeitung

Seminar Internetdienste

Business Intelligence

Bearbeitet von Claudia Schmiedeberg

Sommersemester 2004

Gliederung

1. Informationsmanagement als zentrales Problem der Unternehmensführung
 - 1.1. Information Overload
 - 1.2. Daten, Information & Wissen
 - 1.3. Kommunikation
2. Business Intelligence
 - 2.1. Data Warehouse
 - 2.2. Online Analytical Processing (OLAP)
 - 2.3. Data Mining
3. Business-Intelligence-Anwendungen im Intra- und Internet
 - 3.1. Anwendungen im Intranet
 - 3.2. Anwendungen im Internet
4. Nutzen von Business Intelligence
5. Literaturverzeichnis

1. Informationsmanagement als zentrales Problem der Unternehmensführung

1.1 Information Overload

In Zeiten rapide wachsender Speicher- und Verarbeitungskapazitäten der EDV-Systeme kommt es zu einem Informationsparadox: Daten gibt es im Überfluss, das Problem dagegen besteht in deren Analyse, ohne die das Unternehmen aus dem Datenreichtum keinen Nutzen ziehen kann. Je mehr Daten sich jedoch in einem Unternehmen anhäufen, desto größer ist die Gefahr, dass Daten übersehen werden. Mitarbeiter verschiedener Abteilungen pflegen separate Datenbanken, zwischen den Abteilungen findet ein ungenügender Informationsaustausch statt, so dass ein Teil der Daten an mehreren Stellen gespeichert wird – und dort jeweils gepflegt werden muss – und andererseits die Daten, die sich gegenseitig ergänzen könnten, nicht zueinander in Beziehung gesetzt werden, weil effiziente Zugriffsmöglichkeiten fehlen.

Ein fundamentales Problem im Informationsmanagement besteht in der Divergenz zwischen dem Angebot, dem objektiven Bedarf und dem (subjektiven) Bedürfnis an Information. Wenn aber Informationsangebot und Informationsbedarf nicht deckungsgleich sind, bedeutet das, dass ein Anwachsen der verfügbaren Information allein noch keinen Gewinn bringt, solange nicht genau die Informationen geliefert werden, die für die Entscheidung nötig sind, oder aber wenn der Entscheidende sich nicht bewusst ist, welche Informationen er tatsächlich benötigt. Gerade dadurch, dass menschliche Fähigkeit, Informationen aufzunehmen und zu verarbeiten, begrenzt ist, verringern zusätzliche Informationen die Entscheidungseffizienz statt sie zu erhöhen. Statt sich zuerst zu fragen, welche Aufgabe er zu lösen hat und welches Wissen dazu nötig ist, kämpft sich der Mitarbeiter durch Berge von Datenmüll, die Folge sind Stress und Konfusion. Immer wichtiger wird deshalb zur Bewältigung der Informationsflut die Strukturierung und Selektion der Daten. Hierin besteht die Aufgabe von Informationsmanagement: Informationsangebot und Informationsbedarf sollen aufeinander abgestimmt werden, sowie dem Mitarbeiter Instrumente zur Verfügung gestellt werden, die eine effizientere Verarbeitung der Information ermöglichen.

1.2 Daten, Information und Wissen

Einleitend war stets die Rede von Daten, Information und Wissen, ohne sie zu definieren und gegeneinander abzugrenzen. Häufig werden die Bezeichnungen irrtümlich synonym gebraucht. Denn wie vorausgehend bereits illustriert wurde, ist der Nutzen von operativen Daten für ein Unternehmen gering, solange sie nicht weiterverarbeitet werden. Deshalb sollen diese drei grundlegenden Termini hier kurz definiert werden:

Daten sind nach DIN 44300 zuerst einmal neutrale Zahlenwerte und Fakten, die maschinell verarbeitet werden können. Daten können in strukturierter, d.h. mit einer fest definierten Syntax, oder unstrukturierter Form, beispielsweise als Bilddateien oder Textdokumente, vorliegen. Unterstellt man ein Kontinuum von Daten über Information zu Wissen, stellen die Daten die kleinsten Bausteine des Wissens dar.

Information dagegen ist „zweckorientiertes Wissen“ (Ahlemeyer-Stubbe 2001), also die Antwort auf eine Frage. Je nach Perspektive können dabei allerdings unterschiedliche Aspekte im Vordergrund stehen. Etwa in der Nachrichtentechnik und der Informatik wird Information fast synonym mit Daten gebraucht, so dass alles, was – unabhängig von dessen Bedeutung – in codierter Form von Sender zu Empfänger übermittelt werden kann, Information darstellt (nach: Krempl 1997). Im Folgenden wird Information als die Verbindung zwischen Daten und Wissen betrachtet, d.h. als ein Aggregat von strukturierten Daten, die in einem bestimmten Kontext stehend und einen eindeutigen Inhalt transportieren, die bei der Problemlösung verwendet werden können, aber noch keine Problemlösekompetenz verschaffen. Eine sinnvolle Unterscheidung lässt sich auch zwischen primären und sekundären Informationen treffen; letztere werden aus ersteren durch Aggregation und Bearbeitung gewonnen, wodurch Kennzahlen entstehen wie etwa durchschnittliche Monatswerte, die auf den einzelnen Verkaufszahlen beruhen. Die Datenauswertung erfolgt überwiegend auf Basis von sekundären Daten, gespeichert werden meist sowohl primäre als auch sekundäre Daten. (vgl. Vetschera 1994, S. 5)

Wissen schließlich ist an den Menschen gebunden, es beinhaltet alle Kenntnisse und Erfahrungen einer Person oder einer Gruppe von Personen. Das bedeutet, dass das Wissen nicht problemlos in einer Datenbank gespeichert werden kann. Manche Autoren unterscheiden zwischen implizitem Wissen, das auf persönlichen Erfahrungen basiert und sich tatsächlich nur über direkten persönlichen Kontakt

weitergegeben werden kann, während explizites Wissen sich in Anleitungen und Dokumentationen festhalten lässt (nach: Richter 2001). Der Erwerb von Wissen erfolgt durch die logische Verknüpfung von Informationen, die im Hinblick auf die Lösung von Problemen erfolgt. Dies zeigt die Abhängigkeit jedes Wissens einerseits von der Datenbasis, aus der es gewonnen wird, andererseits von den Problemstellungen, die den Prozess der Wissensgewinnung erst in Gang setzen.

Den Zusammenhang zwischen Daten, Information und Wissen sowie den Transformationsprozess von Daten in entscheidungsrelevantes Wissen, der den Kern von Business Intelligence darstellt, macht folgendes Zitat nochmals deutlich: „Daten operativer Systeme und externer Anbieter werden integriert und in Informationen transformiert. Diese Informationen werden durch die Analyse der Anwender in Wissen umgewandelt.“ (Eberlein 2004)

1.3 Kommunikation

Auch Kommunikation ist ein relativ vieldeutiger und in den Wirtschaftswissenschaften oft verwendeter Begriff. Von Bedeutung ist im Kontext mit Business Intelligence nicht die externe Unternehmenskommunikation, die vor allem zu Marketingzwecken eingesetzt wird, sondern die interne Kommunikation, die dafür sorgt, dass die beispielsweise mittels Data Mining gewonnenen Informationen zu den Mitarbeitern transportiert werden, die sie benötigen. Während bisher also von Daten, Information und Wissen als Produkten die Rede war, muss Kommunikation als Prozess der Vermittlung von Daten, Information und Wissen verstanden werden. Kommunikation steht damit sowohl am Anfang als auch Ende des Business-Intelligence-Prozesses, indem durch sie die Problemstellungen in den Prozess eingebracht und schließlich die Ergebnisse im Unternehmen verbreitet und zur Anwendung gebracht werden.

Es sind zwei Methoden der Informationsversorgung denkbar: Die Pull-Strategie verlangt vom Mitarbeiter, sich die benötigten Informationen selbst zu beschaffen, etwa durch Internet-Recherche, Lesen von Informationsmaterial oder Anfrage an andere Personen. Die Push-Strategie hingegen bedeutet eine Bringschuld derjenigen, die die relevanten Informationen besitzen, so dass beispielsweise bei Gesetzänderungen alle betroffenen Abteilung von der internen Rechtsabteilung eines Unternehmens automatisch darüber informiert werden. Die Kommunikationsstruktur beinhaltet die Festlegung, auf welche Weise welche Informationen zirkulieren, d.h.

die Vergabe von Rechten und Pflichten bezüglich des Zugriffs und der Weitergabe von Informationen.

Nachdem interne Kommunikation lange Zeit vertikal stattfand, indem der Vorgesetzte den Mitarbeiter mit Information versorgte und der wiederum das Ergebnis seiner Arbeit nach oben meldete, findet heute Kommunikation auf und zwischen allen Hierarchieebenen statt. Voraussetzung dafür ist eine Unternehmenskultur, die auf die aktive, selbständige Mitarbeit aller setzt, sowie eine adäquate Informationsinfrastruktur, wie etwa den Datenaustausch über das Intranet. Im Idealfall führt eine optimierte interne Unternehmenskommunikation zu einer besseren Koordination der Arbeit und damit zu einer effizienteren Ressourcennutzung, zudem werden Zusammenarbeit, Leistungsbereitschaft und Verantwortungsgefühl gefördert. (nach: Senger 1999)

Die Verbreitung der durch Business-Intelligence-Tools aufbereiteten Informationen im Rahmen der internen Kommunikation kann durchaus als Teil des Business-Intelligence-Prozesses verstanden werden. Oft werden diese Funktionen allerdings unter dem Schlagwort Wissensmanagement subsumiert. Deshalb wird die Informationsverteilung in dieser Arbeit nicht zentral behandelt werden, sondern nur wo nötig Erwähnung finden.

2. Business Intelligence

Vor dem Hintergrund der Schwierigkeit, aus der Flut operativer Daten die relevanten Informationen herauszufiltern, sowie der Notwendigkeit von Effizienzsteigerungen durch einen immer stärker werdenden Konkurrenzdruck auf den globalisierten Märkten kommen komplexe Softwarelösungen zur Anwendung, die Speicherung und Analyse der Geschäftsdaten unterstützen sollen. Als Begriff dafür wurde der Terminus Business Intelligence 1993 von der Gartner Group geprägt, der ein breites Spektrum von Anwendungen und Technologien umfasst. Die wichtigsten Elemente sind Data Warehouse, OLAP und Data Mining, die im Folgenden vorgestellt werden sollen. (nach: Eberlein 2004)

Der Business-Intelligence-Prozess lässt sich in drei Phasen gliedern:

1. Bereitstellung der Basisdaten, Integration der Daten z.B. im Data Warehouse,
2. Auswertung der Daten, etwa durch OLAP, Data Mining oder Web Mining,
3. Verbreitung und Anwendung der Ergebnisse. (vgl. Preuschoff 2002, S. 9)

Wichtig ist die Integration aller Prozessschritte, um Insellösungen zu vermeiden, sowie deren Verankerung in der Unternehmensstrategie.

2.1 Data Warehouse

Das Data Warehouse stellt die Grundlage für jedes Business-Intelligence-System dar, wo alle relevanten Daten in einer zentralen Datenbank langfristig gespeichert werden. Typischerweise bestehen diese Datenbanken parallel zu den operativen Systemen im Produktivbetrieb und verfügen über eine komplexe multidimensionale Struktur, in der die Heterogenität der Datenbestände abgebildet ist. In einer Definition von Inmon (1996, S. 33) werden die Kennzeichen der Data-Warehouse-Architektur zusammengefasst als „subject oriented, integrated, non-volatile, time-variant“: Im Gegensatz zu den operativen Datenbanken orientiert sich die Datenstruktur im Data Warehouse an unternehmensrelevanten (Geschäfts-)Objekten (subject oriented) statt an Geschäftsprozessen; beispielsweise können als (Objekt-)Dimensionen die Kunden- oder Produktstruktur gewählt werden. Zudem werden sowohl historische als auch neue Daten gespeichert, die jeweils durch einen Zeitstempel gekennzeichnet sind (time-variant). Statt sie bei Veränderungen zu überschreiben oder zu löschen, werden alle historischen Daten gespeichert (non-volatile), so dass beispielsweise die Bonität eines Kunden im Zeitverlauf betrachtet werden kann. Da sie aus den verschiedensten internen wie externen Quellen zusammengeführt werden, müssen die Daten vor der Integration im Data Warehouse geordnet, bereinigt und auf Redundanz und Inkonsistenzen geprüft werden (integrated). Gerade wenn in einem Unternehmen nach Jahren unkontrollierten Wachstums die DV-Strukturen wenig einheitlich sind, gibt es eine Vielzahl an nicht kompatiblen Dateiformaten, unterschiedlichen Datenstrukturen und redundante Datenbeständen, die in einer großen Datenbank integriert werden müssen.

Entscheidend für jede Business-Intelligence-Lösung ist die Modellierung der Datenstrukturen, die im Data Warehouse gespeichert werden sollen, was sowohl Wirtschafts- als auch IT-Kenntnisse erfordert. Wichtig ist dabei, dass von Anfang an auf die zentralen Fragestellungen der Geschäftsführung und aller beteiligten Abteilungen geachtet wird, so dass die Planung des Data Warehouses daher alle Abteilungen involviert, deren Daten darin zur Analyse vorgehalten werden sollen. Definiert werden die Kennzahlen und Kennzahlengruppen, die Dimensionen sowie

die Aggregationsstufen, d.h. erstens welche Kennzahlen sollen gespeichert werden, zweitens in welche Einheiten – wie Produktgruppen, Absatzmärkte oder Zeiteinheiten – sollen sie zusammengefasst werden können, und in welchen Schritten sollen sie sich aggregieren lassen. Die zeitliche Dimension kann beispielsweise in Tage, Wochen, Monate und Jahre gegliedert sein. Als Werkzeuge stehen verschiedene Modellierungssprachen zur Verfügung, insbesondere die multidimensionale Erweiterung des Entity-Relationship-Modells (EM/R), Application Design for Analytical Processing Technologies (ADAPT) sowie Unified-Modelling-Language (UML). (nach: Schmidt-Thieme 2003)

In vielen Fällen bietet sich der Aufbau der Data-Warehouse-Architektur „von unten“ an: Zusätzlich zum unternehmensweiten Data Warehouse werden in einem ersten Schritt sogenannte Data Marts z.B. für die einzelnen Abteilungen kreiert, die nur die für die entsprechende Abteilung relevanten Informationen enthalten und daher über eine wesentlich geringere Komplexität verfügen. Aus diesen dezentralen Einheiten werden die Daten, die bereits einen ersten Aufbereitungsprozess durchlaufen haben, an das Data Warehouse weitergegeben. (nach: Wolf 2003)

2.2 Online Analytical Processing (OLAP)

Nachdem das Data Warehouse steht, ist das nächste Problem, sich in dessen komplexer Datenstruktur zurechtzufinden. Hier kommt OLAP zur Anwendung, ein Instrument zur Datenanalyse und Berichterstellung, das eine Weiterentwicklung der herkömmlichen Datenbankabfragen darstellt. Anwender von OLAP sind nicht Datenbankspezialisten, sondern hauptsächlich Mitarbeiter von Marketing- und Controllingabteilungen mit eher geringen IT-Kenntnissen, deren Aufgabe die Erstellung von ad-hoc-Auswertungen mit komplexen Charakter ist, was sowohl eine intuitive Bedienbarkeit, als auch Vielseitigkeit und Flexibilität des Systems voraussetzt.

Um die Anforderungen an OLAP-Systeme zu standardisieren, wurden 1993 die 12 Regeln des Dr. Codd (sozusagen des Vaters der relationalen Datenbank) sowie 1995 die Definition FASMI publiziert. In dem Kunstwort FASMI fasst der Olap-Report die fünf grundlegenden Eigenschaften von OLAP zusammen: Die Systeme sollen schnell (*fast*) sein, d.h. auch in großen Datenmengen in wenigen Sekunden Abfragen durchführen können, sie sollen statistische Analysen unterstützen (*analysis*), multi-

user-fähig sein und dabei gewisse Sicherheitsstandards berücksichtigen (*shared*), und natürlich in multidimensionaler Form (*multidimensional*) die Informationen darstellen (*information*). (nach: Pendse 2004)

Charakteristisch für OLAP ist also, dass aus der (im Allgemeinen relationalen) Struktur der Datenbank eine Darstellung in mehrdimensionalen Sichten generiert wird. Diese Multidimensionalität ist als Datenwürfel vorstellbar, der z.B. Zeit, Region und Produkt repräsentiert. Die einzelnen Dimensionen können wiederum in Hierarchien gestaffelt sein, etwa in Produktgruppen und -linien oder die verschiedenen regionalen Ebenen Stadt, Landkreis, Bundesland. Gerade solche hierarchischen Strukturen sind geeignet, um große Datenvolumen beherrschbar zu machen, weil bei der Suche nach Informationen vom Groben zum Feinen vorgegangen werden kann. Innerhalb des Datenwürfels kann der Benutzer navigieren, etwa indem er die Aggregationsebene wechselt (roll-up bzw. drill-down), um mehr oder weniger detaillierte Informationen zu erhalten, oder das betrachtete Datenvolumen reduziert: die Funktion „slice“ ermöglicht es, eine sozusagen Scheibe des Würfels abzuschneiden, um etwa nur ein bestimmtes Produkt oder ein Geschäftsjahr zu betrachten, „dice“ reduziert die Daten in mehreren Dimensionen, so dass beispielsweise nur noch der Absatz eines Produkts im vergangenen Monat in Süddeutschland dargestellt wird. Die Funktion „rotate“ verändert nur die Ansicht des Würfels.

Auf dem Markt wird eine Vielzahl von OLAP-Standardlösungen angeboten, häufig findet man auch die Einbindung in handelsübliche Tabellenkalkulationsprogramme (die hier nicht genannt werden sollen). Das Anwendungsspektrum ist vielseitig; es lassen sich sowohl die relevanten Daten für die Planung von Marketingaktionen abrufen, beispielsweise indem ein bestimmtes Kundenprofil erstellt wird, auf das eine Direkt Mail-Aktion abzielt, als auch Produktionsprozesse analysieren oder die Entwicklung der Auftragsdaten für die Vertriebsplanung berechnen.

2.3 Data Mining

Auch Data Mining bezeichnet eine Form der Datenanalyse auf der Basis des Data Warehouse, doch ist die Zielsetzung eine andere als bei OLAP: Durch verschiedene, vorwiegend statistische Verfahren sollen unerwartete Zusammenhänge zwischen den Daten aufgedeckt werden; statt konkreter Fragen existiert als Ausgangspunkt

nur das Ziel der Analyse, das die Auswahl der Instrumente und Techniken beeinflusst.

Das Vorgehen im Data-Mining-Prozess lässt sich in 7 Schritte einteilen: 1. Zielanalyse, 2. Datenauswahl, 3. Datenaufbereitung, 4. Auswahl der Analysetools, 5. Anwendung der Data-Mining-Methoden, 6. Interpretation und Evaluation der Ergebnisse und 7. Ergebnisimplementation. Ein besonderes Augenmerk liegt bei der Auswahl und Aufbereitung der zugrundeliegenden Daten, die circa 60% der Arbeitszeit beanspruchen. (nach: Wilde) Schließlich kann jede Statistik und Analyse nur so gut – sprich: auf die Realität zutreffend – sein wie ihre Datenbasis. Ausgehend von der Problemstellung, die im ersten Schritt erarbeitet wurde, werden deshalb die relevanten Datensätze ausgewählt, Fehler und Lücken entfernt sowie die interessierenden Variablen identifiziert. Einerseits soll versucht werden, keine überflüssigen Daten zu verarbeiten, andererseits lassen sich Schätzfehler und „voreilige Schlüsse“ umso eher vermeiden, je mehr Datensätze verwendet werden, weshalb Data-Mining-Software enorme Datenvolumen bewältigen können muss. Liegen die aufbereiteten Daten vor, wird durch die Auswahl der Analysemethoden und die Modellerstellung festgelegt, wie das eigentliche Data Mining ablaufen wird. Es gibt dafür ein breites Spektrum an Methoden und Techniken, von denen einige kurz vorgestellt werden sollen:

Cluster Analysis: Es wird versucht, die Daten zu klassifizieren, d.h. Gruppen zu bilden, die gleiche Merkmale aufweisen. Dieses Verfahren ist relativ einfach in der Anwendung für alle Arten von Variablen, die Schwierigkeit liegt allerdings in der Interpretation der Cluster. Ein typisches Anwendungsbeispiel ist etwa die Einteilung von Kunden gemäß ihrer Kaufkraft und Bonität.

Lineare Regression: Dieses recht simple statistische Verfahren versucht, unter Verwendung von Standardabweichung und Erwartungswerten eine Variable in Abhängigkeit einer anderen Variablen abzubilden. Allerdings funktioniert die Regression nur mit metrischen Parametern und bei linearen Abhängigkeiten.

Entscheidungsbäume: Die verwendeten Algorithmen, z.B. CHAID (Chi Square Automatic Interaction Detection) oder CART (Classification and Regression Tree), basieren auf der Abbildung aller möglichen Ergebnisse der Parameter in einer Baumstruktur. Für jeden Teilbaum wird dann die unabhängige Variable mit dem größten Einfluss auf das Ergebnis des Teilbaums identifiziert. Mit steigender Anzahl

der zu betrachtenden Variablen oder deren möglichen Ausprägungen wächst die Komplexität des Baumes und damit der Rechenaufwand, weshalb sich das Verfahren besonders für binäre Zielvariablen eignet. (nach: Trumfheller 2003)

Neuronale Netze: Die Daten werden als Netz unabhängiger Neuronen modelliert, mit dem auch unbekannte oder nicht-lineare Zusammenhänge erkannt werden können. Typisch ist die Lernfähigkeit neuronaler Netze, wodurch eine Klassifikation von Daten anhand von Trainingsbeispielen „erlernt“ wird, ohne dass mathematische Regeln notwendig sind. Das Problem ist allerdings, dass oft eine sehr große Menge an Trainingseinheiten benötigt wird, um sinnvolle Resultate zu erzielen. Als weiteres Problem stellt sich die Interpretation der Ergebnisse, da das neuronale Netz wie eine Black Box funktioniert und so nur schwer Aussagen über die Qualität des gefundenen Musters gemacht werden können.

Assoziationsanalyse: Hierzu gehört insbesondere die klassische Warenkorbanalyse, die das Käuferverhalten in Form von Wenn-dann-Aussagen einzuordnen versucht. Ergebnis sind Aussagen wie „Wenn Kunden Bier kaufen, kaufen sie mit einer Wahrscheinlichkeit von 60% auch Chips.“ Mittels Support und Konfidenz kann die statistische Signifikanz der gefundenen Korrelationen von vornherein festgelegt werden.

Welche Ergebnisse das Data Mining liefert, hängt stark von den hier getroffenen Entscheidungen ab, denn die Zahl der theoretisch vorhandenen Muster ist groß, und weder können durch Data Mining alle Muster gefunden werden, noch ist jedes Muster sinnvoll und verwertbar. Deshalb müssen die Ergebnisse interpretiert und bezüglich ihrer Relevanz und Stichhaltigkeit geprüft werden. Hierfür bietet sich, wie auch schon bei der Auswahl und Aufbereitung der Daten, die Anwendung von OLAP an.

Data Mining hat sich als ein schlagkräftiges Instrument vor allem im Marketingbereich erwiesen, das die Personalisierung von Vertriebs- und Marketingaktionen ermöglicht. In Kundendatenbanken werden eine Vielzahl von Kaufinformationen, soziographischen Daten und Kommunikationsdaten gespeichert, die im Bedarfsfall Vorhersagen über das Potential der Kunden und die jeweils optimale Marketingstrategie liefern. Problematisch ist dabei der Umgang mit den Daten zu sehen. Es stellt sich die Frage, ob der gläserne Kunde sich mit den Grundsätzen des Datenschutzes im Einklang steht.

3. Business-Intelligence-Anwendungen im Intra- und Internet

3.1 Anwendungen im Intranet

Das Intranet hat in den letzten Jahren einen gewaltigen Aufschwung erfahren und ist damit zum bedeutendsten Instrument der internen Unternehmenskommunikation avanciert. Erfüllt werden Funktionen sowohl Kommunikations- wie auch des Informationsmanagements, wobei letzteres im Zusammenhang mit Business Intelligence im Vordergrund steht. Werden bereits vorhandene Anwendungen, wie etwa Datenbanken und Berichtssysteme, im Intranet integriert, erhalten die Mitarbeiter einfach und schnell Zugriff auf die Geschäftsdaten über eine einheitliche Bedienoberfläche. (nach: Senger 1999)

Dies optimiert einerseits den Informationsfluss im gesamten Unternehmen, dadurch dass der Zugriff auf die Unternehmensdaten für alle Mitarbeiter transparent wird. Es können nicht mehr nur wenige Spezialisten bestimmter Abteilungen die Daten abfragen, die zum Informationsengpass werden können. Vor allem wird die räumliche Unabhängigkeit der Mitarbeiter realisiert, da die Verfügbarkeit auch bei einer kurzfristigen Verlagerung des Arbeitsplatzes stets garantiert ist. Erheblichen Nutzen können beispielsweise Außendienstmitarbeiter aus Intranetlösungen ziehen, etwa wenn sie während eines Kundengesprächs die aktuellen Daten präsentieren oder Produktbestände und Lieferzeiten abfragen können, ebenso wird Telearbeit und die Vernetzung der Mitarbeiter auch an mehreren Unternehmensstandorten gefördert. Vielversprechende Anwendungsmöglichkeiten ergeben sich durch die Nutzung mobiler Endgeräte wie Palm oder Handy, die die Informationsversorgung außerhalb des Arbeitsplatzes sicher stellen. So könnten beispielsweise Kundendienstmitarbeiter die Verfügbarkeit von Ersatzteilen abprüfen oder Ärzte Patientendaten direkt im Krankenzimmer abfragen oder in die Datenbank eingeben; denkbar ist auch das Versenden automatisch generierter Nachrichten, wenn etwa bestimmte Abweichungen vom Sollwert festgestellt werden. Dabei sollte die zu übertragende Datenmenge auf das verwendete Medium angepasst werden, da lange Antwortzeiten die Akzeptanz von Anwenderseite vermindern. Statt hochverdichtete Daten oder Grafiken zu versenden, sollten die Informationen auf kleine Einheiten reduziert werden.

Da natürlich nicht jeder Mitarbeiter auf alle Daten zugreifen können soll, muss genau geregelt sein, welche Zugriffsmöglichkeiten bestehen sollen. Hierzu ist eine Personalisierung des Intranets nötig, d.h. die Mitarbeiter melden sich mit ihrem Benutzernamen an, dem bestimmte Zugriffsrechte zugeordnet sind. Innerhalb dieses personalisierten Bereichs können auch vertrauliche Informationen bereitgestellt werden, etwa die eigenen Personaldaten wie Fehlzeiten, Betragssätze zur Altersvorsorge und ähnliches. Zudem besteht bei manchen Lösungen für den Einzelnen die Möglichkeit, die Oberfläche an seine Anforderungen anzupassen und Informationsabonnements zu nutzen, so dass automatisch die Informationen bereitgestellt werden, die der Mitarbeiter am häufigsten benötigt. Dabei beschränkt sich der Inhalt nicht ausschließlich auf unternehmensinterne Informationen, sondern es können auch externe Dienste wie etwa Börsenticker oder Wetterberichte integriert werden. Ziel ist die Erleichterung der Informationsbeschaffung durch eine Bereitstellung aller verfügbaren Quellen. Hier treffen sich Business Intelligence und Knowledge Management.

3.2 Anwendungen im Internet

Die Anwendung von Business Intelligence im Internet ist auf zwei Arten möglich: entweder werden Daten aus dem Internet verarbeitet und die Erkenntnisse intern weiterverwendet oder es wird via Internet der Zugriff auf Daten des Data Warehouse ermöglicht.

Letzteres praktizieren beispielsweise einige Banken, die für ihre Kunden online detaillierte Informationen über die Entwicklung ihrer Aktienfonds bereitstellt. Ebenso können derartige Anwendungen die Zusammenarbeit mit den Zulieferern verbessern, indem diese Zugriff auf die Auftragslage erhalten und so frühzeitig darauf reagieren können. Problematisch sind dabei häufig Sicherheitsaspekte, da der Zugriff auf die Datenbank kontrolliert ablaufen muss. Eine Lösung kann hier die Schaffung eines Replikats des Data Warehouse auf einem separaten Server darstellen, wo, ähnlich wie in einem Data Mart, nur der für die Kunden zugängliche Teil der Daten vorgehalten wird. Manipulationen an den Daten des originären Data Warehouse werden so ausgeschlossen. Viele Unternehmen sind allerdings an mehr Transparenz nur wenig interessiert, weshalb die Entwicklung derartiger Systeme eher langsam voranschreitet. (nach: Fritsch 2002)

Recht vielversprechend ist das sogenannte Web Mining, die Anwendung von Data-Mining-Techniken im Internet, bei dem versucht wird, aus den Daten, die die Benutzer einer Webseite hinterlassen haben, Konsumentenprofile und Verhaltensmuster herauszufinden. (nach: Schmidt 2003) Web Mining lässt sich wiederum unterteilen in Web Usage Mining und Web Content Mining, je nachdem, ob die Nutzung der Internetseiten oder deren Inhalt im Vordergrund der Analyse stehen.

Web Usage Mining stützt sich größtenteils auf die Logfiles, die bei jedem Besuch der Internetseite erstellt werden und das Verhalten des Anwenders dokumentieren. So kann der Anbieter eines Online-Shops Rückschlüsse auf die Beliebtheit seiner Produkte ziehen, indem er analysiert, welche Produkte ein Kunde wie lange betrachtet hat, um sie am Ende zu kaufen oder eben auch nicht. Ein Produkt, das regelmäßig lange betrachtet aber relativ selten gekauft wird, ist beispielsweise tendenziell zu teuer. Werden, z.B. bei einem Kauf, zusätzlich persönliche Daten abgefragt, kann das Kundenverhalten mit den erhaltenen soziodemografischen Daten in Zusammenhang gebracht werden und so weiteren Aufschluss geben.

Bei *Web Content Mining* sind der Aufbau und Inhalt von Websites oder die Linkstruktur zwischen den einzelnen Seiten Gegenstand der Betrachtung. Letzteres wird teilweise auch als *Web Structure Mining* bezeichnet. Mit den dort gewonnenen Ergebnissen können Rückschlüsse auf die Interessen der Kunden gezogen und die Internetseiten dementsprechend optimiert werden. Wenn beispielsweise eine Assoziationsanalyse ergibt, dass bestimmte Seiten häufig nacheinander aufgerufen werden, kann durch einen entsprechenden Link den Usern die Navigation erleichtert werden.

Interessant im Hinblick auf Daten aus dem Internet ist auch die Möglichkeit, mit Data Mining nicht nur Zahlenwerte, sondern auch textgebundene Informationen, z.B. Kundenbriefe, Forschungsberichte oder Emails zu analysieren, das sogenannte Text Mining. (nach: Münch 2004) Dadurch können beispielsweise automatisch Abstracts generiert, die wichtigsten Schlagworte, die Aufschluss über den Inhalt des Textes geben, identifiziert oder auch Beziehungen zwischen Dokumenten entdeckt werden. Die typische Anwendung dieser Techniken ist das Auffinden und die Selektion relevanter Dokumente mit Hilfe von Suchmaschinen. Die Anwendungsmöglichkeiten sind hier nahezu unbegrenzt.

4. Nutzen von Business Intelligence

Business Intelligence ist derzeit in Mode: Manager versprechen sich davon verschiedenste Vorteile, von einer erhöhten Schlagkraft ihrer Marketingaktionen oder einer besseren Kundenbindung bis zur Reduzierung der Produktionskosten oder ähnlichem. Andererseits geht aus einer Studie hervor, dass viele Unternehmen keine exakte Vorstellung davon zu haben scheinen, was Business Intelligence genau ist, und worin sein Nutzen besteht. (nach: Frank 2002) Es lohnt sich also die Frage zu stellen, worin der Nutzen von Business Intelligence tatsächlich liegt.

Data Mining wird in Deutschland Großteils vom stark kennzahlenorientierten Controlling genutzt. Durch die Entdeckung relevanter Muster können Abhängigkeiten und Entwicklungspotentiale entdeckt oder Erklärungen für Soll-Ist-Abweichungen gefunden werden. In diesem Zusammenhang kommt auch OLAP zum Einsatz, das die Generierung von Berichten, Statistiken und Grafiken ermöglicht.

Der Anwendungsbereich von Web Mining ist dagegen hauptsächlich im Marketingbereich zu sehen, dadurch dass eine Verbesserung der Kundenorientierung im E-Business erreicht werden kann. Die wichtigsten Ziele von Web Mining sind dementsprechend die Gewinnung von Kundeninformationen und, darauf aufbauend, die Personalisierung des Internetportals.

Der Nutzwert von Business Intelligence ist grundsätzlich auf dem Gebiet der Informationsversorgung zu sehen, die in mehreren Hinsichten optimiert werden kann: Die automatische Integration und Aufbereitung der Daten führt zu einer höheren Aktualität der Daten sowie zu einer größeren Breite der Datenbasis, zudem garantieren die Analyseinstrumente eine verbesserte Informationsselektion, sowie tendenziell höhere Qualität und Flexibilität in der Auswertung und der Darstellung der Ergebnisse. (nach: Vetschera 1994, S. 18) Business Intelligence kann also durchaus ein Hilfsmittel zur besseren Bewältigung des eingangs erwähnten Information Overload sein, dadurch dass die Datenhaltung, -aufbereitung und -analyse zu einem strukturierten Prozess gemacht und so die Datenflut auf wenige, aber dafür die gewünschten Informationen reduziert wird, die sich in einer übersichtlichen Darstellung präsentieren. Allerdings muss der Anwender mit den Business-Intelligence-Technologien auch umgehen können. Sicher wachsen die Anforderungen bezüglich Abstraktionsvermögen und logischem Denken durch den Einsatz derartiger Methoden. Business Intelligence kann nicht den denkenden

Mitarbeiter ersetzen, der das Ziel der Analyse vorgibt: „Der Einsatz von Business Intelligence beginnt nicht mit der Technik, sondern mit der Frage, die man beantwortet haben will.“ (Münch 2004)

Ebenso entscheidend ist die Integration der Business-Intelligence-Lösung in die gewachsene Infrastruktur. Nur wenn die Ergebnisse der Analysen an die entscheidenden Stellen gelangen und dort zur Anwendung kommen, hat sich der Aufwand gelohnt. Dazu müssen die Mitarbeiter in den Prozess eingebunden sein und die Technologien akzeptieren. Insofern ist der Erfolg von Business Intelligence abhängig vom Funktionieren der internen Unternehmenskommunikation und des Wissensmanagement.

5. Literaturverzeichnis

Ahlemeyer-Stubbe, A.: Data Warehousing, IBAI-Report 7/2001, URL:
<http://www.ahlemeyer-stubbe.de/artikel/ibai0700.htm>

Eberlein, Rüdiger: Mehrwert für Unternehmen mit Business Intelligence, 2004, URL:
<http://www.sdm.de/de/it-wissen/themen/bi/>

Frank, Mathias: Status Quo von Business Intelligence 2002, Köln 2002, URL:
<http://www.businessvillage.de/shop/eDocs/Detail/eb-448.html>

Fritsch, Werner: Business Intelligence im Extranet, in: Informationweek, 5/2002, URL:
<http://www.informationweek.de/index.php3?/channels/channel11/021044.htm>

Grotheer, Manfred, Hebben, Hermann: Schrittliste zum Aufbau einer OLAP-Datenbank, in: Controller Magazin 3/1999, URL: <http://www.my-controlling.de/aufsaeetze/olap/olap.htm>

Inmon, William H.: Building the data warehouse, 2. Auflage, New York 1996.

Krempf, Stefan: Was ist Information?, 04/1997, URL: <http://viadrina.euv-frankfurt-o.de/~sk/SoSe97/infosoc/DefInfo.html>

Münch, Vera: Data Mining im Bankkonto, in: c't aktuell, 27.04.2004, URL:
<http://www.heise.de/ct/aktuell/meldung/46880>

Pendse, Nigel: What is OLAP?, 5/2004, URL: <http://www.olapreport.com/fasmi.htm>

Preuschoff, Sarah: Business Intelligence – Gegenstand, Ansätze und Technologien, Stuttgart 2002, URL: [http://www.iuk.hdm-stuttgart.de/nohr/Km/KmAP/BusinessIntelligence .pdf](http://www.iuk.hdm-stuttgart.de/nohr/Km/KmAP/BusinessIntelligence.pdf)

Richter, Wolfgang: Virtual Communities und Customer Relationship Management, Graz 2001, URL: <http://www.iicm.edu/wrichter/thesis-final/node72.html>

Schmidt, Artur P.: Überleben im digitalen Zeitalter. Strategien, Technologien und Innovationen für den neuen Aufschwung, 2. Aufl., Graz 2003, URL:
<http://www.tzw.biz/pdf/3901402330.pdf>

Schmidt-Thieme, Lars: E-Business, Karlsruhe 2003, URL: <http://www.informatik.uni-freiburg.de/cgnm/lehre/eb-03s/eb9.pdf>

Senger, Joachim: Intranet als Produktionsfaktor?, Berlin 1999, URL: http://www.unrast.de/cms/dokumente/10000654/08612c80/intranet_als_produktionsfaktor.pdf

Trumpfheller, Jürgen: Data Mining in der Versicherungswissenschaft. Ausgewählte Ergebnisse, Hannover 2003, URL: http://www.iwi.uni-hannover.de/lv/do_ss03/do-11.pdf

Vetschera, Rudolf, u.a.: Die Entwicklung und Bewertung managementunterstützender Informationssysteme, Konstanz 1994.

Wilde, Klaus D.: Data Warehouse, OLAP und Data Mining im Marketing – Moderne Informationstechniken im Zusammenspiel. URL: http://www.kueichstaett.de/Fakultaeten/WWF/Lehrstuehle/WI/Lehre/dm_v/Sections/content/DM%201.pdf

Wolf, Ilse & Rudolf: Mit Business Intelligence erfolgreicher im Wettbewerb bestehen, in: MONITOR, 5/2003, URL: <http://www.monitor.co.at/index.cfm?storyid=5741>