



FAKULTÄT FÜR **INFORMATIK**

# Fuzzy-Logik für Data Mining im CRM

DIPLOMARBEIT

zur Erlangung des akademischen Grades

**Mag. rer. soc. oec.**

im Rahmen des Studiums

**Wirtschaftsinformatik (175)**

eingereicht von

**Cenk Kurter**

Matrikelnummer 9826240

an der  
Fakultät für Informatik der Technischen Universität Wien

Betreuung:  
*Betreuer:* Univ.-Prof. Dipl.-Ing. Mag. Dr. Dr. Gerald Quirchmayr

Wien, 26.11.2008

\_\_\_\_\_  
(Unterschrift Verfasser)

\_\_\_\_\_  
(Unterschrift Betreuer)



# Fuzzy-Logik für Data Mining im CRM

## Art der Arbeit

Diplomarbeit, vorgelegt zur Erlangung des akademischen Grades  
eines Diplomingenieurs der Studienrichtung Wirtschaftsinformatik (E-175)  
an der Technischen Universität Wien

## verfasst von

Cenk KURTER

Matrikelnummer: 9826240

E-Mail: [cenkkurter@hotmail.com](mailto:cenkkurter@hotmail.com)

## betreut durch

Univ.-Prof. Dipl.-Ing. Mag. Dr. Dr. Gerald Quirchmayr

Universität Wien

Fakultät für Informatik

Distributed and Multimedia Systems

Multimedia Information Systems

A-1010 Wien • Liebiggasse 4/3-4

### **-Selbstständigkeitserklärung-**

Hiermit erkläre ich, die vorliegende Diplomarbeit mit dem Thema „Fuzzy-Logik für Data Mining im CRM“ selbstständig angefertigt zu haben.

Ich versichere, dass ich keine anderen als die im Literaturverzeichnis angeführten Quellen und Hilfsmittel benutzt, und alle von anderen Verfassern wörtlich (Zitate) oder sinngemäß (Paraphrasen) übernommenen Stellen bzw. Textausschnitte als solche kenntlich gemacht, sowie die dazugehörigen Quellen entsprechend angegeben habe.

Wien, 26.11.2008

---

(Unterschrift Verfasser)

**-Danksagung-**

An dieser Stelle möchte mich bei all jenen Personen herzlich bedanken, die zur Realisierung dieser Diplomarbeit beigetragen haben. In erster Linie gilt dieser Dank dem Herrn Univ.-Prof. Dipl.-Ing. Mag. Dr. Dr. Gerald Quirchmayr. Ich bedanke mich bei ihm für seine Betreuung und seine sowohl wissenschaftliche als auch menschliche Unterstützung bei meiner Diplomarbeit. Dank seinem Verständnis und seinem Vertrauen in meine Arbeit gelang es mir, meine Diplomarbeit fertig zu stellen.

Ein weiterer Dank gilt meinem Studentenkollegen Herrn Senol Ari, der mir bei der Erstellung dieser Arbeit immer mit Rat und Tat zur Seite gestanden ist.

Weiters bedanke ich mich bei Frau Elfriede Schmitt für ihren Einsatz beim Korrekturlesen. Nicht zuletzt möchte ich mich bei meinen Eltern bedanken, die mich während meines Studiums stets kompromisslos unterstützten und mir ein Studium ermöglichten.

# INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS .....	9
TABELLENVERZEICHNIS .....	11
BILDVERZEICHNIS .....	12
SUMMARY .....	13
ZUSAMMENFASSUNG .....	15
<b>KAPITEL 1.....</b>	<b>1</b>
<b>EINLEITUNG .....</b>	<b>1</b>
1.1 ZIELSETZUNG UND VORGEHEN .....	3
<b>KAPITEL 2 .....</b>	<b>5</b>
<b>BEGRIFFLICHE ABGRENZUNG .....</b>	<b>5</b>
2.1 KUNDENVERHALTEN IM INTERNETZEITALTER .....	5
2.2 TRADITIONELLE MARKETING vs. ONE-TO-ONE MARKETING .....	7
2.3 WACHSENDE BEDEUTUNG DES KUNDEN FÜR UNTERNEHMEN .....	10
2.4 KUNDENORIENTIERUNG UND -ZUFRIEDENHEIT .....	10
2.5 KUNDENBINDUNG.....	12
2.6 KUNDENBEZIEHUNG .....	13
2.7 WISSENSMANAGEMENT – DIE KUNDENDATEN, DIE KUNDENINFORMATIONEN UND DEREN ANALYSE .....	14
<b>KAPITEL 3 .....</b>	<b>17</b>
<b>CUSTOMER RELATIONSHIP MANAGEMENT (CRM).....</b>	<b>17</b>
3.1 EINFÜHRUNG .....	17
3.2 GRÜNDE FÜR DIE ENTWICKLUNG DES CRM .....	19
3.3 WAS IST CRM EIGENTLICH? .....	22
3.4 GRUNDLAGEN UND ABGRENZUNG DES CRM-ANSATZES .....	24
3.5 ZIELE DES CUSTOMER RELATIONSHIP MANagements.....	33
3.6 KOMPONENTEN UND AUFGABENBEREICHE EINES CRM-SYSTEMS .....	40
3.6.1 Operatives CRM (oCRM) .....	44
3.6.1.1 Marketing-Automation .....	45
3.6.1.2 Sales-Automation.....	51
3.6.1.3 Service-Automation.....	55

3.6.2 Analytisches CRM (aCRM) .....	57
3.6.2.1 Online Analytical Processing .....	60
3.6.2.2 Data Mining .....	63
3.6.3 Kommunikatives CRM (kCRM) .....	66
3.6.3.1 Customer Interaction Center .....	67
3.6.3.2 Internet .....	69
3.7 WETTBEWERBSVORTEILE DURCH CRM .....	71
3.8 ERFOLGSFAKTOREN IM CRM .....	74
<b>KAPITEL 4 .....</b>	<b>81</b>
<b>KNOWLEDGE MANAGEMENT (KM) .....</b>	<b>81</b>
4.1 EINFÜHRUNG .....	81
4.2 DEFINITIONEN UND EIGENSCHAFTEN DES WISSENS .....	82
4.2.2 <i>Wie entsteht Wissen?</i> .....	85
4.2.3 <i>Formen von Wissen</i> .....	88
4.3 WAS IST WISSENSMANAGEMENT? .....	95
4.3.1 <i>Definitionen</i> .....	95
4.3.2 <i>Gestaltungsdimensionen des Wissensmanagements</i> .....	97
4.3.3 <i>Bausteine des Wissensmanagements</i> .....	101
4.4 WISSENSMANAGEMENT IM CRM .....	106
4.5 COMPUTERUNTERSTÜTZUNG IM MANAGEMENT .....	113
4.6 BUSINESS INTELLIGENCE .....	118
4.6.1 <i>Das Data Warehouse-Konzept</i> .....	122
4.6.2 <i>Database Marketing im CRM</i> .....	133
<b>KAPITEL 5 .....</b>	<b>140</b>
<b>DER EINFLUSS VON DATA MINING AUF CRM .....</b>	<b>140</b>
5.1 EINFÜHRUNG .....	140
5.2 WAS BEDEUTET KNOWLEDGE DISCOVERY IN DATABASES? .....	141
5.2.1 <i>Evolution des Data Mining</i> .....	143
5.2.2 <i>Data Mining und herkömmliche Statistik - Parellellitäten und Unterschiede</i> .....	144
5.2.3 <i>Verallgemeinerter Data Mining-Prozessablauf</i> .....	149
5.2.4 <i>Aufgaben des Data Mining</i> .....	156
5.2.5 <i>Methoden des Data Mining</i> .....	160
5.2.6 <i>Statistik und andere Methoden, die im Data Mining ihre Anwendung finden</i> .....	187
5.3 PROZESSMODELL FÜR DIE IMPLEMENTIERUNG EINER ERFOLGREICHEN DATA MINING-LÖSUNG IM CRM .....	191
<b>KAPITEL 6 .....</b>	<b>195</b>
<b>FUZZY-LOGIK .....</b>	<b>195</b>

6.1 EINFÜHRUNG .....	195
6.2 WARUM FUZZY-LOGIK?.....	196
6.3 GRUNDLAGENDE BEGRIFFE.....	197
6.3.1 <i>Klassische Mengenlehre und Logik</i> .....	197
6.3.2 <i>Unscharfe Mengenlehre und Logik</i> .....	200
6.3.3 <i>Zugehörigkeitsfunktion</i> .....	203
6.3.4 <i>Unbestimmtheit, Unsicherheit, Unschärfe, Wahrscheinlichkeit, Möglichkeit</i> ..	205
6.4 FUZZY-SET-THEORIE.....	212
6.4.1 <i>Grundlagen</i> .....	212
6.4.2 <i>Mathematische Grundlage der Fuzzy-Set-Theorie</i> .....	214
6.4.2.1 <i>Festlegung der Zugehörigkeitsfunktion</i> .....	218
6.4.2.2 <i>Operationen mit Fuzzy-Sets</i> .....	221
6.4.3 <i>Fuzzy-Logik als Erweiterung der klassischen Logik</i> .....	223
6.4.3.1 <i>Unscharfe Schlussfolgerung (Approximatives Schließen)</i> .....	227
6.4.3.2 <i>Linguistische Variablen und Terme</i> .....	228
6.5 FUZZY-SYSTEME.....	230
6.6 DATENANALYSE MITTELS FUZZY-LOGIK.....	232
6.6.1 <i>Clustering</i> .....	233
6.6.2 <i>Grundlagen der Fuzzy-Clusteranalyse</i> .....	236
6.6.3 <i>Probabilistische und Possibilistische Clustereinteilung</i> .....	238
6.6.4 <i>Funktionen und Bestandteile einer probabilistischen Clusteranalyse</i> .....	241
6.6.5 <i>Probabilistischer Clustering-Algorithmus</i> .....	245
6.6.6 <i>Fuzzy-C-Means-Algorithmus</i> .....	247
6.6.7 <i>Ähnlichkeits- bzw. Distanzmaße</i> .....	249
6.6.8 <i>Bewertung der Clustereinteilung - Bestimmung der Clusteranzahl</i> .....	253
6.6.9 <i>Ablauf des Fuzzy-C-Means-Algorithmus</i> .....	259
6.6.10 <i>Was kann mit Hilfe der Fuzzy-Datenanalyse im CRM erreicht werden?</i> .....	262
<b>KAPITEL 7 .....</b>	<b>264</b>
<b>FUZZY-LOGIK BASIERTE SEGMENTIERUNG VON KUNDEN FÜR CRM.....</b>	<b>264</b>
7.1 EINFÜHRUNG .....	264
7.2 METHODOLOGIE .....	264
7.3 PROBLEMSTELLUNG .....	265
7.4 ANFORDERUNGEN.....	266
7.5 ANALYSE UND ENTWURF .....	274
7.6 IMPLEMENTIERUNG UND DEMONSTRATION .....	278
7.7 EVALUIERUNG DER ANALYSEERGEBNISSE .....	287
<b>KAPITEL 8 .....</b>	<b>291</b>
<b>SCHLUSSFOLGERUNG .....</b>	<b>291</b>

<b>9 LITERATURVERZEICHNIS .....</b>	<b>294</b>
<b>10 ANHANG .....</b>	<b>309</b>

ABBILDUNGSVERZEICHNIS

Abbildung 1: Transaktions-Marketing vs. Beziehungs-Marketing..... 18

Abbildung 2: Einflussfaktoren des CRM ..... 26

Abbildung 3: Monetäre Nutzenentwicklung im Verlauf einer Kundenbeziehung.....30

Abbildung 4: Berechnung des Kapitalwerts einer Kundenbeziehung ..... 31

Abbildung 5: Abgrenzung des CRM von verwandten Begriffen ..... 32

Abbildung 6: Untergeordnete Ziele des CRM-Konzepts ..... 34

Abbildung 7: Klassifizierung von CRM-Systemen ..... 43

Abbildung 8: Komponenten einer CRM-Lösung ..... 44

Abbildung 9: Kreislauf des Kampagnenmanagements ..... 49

Abbildung 10: Zentrale Bedeutung des Customer Data Warehouse im CRM-Konzept..... 59

Abbildung 11: Stufen des Data Mining-Prozesses ..... 65

Abbildung 12: Kosten- und Nutzenvorteile durch CRM ..... 72

Abbildung 13: Wirkungsebenen und Ansatzpunkte von CRM-Systemen ..... 73

Abbildung 14: Ebenen der Begriffshierarchie ..... 86

Abbildung 15: Formen der Wissensumwandlung..... 91

Abbildung 16: Die Spirale des Wissens..... 93

Abbildung 17: Gestaltungsdimensionen eines ganzheitlichen Wissensmanagements ..... 99

Abbildung 18: Bausteine des Wissensmanagements ..... 102

Abbildung 19: Kundenprozesse ..... 108

Abbildung 20: CRM-Prozesse ..... 109

Abbildung 21: Sicht auf Geschäftsprozesse als Wissensprozesse ..... 109

Abbildung 22: Interdependenzen zwischen Prozessen ..... 110

Abbildung 23: Wissensflüsse im CRM-Prozess ..... 111

Abbildung 24: Erstellung eines Gesamtbildes des Kunden im CKM..... 112

Abbildung 25: Systempyramide ..... 113

Abbildung 26: Bestandteile von Management Support-Systemen ..... 117

Abbildung 27: Einordnung unterschiedlicher Technologien und Abgrenzungen von Business Intelligence..... 119

Abbildung 28: Die Elemente eines Data Warehouse ..... 126

Abbildung 29: Top Down und Bottom Up Implementation im Vergleich ..... 130

Abbildung 30: Von der zweistufigen zur dreistufigen Architektur ..... 132

Abbildung 31: Bestandteile des Database Marketing ..... 138

Abbildung 32: Das logische Modell eines Database Marketing-Systems ..... 139

Abbildung 33: Vorgehensmodell des Knowledge Discovery in Databases ..... 150

Abbildung 34: Klassifikation der Ziele des Knowledge Discovery in Databases..... 153

Abbildung 35: top down-versus bottom up-Analyseverfahren .....	155
Abbildung 36: Aufgabenstellungen des Data Mining .....	157
Abbildung 37: Einbettung von Data Mining im maschinellen Lernen .....	163
Abbildung 38: 3-nächste-Nachbarn-Methode (k=3) .....	168
Abbildung 39: Diskriminanzanalyse .....	170
Abbildung 40: Entscheidungsbaum für die Kreditwürdigkeitsdaten .....	173
Abbildung 41: Aufbau eines Künstlichen Neuronalen Netzes .....	176
Abbildung 42: Hierarchisches Clustering .....	182
Abbildung 43: 3-means Clustering .....	183
Abbildung 44: Zentrale Maßgrößen der Assoziationsanalyse .....	186
Abbildung 45: Generierung einer Regel in der Assoziationsanalyse .....	187
Abbildung 46: Phasen des Referenzmodells CRISP-DM.....	192
Abbildung 47: „starkes Fieber“ als klassische Menge.....	199
Abbildung 48: „starkes Fieber“ als Fuzzy-Menge .....	202
Abbildung 49: Scharfe und Fuzzy-Mengen am Beispiel „Raumtemperatur“.....	204
Abbildung 50: Zugehörigkeitsfunktion für die Fuzzy-Menge „alter Mensch“ .....	215
Abbildung 51: Beispiele für Zugehörigkeitsfunktionen .....	220
Abbildung 52: Durchschnitt zweier Fuzzy-Mengen .....	221
Abbildung 53: Vereinigung zweier Fuzzy-Mengen .....	222
Abbildung 54: Komplement einer Fuzzy-Menge.....	222
Abbildung 55: Wahrheitswerte einer Fuzzy-Aussage .....	226
Abbildung 56: Die linguistische Variable „Alter“ .....	229
Abbildung 57: Darstellung eines Fuzzy-Systems.....	231
Abbildung 58: Unterschiedliche Metriken.....	234
Abbildung 59: Schmetterlingsproblem .....	237
Abbildung 60: Crisp- und Fuzzy-Cluster des Schmetterlings .....	237
Abbildung 61: Klassifikation mit dem Fuzzy-C-Means Algorithmus.....	238
Abbildung 62: Klassifikationsproblem.....	240
Abbildung 63: Streudiagramm DB-Umsatz für 200 Datenobjekte .....	271
Abbildung 64: Client-Server-Architektur.....	274
Abbildung 65: Gütegraphik.....	285

## TABELLENVERZEICHNIS

---

Tabelle 1: Gründe für die Entwicklung des CRM .....	19
Tabelle 2: Erfolgsfaktoren bei der Einführung von CRM-Systemen .....	80
Tabelle 3: Unterschiede zwischen implizitem und explizitem Wissen .....	90
Tabelle 4: Gegenüberstellung der operativen u. managementunterstützenden Datenverarbeitungs- Systeme .....	114
Tabelle 5: Star- und Snowflake-Schemen.....	131
Tabelle 6: Evolutionsschritte des Data Mining .....	144
Tabelle 7: Konzepte von Unsicherheit .....	207
Tabelle 8: Anwendungsbereiche der Fuzzy-Set-Theorie.....	213
Tabelle 9: Verteilung von 200 Datenobjekten .....	287
Tabelle 10: Die gefundenen Clusterzentren für vier Merkmale (Dimensionen) nach der Analyse .....	288

## BILDVERZEICHNIS

---

Bild 1: Datentabelle für die Kundensegmentierung .....	269
Bild 2: Benutzeroberfläche für die Konfiguration.....	279
Bild 3: Die Konfiguration-GUI nach der Datenbankanbindung und Auswahl der Customer Tabelle....	279
Bild 4: Die Konfiguration-GUI mit auswählbaren Feldern der Customer Tabelle .....	280
Bild 5: Die Konfiguration-GUI nach der Auswahl der Felder (Dimensionen) und Fuzzyparameter .....	280
Bild 6: Die Ergebnis-GUI für die Darstellung der Fuzzy-Ergebnisse der Clusterbildung (a).....	281
Bild 7: Die Ergebnis-GUI für die Darstellung der Fuzzy-Ergebnisse der Clusterbildung (b).....	281

## SUMMARY

---

The goal of this thesis is to analyze existing approaches to customer relationship management techniques and their support with Fuzzy Logic based techniques. To demonstrate how this approach can work in practice, a small prototype was implemented.

Today, a primary challenge of companies is how to turn their customer databases into a competitive business advantage by converting masses of existing data into useful information. How this challenge is met is critical, because companies are increasingly relying on an effective analysis of customer data to remain competitive. As the leading competitive business strategy of today's companies, Customer Relationship Management (CRM) aims at targeting profitable customers to increase the revenue to the company. Technologically, it primarily depends on various analytics to identify profitable customers using their historic data.

The rapid development of data mining methods and techniques enables using large amounts of customer data to extract the knowledge. The extracted knowledge from data could be used for supporting the marketing decision process. Therefore, many companies have already recognized data mining as an important procedure that will have an impact on their performance. Data mining could be defined as the process of nontrivial extraction of implicit, previously unknown and potentially useful information such as knowledge rules, constraints, and regularities from customer data stored in repositories using pattern recognition technologies as well as statistical and mathematical techniques.

Clustering is one of the major data mining functions. That is why master thesis focuses on cluster analysis, which is probably the most applied methodology for customer analysis in CRM. It is used to identify various types of customers in different segments. Accordingly, companies try to segment their customers by identifying groups of customers with different structures and properties that are as homogeneous as possible within each group and significantly heterogeneous between groups. These groups can then be addressed with a specially designed marketing strategy. It is essential to build refined marketing strategies for customers based on their different characteristics.

Consequently, the aim of this thesis is to examine fuzzy theory as a methodological approach for data mining in Customer Relationship Management. The main advantage of fuzzy theory is that uncertainty resulting from the complexity of the data could be more efficiently handled than with traditional data mining techniques. Further on, the thesis

investigates whether fuzzy clustering could be effectively applied to customer segmentation for analytical Customer Relationship Management (aCRM).

Fuzzy clustering is based on „Fuzzy Set Theory“, which was first introduced in 1965 by Lotfi A. Zadeh. It may be regarded as both, a generalization of classical set theory and a generalization of dual logic. The main issue of clustering is to find patterns or structures contained within groups of data. These structures are usually classes to which objects from the data set are assigned. Classical clustering techniques assign each data object to exactly one class, whereas in fuzzy clustering the data objects could be assigned different degrees of membership to the different classes. This gradual membership of an object to different classes makes it possible to achieve comprehensible and meaningful results in cluster analysis. Furthermore, it provides many useful advantages compared to conventional clustering techniques in numerous instances, and contributes critically to the data analysis and classification.

The master thesis is structured into two main parts, a theoretical and a practical one. The first part gives an insight into the aspects of CRM, Knowledge Management, Data Mining and especially Fuzzy Logic, and establishes the basis of the practical part. The second part, which contains the author’s own major contribution to the diploma thesis, contains the implementation and demonstration of a program, which implements the Fuzzy C-means Algorithm chosen for the customer segmentation in the case study. Fuzzy clustering and its benefits for the customer analysis are shown in an application using a case study on a retail company. The case-specific interpretation of identified customer segments as the result of fuzzy analysis is also given there. The thesis ends with chapter eight making some conclusions about fuzzy clustering and its significance for customer relationship management.

## ZUSAMMENFASSUNG

---

Ziel dieser Diplomarbeit ist es, vorhandene Ansätze zu Kundenbeziehungsmanagement-Techniken und ihre Unterstützung mit Fuzzy-Logik-basierten Techniken zu analysieren. Um zu demonstrieren, wie dieser Ansatz in der Praxis funktionieren kann, wurde ein kleiner Prototyp implementiert.

Eine primäre Herausforderung der sich Firmen stellen müssen, ist heute, wie sie ihre Kundendatenbanken zu einem Wettbewerbsvorteil machen, indem sie Massen vorhandener Daten in nützliche Information umwandeln. Wie dieser Herausforderung begegnet wird, ist insofern kritisch, als Firmen in zunehmendem Maße auf eine wirkungsvolle Analyse von Kundendaten angewiesen sind, um konkurrenzfähig zu bleiben. Als eine der führenden Geschäftsstrategien der heutigen Firmen strebt Kundenbeziehungsmanagement (Customer Relationship Management) an, auf rentable Kunden zu zielen, um die Firmenerträge zu erhöhen. Technologisch gesehen, orientiert sich CRM hauptsächlich an verschiedenen Analysetechniken zur Identifizierung rentabler Kunden durch die Verwendung ihrer historischen Daten.

Die schnelle Entwicklung von Data Mining-Methoden und -Techniken ermöglicht unter Verwendung großer Mengen von Kundendaten, das Wissen aus den Daten zu extrahieren. Das aus diesen Daten extrahierte Wissen kann zur Unterstützung des Marketing-Entscheidungsprozesses verwendet werden. Daher haben viele Firmen bereits Data Mining als ein wichtiges Verfahren erkannt, das eine Auswirkung auf ihre Leistung hat. Data Mining kann als der Prozess der nicht trivialen Extraktion von impliziten, vorher unbekannter und potentiell nützlicher Information, wie Wissensrichtlinien, -begrenzungen und -regelmäßigkeiten aus in Repositories gespeicherten Kundendaten unter Verwendung von Mustererkennungstechnologien sowie von statistischen und mathematischen Techniken angesehen werden.

Segmentierung ist eine der Hauptfunktionen von Data Mining. Das ist der Grund, warum sich die Diplomarbeit auf die Clusteranalyse konzentriert, welche vermutlich die am meisten angewandte Methodologie für Kundenanalyse in CRM ist. Sie wird verwendet, um verschiedene Typen von Kunden in verschiedenen Segmenten zu identifizieren. Dementsprechend versuchen Firmen, ihre Kunden zu segmentieren, indem sie Gruppen von Kunden mit unterschiedlichen Strukturen und Eigenschaften identifizieren, die so homogen wie möglich innerhalb jeder Gruppe und erheblich heterogen zwischen Gruppen sind. Diese Gruppen können dann mit einer besonders entworfenen Marketingstrategie angesprochen werden. Es ist essentiell, präzise Marketingstrategien für die Kunden zu

entwickeln, die auf ihren verschiedenen Eigenschaften basieren.

Infolgedessen ist das Ziel dieser Arbeit, die Fuzzy-Theorie als eine methodologische Annäherung für Data Mining im Kundenbeziehungsmanagement zu untersuchen. Der Hauptvorteil der Fuzzy-Theorie ist, dass die Ungewissheit, die aus der Komplexität der Daten resultiert, mit ihr effizienter als mit traditionellen Data Mining-Techniken behandelt werden kann. Weiterhin untersucht die Arbeit, ob das Fuzzy-Clustering zur Kundensegmentierung für analytisches Kundenbeziehungsmanagement effektiv angewendet werden kann.

Das Fuzzy-Clustering basiert auf „Fuzzy Set Theory“, die zuerst 1965 von Lotfi A. Zadeh vorgestellt wurde. Es kann als eine Verallgemeinerung der klassischen Mengenlehre und als eine Verallgemeinerung der Dual-Logik angesehen werden. Die Kernfrage des Clustering ist, die Muster oder Strukturen, die innerhalb der Gruppen von Daten enthalten sind, zu finden. Diese Strukturen sind normalerweise Klassen, denen Objekte eines Datensatzes zugewiesen werden. Klassische Segmentierungstechniken weisen jedes Datenobjekt genau einer Klasse zu, während in Fuzzy-Segmentierung den Datenobjekten verschiedene Grade der Zugehörigkeit zu verschiedenen Klassen zugewiesen werden können. Diese graduelle Zugehörigkeit eines Objektes zu den verschiedenen Klassen macht es möglich, verständliche und sinnvolle Resultate in der Clusteranalyse zu erzielen. Außerdem liefert Fuzzy-Clustering im Vergleich zu herkömmlichen Clustering-Techniken in zahlreichen Fällen viele nützliche Vorteile, und trägt kritisch zur Datenanalyse und zur Klassifikation bei.

Diese Diplomarbeit wurde in zwei Hauptteile, in einen theoretischen und praktischen Teil untergeteilt. Der erste Teil gibt einen Einblick in die Aspekte von CRM, Wissensmanagement, Data Mining und besonders Fuzzy-Logik, und stellt die Basis des praktischen Teils her. Der zweite Teil, welcher den eigenen Hauptbeitrag des Autors zur Diplomarbeit enthält, beinhaltet die Implementierung und die Demonstration eines Prototyps, der den Fuzzy-C-Means-Algorithmus implementiert, welcher für die Kundensegmentierung in der Fallstudie gewählt wurde. Das Fuzzy-Clustering und sein Nutzen für die Kundenanalyse wurden in einer Anwendung anhand einer Fallstudie für ein fiktives Retail-Unternehmen gezeigt. Die fallspezifische Interpretation der identifizierten Kundensegmente als das Ergebnis der Fuzzy-Analyse wurde auch in diesem Teil der Arbeit dargestellt. Die Diplomarbeit endet mit Kapitel 8, das einige Schlussfolgerungen über das Fuzzy-Clustering und seine Bedeutung für das Kundenbeziehungsmanagement enthält.

# KAPITEL 1

---

## **EINLEITUNG**

In den letzten Jahren haben zahlreiche Entwicklungen die Basis für den Einsatz des Customer Relationship Management gelegt. Durch die Marktveränderungen der 90er Jahre wurde der Wettbewerbsdruck sehr verstärkt. Das Internet hat die Globalisierung beschleunigt und die Unternehmen können über dieses Medium weltweit ihre Produkte anbieten. Es erleichtert dem Kunden die Informationsgewinnung und den Preisvergleich und steigert die Transparenz der Märkte erheblich. Dementsprechend wird das Eingehen auf die Wünsche und Bedürfnisse der Kunden zum entscheidenden Differenzierungsmerkmal, wenn die Produkte sich kaum mehr voneinander unterscheiden, die Märkte gesättigt sind und Markenloyalität sowie Werbeeffizienz sinken. Kundenzentrisch zu agieren wird nun zum primären Unternehmensziel. Außerdem erfordert die neue kundenzentrische Unternehmenskultur neue automatisierte und präzise Analyseverfahren, womit zukünftige Käuferverhalten von Kunden besser eingeschätzt werden können. Durch den Einsatz dieser Analyseverfahren werden die Unternehmen gegenüber der Konkurrenz im Kampf um profitable Kunden einen großen Vorsprung haben.

Ausgehend von dieser Situation wird im Rahmen dieser Diplomarbeit untersucht, welche Bedeutung die Kunden für Unternehmen haben, welche Potentiale sich durch den Einsatz verschiedener Verfahren zur Datenanalyse im CRM ergeben und ob die unscharfe Beschreibung von Kundendaten durch linguistische Terme und die unscharfe Analyse von (scharfen und unscharfen) Kundendaten überhaupt sinnvoll ist.

Die ökonomischen und technischen Geschäftsprozesse in Unternehmen und damit verbundenen Entscheidungen werden ständig noch komplexer. Einerseits sind analytische Modelle häufig unzulänglich, diese Komplexität zu beschreiben und damit umzugehen und andererseits häufen sich die Datensätze mit wachsender

Geschwindigkeit an und stellen dadurch eine potentielle Quelle des Wissens zum Verstehen und Managen der Prozesse bereit. Das allgemeine Problem dabei ist, aus dieser Datensammlung passende Modelle zu entwickeln, welche leicht zu verstehen sind.

Um solche Modelle von den Daten abzuleiten, existieren unterschiedliche Methoden wie z. B. Statistische Regression, Neuronale Netze und andere Data Mining basierte Ansätzen. Die resultierenden Modelle sind jedoch entweder leicht zu verstehen, wie dies in Regel-basierten Ansätze der Fall ist oder genau bzw. präzise (z. B. Neuronale Netze). In den Anwendungen der realen Welt sind aber beide Aspekte von besonderer Bedeutung und spielen durchaus eine wichtige Rolle. Als Konsequenz entsteht die Notwendigkeit, Modelle zu entwickeln, welche beide Aspekte, Verständlichkeit und Genauigkeit gleichzeitig berücksichtigen können. Modelle basierend auf einem Fuzzy-Logik-Ansatz können bei der Bewältigung dieser Herausforderung sehr nützlich sein.

Eine andere und sehr wichtige Herausforderung ist noch, die Aussage, dass die Ereignisse der realen Welt nicht immer als richtig oder falsch beurteilt werden können; sondern dass die reale Welt neben den Extremwerten ja/nein, wahr/falsch auch Zwischenstufen für jeweilige Ereignisse besitzt. Mit dieser Definition beansprucht die Fuzzy-Logik, umfassender zu sein als die klassische zweiwertige Logik, und dadurch lässt sich die reale Welt viel besser modellieren.

Dieser Ansatz findet seine Anwendung auch im Customer Relationship Management. CRM ist eine Strategie mit dem Ziel, alle Unternehmensprozesse konsequent auf die profitablen Kunden auszurichten in der Absicht, deren Bedürfnisse und Erwartungen zu erkennen, individuell darauf reagieren zu können und diese Kunden langfristig an das Unternehmen zu binden. Mit Hilfe der Data Mining Techniken kann diese Strategie leichter formuliert werden. Im Rahmen dieser Arbeit wird festgestellt, ob die Fuzzy-Theorie über ein Rationalisierungspotential für Data Mining Anwendungen im CRM verfügt und was mit Hilfe dieser Theorie im Wissensmanagement tatsächlich erreicht werden kann.

## **1.1 ZIELSETZUNG und VORGEHEN**

Das Ziel dieser Diplomarbeit ist es, den Leser in die Thematik der unscharfen Logik einzuführen und ihm einen Überblick über die Entwicklung, Bedeutung, Grundlagen des Fuzzy-Logik-Ansatzes und dessen Implementierung für Customer Relationship Management zu verschaffen. Zudem sollen die Besonderheiten dieses Ansatzes anhand eines Fallbeispiels zur Analyse von Kunden im CRM aufgezeigt werden.

Die Diplomarbeit ist grundsätzlich in einen theoretischen und praktischen Teil gegliedert. Der theoretische Teil bereitet das nötige Wissen aus verschiedenen Disziplinen wie z. B., das Data Mining - die Extraktion von Wissen aus großen Datenmengen, die unscharfe Logik - ein formalisierter Ansatz aus numerischer Mathematik zur vagen Beschreibung des Systemverhaltens auf und bildet somit die Grundlage für das Verständnis des praktischen Teils, welcher ein Fuzzy-Logik basiertes Rahmenwerk zur Datenanalyse im CRM ermittelt.

Die vorliegende Diplomarbeit besteht aus acht Kapiteln, wobei jedoch der Schwerpunkt der Arbeit im sechsten und siebten Kapitel liegt. Die Einleitung beschreibt den aktuellen Wettbewerb um Kunden im freien Markt, betont die Wichtigkeit des Kunden für strategische Entscheidungen der Unternehmen und leitet daraus die Forderung nach einer gezielten Analyse von Kundeninformationen mit Hilfe von Data Mining Techniken im CRM ab. Sie erläutert ferner die Zielsetzung und Vorgehensweise der Arbeit. Um den weiteren Kontext besser verstehen zu können, werden zunächst einige Begrifflichkeiten und Themen angeführt und erklärt. Der anschließende Abschnitt befasst sich mit dem Thema Customer Relationship Management und stellt die grundlegenden Konzepte und Begriffe dieses Management-Konzeptes vor.

Nach der Erläuterung der Bedeutung und Kennzeichen des CRM wird auf die Business Intelligence im CRM eingegangen. Das vierte Kapitel beschäftigt sich also mit dem analytischen CRM, welches auch als Business Intelligence bezeichnet wird. Business Intelligence ist eines der wichtigsten Konzepte zur Generierung von Wissen über das Unternehmen und bildet damit den Ausgangspunkt und die Grundlage der Unternehmensentwicklung. Das Ziel ist dabei die Gewinnung von Erkenntnissen, die in Hinsicht auf die Unternehmensziele bessere operative, taktische oder strategische Entscheidungen ermöglichen. Es werden die Relevanz von Daten und deren Analyse im Hinblick auf den gewünschten Erkenntnisgewinn verdeutlicht. Zudem werden diverse analytische Konzepte, die im CRM eingesetzt werden können, behandelt.

Das fünfte Kapitel konzentriert sich auf das Data Mining und seinen Einfluss auf CRM.

Es erklärt die zugrundeliegenden Ziele und Konzepte dieser Technologie in Abhängigkeit von den Einsatzmöglichkeiten im CRM.

Das sechste Kapitel befasst sich mit der Fragestellung, ob der Fuzzy-Logik-Ansatz als eine erfolgreiche Methodik zur Datenanalyse im Informationsmanagement überhaupt angewendet werden kann. Zunächst wird unter Bezugnahme auf die grundlegenden Arbeiten des Erfinders der unscharfen Logik, A. L. Zadeh, die Idee der *Theorie unscharfer Mengen* erklärt. Danach werden die wichtigsten Begriffe der unscharfen Logik eingeführt. Die Fuzzy-Set-Theorie wird detailliert beschrieben, und darüber hinaus wird versucht, sämtliche Fragestellungen über den Fuzzy-Logik-Ansatz zu beantworten. Typische Fragestellungen seien z. B.: Was für Einsatzmöglichkeiten und Nutzenpotentiale besitzt der Fuzzy-Logik-Ansatz? Was kann mit Hilfe der Fuzzy-Set-Theorie bei der Daten- bzw. Kundenanalyse im CRM erreicht werden? Was sind die Vorteile gegenüber klassischer Theorie? usw.

Zur Vertiefung des Verständnisses wird im siebten Kapitel anhand eines selbst gewählten Fallbeispiels eine Fuzzy-Logik basierte Datenanalyse für CRM durchgeführt und die Analyseergebnisse ausgewertet.

Zum Abschluss der Arbeit wird noch versucht, die Schlussfolgerungen aus den gewonnenen Erkenntnissen zu ziehen und vor allem einige Vorschläge und Forderungen bezüglich des Einsatzes der Fuzzy-Logik für Data Mining im CRM zu machen.

# KAPITEL 2

---

## **BEGRIFFLICHE ABGRENZUNG**

### **2.1 KUNDENVERHALTEN IM INTERNETZEITALTER**

Das Internet als universelles Medium hat in einer kurzen Zeit einen weltweiten Aufschwung des elektronischen Handels bewirkt. Es dient zunehmend nicht nur der Informationsvermittlung. Vielmehr wird es zu einer Plattform für den interaktiven Kauf und Verkauf aller Arten von Produkten und Dienstleistungen. So stellen neue Geschäftsmodelle bisherige Konzepte des Austausches von Gütern und Dienstleistungen in Frage, die Beziehungen zwischen Anbietern und Nachfragern müssen neu definiert und die Spielregeln des Wettbewerbs neu festgelegt werden. Die Art und Weise, wie Geschäfte zwischen Unternehmen und ihren Kunden abgewickelt werden, verändert sich radikal. Räumliche und zeitliche Restriktionen werden durch die globale Vernetzung aufgehoben, Unternehmungsgrenzen überwunden. Alle diese Entwicklungen führen allerdings dazu, dass sich überall neue Formen der kollaborativen Wertschöpfung etablieren.

Mit seinen vielfältigen Möglichkeiten der Vernetzung von Marktteilnehmern bietet das Internet Unternehmen großartige Aussichten, Geschäftsprozesse schneller, kostengünstiger und effektiver zu gestalten und die Beziehungen zu ihren Kunden und Lieferanten zu verstärken, um letztlich im globalen Wettbewerb zu bestehen und langfristig im Markt profitabel zu agieren. Die Verheißungen des Internets, mittels E-Commerce neue Markt- und Umsatzpotentiale zu nutzen, Vertriebskosten drastisch zu senken sowie Kunden enger und dauerhafter an sich zu binden, werden von Unternehmen sehr schnell akzeptiert.

Viele Firmen aus unterschiedlichsten Branchen setzen große Erwartungen in den neuen Vertriebskanal Internet. Sämtliche Vorteile des Internets als Vertriebskanal

können Folgendes sein:

- Grenzenlose Marktpräsenz rund um die Uhr und um den Globus,
- Beschleunigung von Vertriebsprozessen durch automatisierte Abläufe,
- Senkung der Transaktionskosten durch Automatisierung der Vertriebsprozesse über Unternehmensgrenzen hinweg,
- Senkung der Personalkosten durch Reduzierung telefonischer Bestellungen und Anfragen,
- gezieltere, direkte Kundenansprache durch umfangreichere Informationen über Kunden,
- höhere Kundenzufriedenheit und stärkere Kundenbindung durch besseren Kundenservice,
- höhere Umsätze durch die Erschließung neuer Kundengruppen und die Ausschöpfung des Cross/Up Selling-Potentials.<sup>1</sup>

Das Kundenverhalten in traditionellem und online Handel weist grundsätzliche Unterschiede auf, die für die zielgerechte Ausrichtung des Customer Relationship Management von besonderer Bedeutung sind. In der Regel sucht der Onlinekunde bestimmte Informationen oder Produkte und besucht dafür auch ganz gezielt einzelne Websites. Vor diesem Hintergrund soll ein Unternehmen dem Kunden ermöglichen, dass er erstens die Website des Unternehmens im Internet überhaupt findet und zweitens online möglichst einfach auf Informationen zugreift, die er letztendlich sucht. Aus der Tatsache, dass dem Kunden das Anfassen und Erleben der Ware im Internet nicht möglich ist, resultieren Spontankäufe seltener als im traditionellen Handel.

Das Internet ermöglicht den freien Zugang zu Informationen und eröffnet dem Kunden noch die Möglichkeit, dass er bewusst eine Webseite oder eine Information sucht, wo immer und wann immer er sie braucht. Aus dieser Perspektive kann das Internet als ein Pull-Medium bezeichnet werden.<sup>2</sup> Zudem werden Informationen von einer Vielzahl von Anbietern in unterschiedlicher Art und Weise zur Verfügung gestellt. Der Kunde ist also viel weniger abhängig vom einzelnen Anbieter. Aufgrund dieser Verkaufsflexibilität bzw. -unabhängigkeit des Kunden herrscht hier eine höhere

---

<sup>1</sup> Vgl. Buck-Emden/Zencke (2004), S. 232.

<sup>2</sup> Vgl. Förster (2001), S. 33.

Wettbewerbsintensität und Markttransparenz als im traditionellen Handel.

Moderne IuK-Technologien eröffnen zahlreiche neue Möglichkeiten.<sup>3</sup> Leistungsfähige Kundendatenbanken ermöglichen dialogorientierte Formen des Marketings. Internetseiten, Instant Messaging, E-Mails und Newsletters erlauben es den Online-Kunden, direkt und in erster Linie schnell mit Unternehmen in Kontakt zu treten. Die Internetökonomie mit ihren speziellen Markt- und Rahmenbedingungen macht die Kundenbindung im Electronic Business zu einem zentralen Thema.<sup>4</sup> Die Grundüberlegung hier ist, dass gute und dauerhafte Beziehungen zu Kunden oft das wichtigste Kapital eines Unternehmens sind und es Ziel sein muss, solche Beziehungen aufzubauen und langfristig weiterzuentwickeln. Dies ist gerade im Internetmarketing von großer Bedeutung. Neue Kunden zu gewinnen ist bekanntlich etwa sieben bis zehn Mal teurer, als bestehende Kunden zufrieden zu stellen.<sup>5</sup> Schließlich ist der Konkurrent nur einen Klick entfernt.

Die langfristige Bindung profitabler Kunden an das Unternehmen kann dessen Ertragsituation deutlich verbessern.<sup>6</sup> Aus diesem Grund substituiert im Electronic Business die Sicherung einer bestehenden Kundenbasis weitgehend die Neukundenakquisition.<sup>7</sup> Die Kundenbindung und mithin das Kundenbindungsmanagement in der Internetökonomie wird somit zu einem zentralen Erfolgsfaktor<sup>8</sup>, denn Beziehungen zu verstärken ist die zentrale ökonomische Forderung der Internetökonomie.

## **2.2 TRADITIONELLE MARKETING vs. ONE-TO-ONE MARKETING**

Die Globalisierung der Märkte hat in der Ökonomie zu bedeutsamen Veränderungen geführt. Die Staatsgrenzen stellen für die Handelsbeziehungen zwischen den Unternehmen kein Hindernis mehr dar. So beteiligen sich Unternehmen, die sich in geographisch verschiedenen Orten der Welt niedergelassen haben, an einer Produktherstellung oft gemeinsam. Der Austausch von Informationen wird durch den gezielten Einsatz moderner Informations- und Kommunikationstechnologien ganz erleichtert. Einerseits befinden sich Unternehmen in einer internationalen Wettbewerbssituation mit vielen Konkurrenten. Andererseits sind aber erhebliche

---

<sup>3</sup> Vgl. Loebbecke (2001), S. 105.

<sup>4</sup> Vgl. Wirtz (2000), S. 89.

<sup>5</sup> Vgl. Kreuz/Förster (2001), S. 10.

<sup>6</sup> Vgl. Kreuz/Förster (2001), S. 10.

<sup>7</sup> Vgl. Wirtz (2001), S. 502.

<sup>8</sup> Vgl. (eigene Übersetzung) Wirtz (2001), S. 502.

Kostensenkungen und Effizienzsteigerungen realisierbar. Dieser Sachverhalt führt zu einer höheren Souveränität des Kunden, der Produkte und Leistungen je nach seinen Wünschen und Anforderungen von unterschiedlichen Anbietern verkaufen kann.

Umgeben von flexiblen Kunden und zahlreichen Konkurrenten versuchen Unternehmen wirtschaftlich zu agieren, um wettbewerbfähig auf dem Markt zu bleiben. Das Verlangen der Kunden nach individuellen Leistungen und Produkten sowie der steigende Wettbewerbsdruck stellen Unternehmen immer wieder vor neue Herausforderungen.

Die große Angebotsvielfalt an Produkten verursachte in den 80er Jahren einen Wandel des bisherigen Käuferverhaltens. Die herkömmlichen Werbemaßnahmen des Massenmarketing wurden von vielen Kunden nicht mehr in Anspruch genommen. Dementsprechend sahen Unternehmen die Notwendigkeit dafür, ihre Marketingaktivitäten auf die speziellen Wünsche des Kunden auszurichten und diese in der Geschäftsbeziehung noch stärker zu berücksichtigen. So verfolgten Unternehmen nicht länger die kurzfristigen Generierung von Umsätzen, sondern konzentrierten sich auf den Aufbau und Ausbau von Kundenbeziehungen. Die erforderlichen Maßnahmen zur gezielten Befriedigung der Kundenbedürfnisse waren jedoch meist sehr teuer und zeitaufwendig. Unternehmen wollten zwar ihre Kunden in den Prozess der Leistungserstellung stärker integrieren, doch scheiterten sie bei dieser Zielsetzung oft an mangelhaften Kenntnissen über ihren Kunden.

Zur gleichen Zeit machten sogenannte IuK-Technologien einen riesigen Entwicklungsfortschritt und leisteten Unternehmen schließlich die notwendige Hilfe bei der Realisierung der kundenorientierten Ziele. Zusätzlich zu den herkömmlichen Massenmedien wie, Fernsehen, Rundfunk und Druckmedien entstanden neue Arten von Medien, die individuell adressierbar sind. Die charakteristischen Eigenschaften dieser so genannten One-to-One Medien sind, die Bidirektionalität des Informationsaustausches, ihre Individualisierbarkeit und die geringen Nutzungskosten.<sup>9</sup>

Die Nutzung dieser Art von Medien eröffnet Unternehmen die Möglichkeit, Kunden bedarfsgerecht und individuell anzusprechen, was in der Massenkommunikation nicht möglich ist. Der wechselseitige Informationsfluss trägt zur Erkenntnisgewinnung über den Kunden bei. Die aus dem Dialog mit dem Kunden gewonnenen Informationen können zur differenzierten Marketingansprache des einzelnen Kunden genutzt werden. Das individuelle Zuschneiden von Marketingmaßnahmen auf jeden einzelnen Kunden

---

<sup>9</sup> Vgl. Peppers/Rogers (1996), S. 28.

nennt man in diesem Zusammenhang One-to-One Marketing.

In den vergangenen Jahren sind Produkte immer besser geworden. Gleichzeitig ähneln sie aber einander in hohem Maße. Im Kampf um die Aufmerksamkeit wird der Kunde nun endlich wirklich zum König. Daher muss die Kommunikation in erster Linie eines tun: Eingehen auf die Bedürfnisse jedes Einzelnen. Sie muss persönlich relevant sein, einen direkten Mehrwert bieten und einen Unterhaltungswert besitzen. Da eine Beziehung nicht von heute auf morgen aufgebaut werden kann, muss sie auch langfristig orientiert sein. Eine Beziehung muss zunächst reifen, damit sie dann Früchte trägt - für beide Seiten, Unternehmen und Kunden.

Das traditionelle Marketing kann mit einer Badewanne verglichen werden: Es herrscht ein ständiger Abfluss von bestehenden Kunden. Doch anstatt den Abfluss zu stopfen, versuchte das klassische Marketing immer, den Zulauf zu erhöhen und noch mehr neue Kunden zu generieren. Dementsprechend wurden bestehende Kunden ignoriert und nicht mehr berücksichtigt. Je mehr Kunden unzufrieden waren, desto mehr neue Kunden mussten gewonnen werden. So wurde stets das Symptom bekämpft, nicht aber das Kernproblem.

Viele Studien belegen heute einheitlich, dass das Gewinnen eines neuen Kunden wesentlich aufwendiger und teurer ist als das Behalten eines bestehenden Kunden. Das langfristig orientierte, moderne One-to-One Marketing lässt daher die bestehenden Kunden nicht außer acht, sondern es konzentriert sich auf sie und sorgt dafür, dass sie völlig zufrieden gestellt werden.

Die Integration moderner Informations- und Kommunikationstechnologien (IuK) in Relationship Marketing und mithin One-to-One Marketing ermöglichte es neue Wege der Kundenbearbeitung und -betreuung zu entwickeln und umzusetzen. So können z. B. aus der Kaufhistorie wichtige Erkenntnisse über die Vorlieben der Kunden für bestimmte Produkte gewonnen werden. Dies ist aber nur dann möglich, wenn das vergangene Kaufverhalten des Kunden in einer Kundendatenbank gespeichert ist. Auf Basis des gewonnenen Wissens werden dann dem einzelnen Kunden gezieltere Angebote zukommen gelassen.

Die Voraussetzung für die erfolgreiche Anwendung dieser Strategie ist erstens das Data Warehousing und zweitens das Data Mining. Unter dem Begriff Data Warehousing versteht man die Tatsache, dass alle Kundendaten systematisch erfasst und in einer Datenbank abgelegt werden. Der Begriff Data Mining bezeichnet den kontinuierlichen Prozess der Analyse von Kundendaten zur gezielten und individuellen Kundenansprache.

## **2.3 WACHSENDE BEDEUTUNG DES KUNDEN FÜR UNTERNEHMEN<sup>10</sup>**

Die vergangenen Jahrzehnte brachten große Veränderungen hinsichtlich der Bedeutung, die Kunden für Unternehmen haben. So wechselte man in den siebziger und achtziger Jahren von der reinen Produktorientierung zur Marktorientierung.

In den siebziger Jahren standen die behavioristischen Modelle im Mittelpunkt, die statisches und dynamisches (Wieder-)Kaufverhalten untersuchten. Bedingt durch die Heterogenität der Bedürfniswelten wurden verschiedene Kundenbereiche in Segmente unterteilt. Mit Ende der achtziger Jahre rückte der Kunde als wesentlicher Erfolgsfaktor im Zuge der neuen Kundenorientierung in den Mittelpunkt unternehmerischen Handelns. Verhaltenswissenschaftliche Modelle und Ansätze standen hier im Mittelpunkt.

Aus dieser Phase ging schließlich jene der individuellen Kundenorientierung der neunziger Jahre hervor. Man konzentrierte sich zuerst auf psychologische Perspektiven der Kundenloyalität, wobei die Nachfragersicht im Vordergrund stand. In der Folge wurde auch auf die Handlungsnotwendigkeit der Unternehmen hingewiesen und somit auch der Begriff des Kundenbindungsmanagements geprägt. Dieses Konzept hat sich in den vergangenen Jahren zunehmend weiterentwickelt und wurde in zahlreichen Unternehmen im Zuge der Kundenorientierung implementiert. Auch zukünftig wird es von großer Bedeutung sein, die zurzeit noch isoliert eingesetzten Instrumente des Kundenbindungsmanagement in ein Gesamtkonzept zu integrieren.

## **2.4 KUNDENORIENTIERUNG und -ZUFRIEDENHEIT<sup>11</sup>**

Der Erfolg auf dem Markt basiert auf den Fähigkeiten des Unternehmens, seine Kunden anzulocken, zufrieden zu stellen und schließlich zu binden. Kundenzufriedenheit ist die primäre Determinante der Kundentreue und der daraus resultierenden Kundenbindung, und sie ist der Schlüssel zum Aufbau eines gut gehenden Unternehmens.

Für den stetigen Unternehmenserfolg ist es entscheidend, sich in den Augen der Kunden immer wieder neu zu erfinden und sich deren verändernden Bedürfnissen anzupassen. Denn es reicht nicht aus, nur zu wissen, welche Produkte oder Produktmerkmale die Kunden aktuell wünschen. Die Unternehmen müssen antizipieren können, in welche Richtung sich die Kundenpräferenzen entwickeln. Dies

---

<sup>10</sup> Vgl. für diesen Abschnitt Homburg/Bruhn (1999).

<sup>11</sup> Vgl. für diesen Abschnitt Scharnbacher/Guido (2003).

setzt voraus, dass der Unternehmer das Marktverhalten des Kunden versteht und diese Informationen beim Agieren auf dem Markt zum eigenen Nutzen verwendet.

Ein verschärfter Wettbewerb, engere Märkte und ein steigendes Anspruchsdenken der Kunden begründen die zunehmende Bedeutung eines stabilen Kundenstamms. Güter und Dienstleistungen im Markt sind leicht austauschbar geworden, und für Kunden stellt es heute kein großes Risiko mehr dar, ein Unternehmen oder einer Marke zu wechseln. Aus diesem Grund ist es für ein Unternehmen von grundlegender Bedeutung, einen stabilen Kundenstamm aufzubauen und zu sichern. Er ist die Grundlage für wirtschaftlichen Erfolg und langfristiges Wachstum.

Der Schlüssel zum Aufbau eines stabilen Kundenstamms sind zufriedene Kunden. Unternehmen, denen es gelingt, ihr Leistungsspektrum genau den Vorstellungen und Erwartungen der Kunden anzupassen oder sogar die Vorstellungen der Kunden zu übertreffen, erzeugen Kundenzufriedenheit und legen somit den Grundstein für zukünftige Geschäfte. Zufriedene Kunden tätigen Wiederholkäufe, entwickeln zusätzliche Bedürfnisse gegenüber dem Unternehmen und berichten Freunden und Bekannten positiv von dem Unternehmen.

Kundenzufriedenheit wird durch strikte Orientierung an der Person des Kunden erreicht. Kundenorientierung bedeutet, auf Kunden individuell einzugehen und durch Erfahrungen mit Kunden zu lernen und dadurch Kundenzufriedenheit nachhaltig und vorausschauend gewährleisten zu können. Auf Grund des zunehmend stärkeren Wettbewerbs können in Konkurrenz stehende Anbieter auf Dauer nur dann überleben, wenn sie das Problem des Nachfragers besser als die Konkurrenz lösen können. Für ein Unternehmen bedeutet das, sich konsequent an den Bedürfnissen der Nachfrager zu orientieren. Eine derartige Orientierung zielt auf das Erreichen eines geplanten Zufriedenheitsniveaus der Kunden, also auf Kundenzufriedenheit ab.

Kundenorientierung wird zwar schon seit Jahrzehnten gefordert und umgesetzt, jedoch nicht in der heute notwendigen Konsequenz und Tragweite. Obwohl sie zur Unternehmensstrategie erhoben wurde, sind bisher das Marketing und der Verkauf für die Kundenorientierung und die Kundenzufriedenheit zuständig gewesen. Die Kundenorientierung muss nun stärker als bisher auf alle Bereiche des Unternehmens bezogen werden. Nicht nur das Marketing und der Verkauf, sondern auch z. B. die Forschung, die Beschaffung und die Produktion müssen sich mit den Vorstellungen und Wünschen der Kunden auseinandersetzen und die Kundenzufriedenheit anstreben. Kundenorientierung und damit Kundenzufriedenheit muss in zunehmendem Umfang ein operatives Ziel und eine Richtschnur für alle unternehmerischen Entscheidungen

werden, um eine dauerhafte Befriedigung der Kundenbedürfnisse zu erreichen und komparative Wettbewerbsvorteile zu realisieren.

## 2.5 KUNDENBINDUNG

Die Tatsachen wie z. B. der verstärkte Wettbewerbsdruck, technologische Fortschritte und der Wandel zur Informationsgesellschaft sowie verändertes Käuferverhalten und eine zunehmende Individualisierung des Konsums stärken nachhaltig die Position des Käufers und „machen den Kundennutzen zum bestimmenden Faktor für den Markterfolg eines Unternehmens“.<sup>12</sup>

Die Ergebnisse zahlreicher Studien weisen eindrucksvoll in die Richtung einer „(...) Abkehr vom einseitigen Fokus auf die Neukundengewinnung zugunsten einer stärkeren Bindung bestehender Kunden an die Unternehmung“.<sup>13</sup>

Langfristige Beziehungen zum Kunden und die Entwicklung und Ausschöpfung der Potentiale dieser Beziehungen haben mittlerweile denselben Stellenwert, der bisher Marktanteilen zugemessen wurde. Dennoch erkennen die Autoren Homburg und Bruhn folgende positive Wirkungen von Kundenbindung auf die Zielgröße Erfolg im Unternehmen. Sie gehen davon aus, dass gebundene Kunden eher bereit sind, auch höhere Preise zu bezahlen als nicht-gebundene, d. h., es existiert ein Preissteigerungspotential. Gleichzeitig wirkt sich eine hohe Kundenbindung positiv auf die Verkaufsmenge eines Unternehmens aus. Bei gebundenen Kunden ist oft eine steigende Kauffrequenz zu beobachten und auch Cross Buying-Potentiale lassen sich leichter ausschöpfen, d. h., dass auch Produkte in anderen Geschäftsbereichen des Unternehmens nachgefragt werden, die nicht im primären Interesse des Kunden lagen. So können durch eine steigende Menge an verkauften Produkten sowie die Möglichkeit, Preise zu erhöhen, in der Folge der Umsatz und auch der Gewinn eines Unternehmens steigen.

Weitere Vorteile lassen sich auf der Kostenseite erkennen, wo v. a. sinkende Kundenbetreuungskosten durch eine Konzentration auf die wichtigen treuen Kunden zu vermerken ist. Wenn es gelingt, Kunden dazu zu bringen, ihre Meinungen, Wünsche und Vorschläge zu den Produkten einzubringen, können durch diese Integration Kosten im Entwicklungs- und Produktionsprozess gespart werden.<sup>14</sup>

Dank neuer Informationstechnologien und damit verbundener Möglichkeiten der

---

<sup>12</sup> Picot/ Reichwald/Wigand (2001), S. 6.

<sup>13</sup> Diller/Müllner (1998), S. 1220.

<sup>14</sup> Vgl. Homburg/Bruhn (1999).

Individualisierung von Leistungen und Kundenbeziehungen wird die Kundenbindung zu einem zentralen Erfolgsfaktor für Unternehmen. Nun steht eine große Chance Unternehmen gegenüber, die früher nicht voll genutzt werden konnte: Der Aufbau enger Bindungen zum Kunden.

Die Notwendigkeit und der Wert langfristiger Kundenbeziehungen wurden zwar schon lange als kritischer Erfolgsfaktor betont. Eine hohe Produktqualität oder ein besonderer Markenname reichen aber immer weniger aus, dauerhafte Kundenbindung zu erreichen. Notwendig sind heute vielmehr die Schaffung von Mehrwert für die Abnehmer und ein systematisches Customer Relationship Management, das an der integrierten Gestaltung der zwischen einem Anbieter und Abnehmer im Laufe der Zeit stattfindenden Interaktionen und Transaktionen ansetzt. Ziel ist eine individuelle Interaktion je nach Profil und Ansprüchen (und Wert) des jeweiligen Kunden.<sup>15</sup>

## **2.6 KUNDENBEZIEHUNG**

Die Erkenntnis, dass eine intensive Kundenbeziehung der wichtigste Wettbewerbsvorteil ist und den Wert eines Unternehmens messbar steigert, hingegen eine mangelnde Kundenbeziehung den Kostenfaktor enorm erhöht, ist keinesfalls neu. Neu ist allerdings die Erkenntnis, dass man sich nicht so sehr auf die Phase vor dem Kauf konzentrieren sollte, sondern eher auf die Phase nach dem Kauf. Erst das ermöglicht ein Servicedenken seitens der Unternehmen.

Die Kundenbeziehung ist unter ökonomischen Aspekten zum interessanten Instrument für Analysten und den Aktienmarkt geworden. Es wird versucht, sämtliche Fragen bezüglich der Kundenbeziehung zu beantworten wie z. B.: Wie wertvoll ist eine Kundenbeziehung, wie lassen sich ihre Potentiale sinnvoll nutzen, und wie werden unterschiedliche Kundenbeziehungsmodelle aufgebaut, die finanziell überprüfbar und langfristig profitabel sind? „Kundenbeziehungen stehen damit nicht mehr im Interessenfokus von Personen, die Unternehmen bewerten und kaufen, sondern werden zum Leitfaden für das zukunftsgerichtete Management.“<sup>16</sup>

Die Kunden haben mehr Zugang zu immer detaillierteren und besseren Informationen als je zuvor und machen davon intensiven Gebrauch. Kaum merklich übernehmen sie die Kontrolle ihrer Geschäftsabwicklungen. Auf der Unternehmensseite wird es davon gezwungen versucht, ein bisher nicht gekanntes Maß an Flexibilität zu entwickeln, mit höchster Geschwindigkeit auf Anforderungen zu reagieren und immer stärker

---

<sup>15</sup> Vgl. Piller/Schaller (2004), S. 72 ff.

<sup>16</sup> Rapp (2005), S. 12.

personalisierte Produkte und Dienstleistungen den Kunden anzubieten, wie und wann sie diese möchten. Unternehmen, die es schaffen, diesen Wandel zu nachvollziehen, werden von dem veränderten Stellenwert der Kundenbeziehung enorm profitieren und durch ihn wachsen. Unternehmen, die sich dagegen sträuben oder darauf nicht richtig reagieren, werden scheitern. Unternehmen strengen sich an, trotzdem erzielen sie keine Verbesserungen der Kundenbeziehungen und ihres ökonomischen Ergebnisses in der gewünschten Form, da sie immer wieder auf herkömmliche Strategien und Kundenbindungsmaßnahmen zugreifen. Damit ein Unternehmen den Vorteil von Kundenbeziehungen voll ausschöpft, ist eigentlich eine konsequente Neuorientierung zum Customer Relationship Management hin notwendig. Das Kundenbeziehungsmanagement wird dann als Strategie zur Unternehmenswertsteigerung zum entscheidenden Wettbewerbsvorteil und sichert das zukünftige Wachstum des Unternehmens.<sup>17</sup> Wenn dieses Konzept vom Topmanagement verinnerlicht und von den Mitarbeitern verstanden wird, lassen sich dessen Potentiale und der mit ihm verbundenen Technologien - wie Data Warehousing, Data Mining oder des Internets erst sinnvoll nutzen.<sup>18</sup>

## **2.7 WISSENSMANAGEMENT – DIE KUNDENDATEN, DIE KUNDENINFORMATIONEN und DEREN ANALYSE**

Es vergeht fast kein Tag, an dem nicht Unmengen von Daten aus verschiedenen Bereichen auf Unternehmen einfließen. Jede Unternehmensabteilung verarbeitet bewusst und unbewusst eine Flut von Informationen und Daten. Diesbezüglich stellt sich oft die Frage, wie man sich vor der wachsenden Gefahr der Informationsüberflutung schützen kann.

Aufgrund ihrer Erfahrung wissen viele Abteilungschefs und Mitarbeiter eines Unternehmens, welche Informationen für Unternehmenserfolg wichtig sind und welche nicht. Und doch müssen sie sich vor einer Datenselektion zuerst oft auch mit unwichtigen Informationen beschäftigen.

Im Wissensmanagement wird es oft versucht, die Begriffe *Daten*, *Informationen* und *Wissen* voneinander abzugrenzen, um das Management von Wissen von dem klassischen Informationsmanagement zu unterscheiden.<sup>19</sup>

Laut Angaben sämtlicher Autoren (siehe z. B. Reichold 2006) werden *Daten* als einfache, quantitative oder qualitative Fakten definiert, die durch einzelne Zeichen,

---

<sup>17</sup> Vgl. Rapp (2005), S. 14.

<sup>18</sup> Vgl. Rapp (2005), S. 15.

<sup>19</sup> Vgl. Reichold (2006), S. 28.

wie „1“, „T“ oder „+“, repräsentiert werden können. In Unternehmensalltag entstehen Daten als strukturierte Beschreibungen bei Vorgängen, Ereignissen oder Transaktionen. Naturgemäß enthalten sie keine Aussage über ihre Bedeutsamkeit, sind kontextfrei und meist in maschinenlesbarer Form gespeichert.

Je nach Umständen können Menschen vorliegenden Daten eine Bedeutung geben, was wiederum dazu veranlasst, dass es durch diesen Prozess *Informationen* entstehen oder anders formuliert, Daten werden in Verbindung mit anderen Daten zu Informationen, da sie dann innerhalb eines Kontextes stehen. Informationen formen eine Botschaft bzw. beantworten eine Frage. In diesem Sinne folgt daraus, dass Datenbanken Informationen speichern, da die Datenstrukturen in den Datenbanken den gespeicherten Daten einen Kontext verleihen (siehe. Reichold 2006).

Aus Daten Informationen zu gewinnen kann auf unterschiedliche Weise geschehen. So können z. B. den Daten, die im Rahmen eines Versuchs gemessen wurden, bspw. einen Graph zu ordnen, der dadurch entsteht, dass für verschiedene Zeitpunkte die zugehörigen Messergebnisse in einem Diagramm dargestellt werden.

Der verbleibende Begriff *Wissen* entsteht, wenn Menschen Informationen aufnehmen und diese mit bereits verfügbaren Informationen unter Einbeziehung ihrer Einstellungen und Erfahrungen vernetzen.<sup>20</sup> Laut Ackerschott steht bei der systematischen Verknüpfung von Informationen zu Wissen die Zweckorientierung im Vordergrund. Durch Wissen wird es ermöglicht, Sachverhalte zu verstehen und zu erklären, Entwicklungen zu prognostizieren, Entscheidungen zu treffen oder in der optimalen Art und Weise zu handeln. Es ist zu beachten, dass die Informationen aus qualitativer Sicht einen punktuellen Charakter aufweisen, wogegen Wissen das Verständnis von Zusammenhängen ermöglicht.

Der Weg von der Produkt- zur Kundenorientierung wird durch die neuen IT-Technologien wesentlich erleichtert. Diese Orientierung ermöglicht, die Kunden neu zu sehen, die wertvollsten Kunden herauszufiltern bzw. die in Zukunft sich als Potential erweisenden zu erfassen. Die Voraussetzung dafür ist, so gute Kundeninformationen zu gewinnen, dass man damit in der Lage ist, die vorhandene Kundenbasis nach verschiedenen Kriterien zu segmentieren, fokussierte Kampagnen zu entwerfen und kundenspezifische Dienstleistungen anzubieten. Dazu müssen die Kundendaten aus verschiedenen Quellen wie z. B. aus den Frontend-Anwendungen im Internet und den Backend-Systemen, erhoben und in einer gemeinsamen Datenbankanwendung zusammengeführt werden, die einen 360-Grad-Blickwinkel auf den Kunden ermöglicht.

---

<sup>20</sup> Vgl. Ackerschott (2001), S. 13.

Das ist die Aufgabe des Kundeninformationsmanagements (CIM). In den CIM-Prozessen werden Kundeninformationen nicht nur erhoben und aufbereitet, sondern auch gepflegt und auf dem neuesten Stand gehalten.

Die Datenbankanwendung kann, je nach Anbieter, Data Warehouse, Data Mart oder Information Store genannt werden. Das Grundprinzip hier ist folgendes: CIM-Prozesse extrahieren Kundendaten aus verschiedenen Geschäftsvorfällen und Anwendungen im Backend- und Frontend-Bereich, aggregieren sie und speichern sie in einer Datenbank. Diese Datenbank schafft die Basis für den Hauptzweck, Daten abzufragen und zu analysieren. Sobald die vorhandene Datenbank genügend Kundendaten enthält, besteht die Möglichkeit, neue Einblicke in das Kaufverhalten und die Kaufmuster der Kunden zu gewinnen. Man kann herausfinden, welche Produkte oder Dienstleistungen die wichtigsten Zielgruppen bevorzugen oder wie erfolgreich bestimmte Kampagnen waren. All das führt letztlich zu einem effektiveren Marketing.<sup>21</sup>

Es stehen sämtliche Analysetools zur Verfügung, um Daten unter verschiedenen Blickwinkeln, etwa um Kunden nach bestimmten Kriterien zu segmentieren (Produktkategorie, letzter Kauf) oder Gruppen von Kunden mit ähnlichen Kaufmustern zu bilden. Die zwei wichtigsten davon sind OLAP (Online Analytical Processing) und Data Mining.

Der wesentliche Unterschied zwischen den beiden Verfahren besteht darin, dass OLAP mit bereits definierten Dimensionen und Zusammenhängen arbeitet, wogegen Data Mining nach neuen, bisher unbekanntem Mustern forscht. Dabei ist OLAP ein nutzergesteuertes „Top Down“-Analyseverfahren, wogegen Data Mining datengetrieben, also „Bottom Up“, unbekannte oder nicht offensichtliche Datenzusammenhänge erkennt.<sup>22</sup> Auf Grund der besonderen Relevanz für CRM wird das Data Mining-Verfahren im fünften Kapitel eingehender behandelt.

---

<sup>21</sup> Vgl. Zingale/Arndt (2002), S. 155 ff.

<sup>22</sup> Vgl. Zipser (2001), S. 43.

# KAPITEL 3

---

## CUSTOMER RELATIONSHIP MANAGEMENT (CRM)

### 3.1 EINFÜHRUNG

*„Der wichtigste heute vernachlässigte Managementgrundsatz ist die Nähe zum Kunden. Seine Bedürfnisse zu erfüllen und seinen Wünschen zuvorzukommen: Darum geht es! Für allzu viele Unternehmen ist der Kunde zum lästigen Störenfried geworden. Sein unberechenbares Verhalten wirft wohldurchdachte strategische Pläne über den Haufen, seine Handlungen bringen die EDV durcheinander, und obendrein besteht er auch noch hartnäckig darauf, gekaufte Produkte müssten funktionieren!“*  
[Lew Young, amerikanischer Journalist, Chefredakteur von Business Week]

*„Das eigentliche Ziel des Marketing ist, das Verkaufen überflüssig zu machen. Das Ziel des Marketing ist, den Kunden und seine Bedürfnisse derart gut zu verstehen, dass das daraus entwickelte Produkt genau passt und sich daher selbst verkauft.“*  
[Peter F. Drucker, US-amerikanischer Ökonom österreichischer Herkunft]

In den letzten Jahren hat die Bedeutung des klassischen Marketings für den Unternehmenserfolg drastisch nachgelassen. Marketing als Unternehmensführungskonzept steht kurz vor dem Ende seiner Dominanz und verliert nach und nach seine zentrale Rolle im Unternehmensgeschehen.

Große, internationale Unternehmen unter denen wie z. B. BMW, die Deutsche

Lufthansa oder British Telecom haben damit begonnen, ihre Marketingvorstandsbereiche aufzulösen und demzufolge Marketing an verschiedenen und dezentralen Schwerpunkten und Positionen im Unternehmen anzusiedeln. Parallel dazu werden das Budget und die Ressourcen für das Marketing enorm reduziert. Strategische Themen und Projekte werden nicht mehr vom Marketing gelöst. Beispiele hierfür sind der Internet-/E-Commerce-Ansatz in vielen Unternehmen oder Elemente der strategischen Vertriebssteuerung wie Sales Force Automation (SFA). In vielen Unternehmen sind bei strategischen Projekten der Strategiebereich oder spezielle Gruppen im operativen und dezentralisierten Vertrieb federführend. Das Marketing erfüllt nur noch eine Kommunikationsfunktion und wirkt nicht als der aktive Treiber solcher Projekte. Mehr und mehr wird die Rolle des klassischen Marketings in allen Unternehmen in Frage gestellt.<sup>23</sup> Hinter dieser Neugestaltung des Marketings steckt ein neues Paradigma, das sich Anfang 90er Jahre herausgebildet hat – Das Konzept des Relationship Marketing, welches sich durch seine Ausrichtung auf Geschäftsbeziehungen und die damit einhergehende langfristige Perspektive auszeichnet.

	Transaktions-Marketing	Relationship-Marketing
Ziel	<ul style="list-style-type: none"> <li>• „To make a sale“, Verkauf (als Abschluss der Kundenbeziehung und Erfolgskriterium)</li> <li>• Bedürfnisbefriedigung (Kunde kauft Werte)</li> </ul>	<ul style="list-style-type: none"> <li>• „To create a customer“, Gewinn eines Kunden (Verkauf als Beginn einer andauernden Beziehung)</li> <li>• Kundenintegration (Interaktive Wertgewinnung)</li> </ul>
Kundenverständnis	<ul style="list-style-type: none"> <li>• Anonymer Kunde</li> <li>• Unabhängigkeit von Verkäufer und Käufer</li> </ul>	<ul style="list-style-type: none"> <li>• Kenntnis individueller Kunden</li> <li>• Interdependenz von Verkäufer und Käufer</li> </ul>
Marketers Aufgabe und Erfolgskriterium	<ul style="list-style-type: none"> <li>• Bewertung auf der Basis von Produkten und Preisen</li> <li>• Fokus auf Neukundengewinnung</li> </ul>	<ul style="list-style-type: none"> <li>• Bewertung auf Basis Problemlösungskompetenz</li> <li>• Fokus auf Wertsteigerung in bestehenden Beziehungen</li> </ul>
Kernaspekte des Austausches	<ul style="list-style-type: none"> <li>• Fokus auf Produkt (Mass Production)</li> <li>• Verkauf als Eroberung</li> <li>• Kundenkontakte als episodische Ereignisse</li> </ul>	<ul style="list-style-type: none"> <li>• Fokus auf Service (Mass Customization)</li> <li>• Kauf als Vereinbarung</li> <li>• Kundenkontakt als kontinuierlicher Prozess</li> </ul>

**Abbildung 1: Transaktions-Marketing vs. Beziehungs-Marketing**

Quelle: in Anlehnung an Wehrli et al. (1994), S. 12. bei Meyer et al. (2001), S. 62

In Abkehr vom dem transaktionsorientierten Marketing, bei dem einzelne Verkaufsabschlüsse maximiert werden und der Kauf als einmalige Transaktion verstanden wird, rückte mit diesem neuen Konzept zum ersten Mal die Kundenorientierung und die damit verbundene Orientierung an den

<sup>23</sup> Vgl. Rapp (2005), S. 21 f.

Kundenbeziehungen in den Mittelpunkt marketingpolitischer Bemühungen (siehe Abb. 1). Das Beziehungsmarketing konnte sich im Zeitverlauf als einer der dominanten Forschungsansätze innerhalb der Marketing Disziplin etablieren und parallel dazu wurde dieses Marketingkonzept insbesondere von der Entwicklung innovativer Informations- und Kommunikationstechnologien (IuK) beeinflusst, was wiederum dazu beiträgt, eine große Anzahl von Kunden individuell anzusprechen und zu betreuen.<sup>24</sup>

### 3.2 GRÜNDE FÜR DIE ENTWICKLUNG DES CRM

In den letzten Jahren haben zahlreiche Entwicklungen die Basis für den Einsatz des CRM gelegt. Die folgende Tabelle zeigt die Änderungen der Marktsituation, des Konsumentenverhaltens und der Kommunikationstechnologien als Auslöser für die Entwicklung des CRM auf.

<b>Marktsituation</b>	<ul style="list-style-type: none"> <li>• Verstärkter Wettbewerbsdruck</li> <li>• Gesättigte Märkte</li> <li>• Globalisierung</li> <li>• Kooperationen und Fusionen</li> <li>• Steigende Markttransparenz durch neue Kommunikationstechnologien</li> <li>• Fehlende USP's</li> </ul>
<b>Konsumentenverhalten</b>	<ul style="list-style-type: none"> <li>• Demographische Verschiebungen und gestiegenes Informationsniveau der Bevölkerung</li> <li>• Preis- versus Qualitätsbewusstsein</li> <li>• Versorgungs- versus Erlebniskauf</li> <li>• Individualisierung und Differenzierung</li> <li>• Wertewandel</li> <li>• Hybrider Verbraucher</li> <li>• Abbau von Dissonanzen (Bedeutung der Theorie der kognitiven Dissonanz)</li> <li>• Aufbau und Sicherung von Beziehungen</li> </ul>
<b>Kommunikationstechnologien</b>	<ul style="list-style-type: none"> <li>• Steigende Akzeptanz und Nutzung neuer Kommunikationstechnologien</li> <li>• Internet</li> <li>• Direktvertrieb</li> <li>• E-Commerce</li> </ul>

**Tabelle 1: Gründe für die Entwicklung des CRM**

Quelle: in Anlehnung an Holland (2004), S. 182.

- *Veränderte Marktsituation:*

Die fortschreitende Globalisierung, die Liberalisierung des Wettbewerbsumfelds, die Auflösung traditioneller Branchengrenzen und die Penetration neuer und aggressiver Wettbewerber im Markt sind nur einige der weitreichenden Veränderungen.<sup>25</sup>

Große Fusionen und Kooperationen der Unternehmen führten zu neuen Marktsituationen mit hohem Wettbewerbsdruck auf gesättigten Märkten. Das

<sup>24</sup> Vgl. Sexauer (2004), S. 19 f.

<sup>25</sup> Vgl. Kreuz/Förster (2001), S. 9.

Internet hat die Globalisierung beschleunigt. Unternehmen können über dieses Medium weltweit ihre Produkte anbieten. Das Internet erleichtert dem Kunden die Informationsgewinnung und den Preisvergleich und steigert die Transparenz der Märkte erheblich.

Die faktischen Vorteile einzelner Produkte werden immer seltener, die Produkte selbst immer austauschbarer in Preis und Leistung. Produktqualität taugt nicht mehr zur Unterscheidung, wir erleben eine Produktflut und eine Produktinflation, welche den Mehrwert des Produktes in den Mittelpunkt der Kaufentscheidungen rückt. Es fehlt vielen Angeboten eine wirkliche Unique Selling Proposition (USP). Eine USP kann somit über eine individuelle Kundenbetreuung geschaffen werden. Wo die Produktqualitäten immer weniger zum qualifizierenden Merkmal werden, kauft der Kunde heutzutage einen vertrauenswürdigen Partner. Die Beziehung zum Anbieter und die Qualität der Kundenbetreuung werden zum entscheidenden Erfolgsfaktor für die Unternehmen.<sup>26</sup>

- *Verändertes Konsumentenverhalten:*

Traditionelle Wertvorstellungen verlieren heute zunehmend an Bedeutung, Wegen des höheren Bildungsstandes und des geänderten Informationsverhaltens der Gesellschaft sind immer mehr Menschen in der Lage, ihr Verhalten an den eigenen Wertvorstellungen auszurichten. Die zunehmende Gruppierung innerhalb der Gesellschaft macht es notwendig, den Kunden immer individueller anzusprechen. Bemühungen um eine möglichst stärkere Individualisierung und Differenzierung des Markets müssen noch forciert werden, damit es durchaus klappt, den Wertvorstellungen des einzelnen Kunden gerecht zu werden.

Des Weiteren sind die potentiellen Kunden jeden Tag mit einer unglaublich großen Menge an Informationen über die unterschiedlichsten Medien konfrontiert und angesichts dieser Informationsüberflutung muss die Botschaft auf das eigentliche Interesse des einzelnen Kunden ausgerichtet sein, um die gewünschte Wirkung zu erzielen. Die Fokussierung auf das Interesse des Kunden, die Zusendung von personalisierten und adressierten Werbemitteln an Kunden und die Kontakthäufigkeit und -regelmäßigkeit dieser Direktmarketing-Aktion verstärken die bestehenden Beziehungen zu Kunden, was sich anschließend auf die Geschäftsbeziehung mit denen positiv auswirkt.

Ein interessantes Ereignis des typischen Konsumentenverhaltens ist, dass viele

---

<sup>26</sup> Vgl. Holland/Huldi/Kuhfuß/Nietsche (2001), S. 14 ff.

Konsumenten nach dem Kauf eines Produktes an der Richtigkeit ihrer Entscheidung zu zweifeln beginnen. Oft empfinden sie den Kauf als eine Störung ihres inneren Gleichgewichts. Die Theorie der kognitiven Dissonanz versucht dieses Phänomen zu erklären, indem sie hypothetisch davon ausgeht, dass Individuen ein dauerhaftes Gleichgewicht ihres kognitiven Systems anstreben. Hier gewinnt der CRM-Ansatz eine große Bedeutung, um dem Kunden zu verdeutlichen, dass er die richtige Kaufentscheidung getroffen hat, welche seinen individuellen Ansprüchen gerecht wird und dass er von dem Unternehmen bestens betreut wird.

- *Kommunikationstechnologien:*

Eine weitere wesentliche Triebkraft für die Entwicklung des CRM ist das Internet mit seiner rasanten Ausbreitung. Es ist bemerkenswert, dass keine andere Kommunikationstechnologie (sei es Telefon, Fernsehen oder Fax) in der Vergangenheit eine so schnelle Verbreitung gefunden hat.

Nach Holland bilden die interaktiven Medien die Basis für die zunehmende Bedeutung des Direktvertriebs und des E-Commerce. Moderne Informations- und Kommunikationstechnologien eröffnen Unternehmen neue Wege bei der Gestaltung von Kundenbeziehungen. Neue Arten der Kommunikation, wie E-Mail oder Instant Messaging, erlauben es den Nutzern, direkt und in erster Linie schnell mit Unternehmen in Kontakt zu treten. Im Internet sind die Informationen ständig abrufbar und jede Änderung ist augenblicklich für jeden sichtbar; es gibt keine Geschäfts- bzw. Ladenöffnungszeiten. Die Kommunikation mit dem Kunden durch E-Mails oder online Chats erfordert Echtzeit-Antwort-Zeiten seitens des Unternehmens oder zumindest, im Fall des E-Mails, sehr schnelle Antworten.

Unter einem finanziellen Aspekt betrachtet, kann das Unternehmen durch die elektronische Kommunikation seine Kommunikationskosten senken. Auf der Kundenseite steigen aber durch die Internetnutzung Kommunikationskosten, welche den Willen für die Aufnahme der gesendeten Nachrichten bei den meisten Kunden einfach beschränken. Falls aber diese Nachrichten intelligent verpackt und mit einem echten Mehrwert versehen werden können, werden Kunden demzufolge auch willig sein, diese aufzunehmen.

Ferner kann Online-Kommunikation mit den übrigen existierenden traditionellen Kommunikationswegen innerhalb einer Kommunikationspolitik kombiniert und abgestimmt werden, wie es bei der Cross Media-Interaktion der Fall ist.<sup>27</sup> So

---

<sup>27</sup> Vgl. Holland/Huldi/Kuhfuß/Nietsche (2001), S. 19.

stehen dem Kunden diverse Kommunikationskanäle (Internet, Zeitung, Fernseher, Telefon, usw.) zur Verfügung, das Unternehmen zu kontaktieren. Der Kunde kann z. B. das Internet nutzen, um schnell und einfach weitere Informationen über gewünschte Produkte und Dienstleistungen zu erhalten und sich interaktiv mit denen zu beschäftigen. Und das CRM strebt an, all diese Kundenkontakte strategisch zu verwalten und eine optimale Gesamtwirkung zu erzielen.

### **3.3 WAS IST CRM EIGENTLICH?**

Der Begriff des Customer Relationship Management (CRM) ist ein neues Schlagwort in der Wirtschaft. Es hat sich jedoch noch keine allgemein anerkannte Definition oder Abgrenzung herausgebildet.

Wenn man nun versucht, den Begriff zu analysieren, fällt auf, dass CRM im Allgemeinen sehr unterschiedlich verstanden und interpretiert wird. Demzufolge wird darunter je nach Umfeld alles von unternehmensweiten strategischen Ansatz über Call Center bis zur Software für die Vertriebsunterstützung verstanden werden.

In Anlehnung an Kotler und Bliemel könnte CRM auch wie folgt definiert werden: „Customer Relationship Management umfasst alle Aktivitäten, die ein Unternehmen gezielt einsetzt, um jeden einzelnen seiner Kunden besser kennenzulernen, wert zu schätzen, zu seiner Zufriedenheit zu bedienen und mit ihm zusammenzuarbeiten.“<sup>28</sup>

Holland definiert Customer Relationship Management als ein strategischer Ansatz, der zur vollständigen Planung, Steuerung und Durchführung aller interaktiven Prozesse mit den Kunden genutzt wird. Laut dieser Definition umfasst CRM das gesamte Unternehmen und den gesamten Kundenlebenszyklus und beinhaltet das Database-Marketing und die entsprechende CRM-Software als Steuerungsinstrument. Eine optimale Kundenorientierung zu erreichen ist das Ziel des Customer Relationship Management.<sup>29</sup>

Diese oben angeführte Erklärung weist darauf hin, dass das CRM kein isoliertes Instrument darstellt, sondern als Philosophie in die Unternehmensprozesse einfließt und weiters, dass „CRM die traditionelle Herangehensweise ‚Herstellen und Verkaufen‘ durch die neue Philosophie ‚Verstehen und Zufriedenstellen‘ ersetzt“<sup>30</sup>.

Schumacher und Meyer bezeichnen CRM als die umfassende Gestaltung der Anbieter-

---

<sup>28</sup> Kotler/ Bliemel (1999), S. 75.

<sup>29</sup> Vgl. Holland/Huldi/Kuhfuß/Nietsche (2001), S. 13.

<sup>30</sup> Kotler (2001), S. 6, zit. in: Knechtenhofer, Oliver (2001), S. 6.

Kunden-Beziehungen eines Unternehmens zu dessen Kunden und Interessenten. Nach diesem Gesichtspunkt wird CRM „als kundenorientierter Ansatz der Unternehmensführung aufgefasst und es beinhaltet die Ausrichtung des unternehmerischen Handelns auf Interessen, Anforderungen und Präferenzen der Kunden, mit dem Ziel der optimalen Gestaltung langfristiger, partnerschaftlicher Kundenbeziehungen, um wirtschaftliche Vorteile zu generieren<sup>31</sup>“.

Rapp erklärt, dass die auftretende Notwendigkeit des Wandels vom traditionellen Marketing hin zum CRM-Ansatz durch das geänderte Kundenbild im Markt entstanden ist.<sup>32</sup> Seiner Meinung nach ist der Kunde heute anspruchsvoller, verfügt über mehr Informationen als zuvor und möchte außerdem Zusatzleistungen entgegennehmen, die über das Produkt hinausgehen bzw. es ergänzen. Diesem gewandelten Kundenbild kann das klassische Marketing nicht gerecht werden, weil es sich produktorientiert auf Märkte konzentriert und gezielt versucht, mit einem Produkt den möglichst größten Marktanteil für sich zu beanspruchen. Obwohl dieses gewandelte Kundenbild schon bekannt ist, reagieren Unternehmen nach wie vor mit den alten Kommunikationsinstrumenten des Massenmarketings.

Seth Godin<sup>33</sup> äußert sich über die Gründe für die Wirkungslosigkeit der klassischen Marketingkommunikation. Klassische Marketing- und Vertriebsmaßnahmen gelangen den Kunden zu jeder Zeit, aber diese stehen nicht dann zur Verfügung, wenn der Kunde solche Nachrichten über Produkte und Dienstleistungen tatsächlich empfangen möchte. Um die Aufmerksamkeit des Kunden zu erlangen, werden die Marketingausgaben erhöht, was wiederum einerseits die zuvor vorhandene Informationsmenge weiters erhöht und andererseits den Wettbewerb um die Ausgaben zwischen Unternehmen beschleunigt.<sup>34</sup>

Godin geht von der Annahme aus, dass die Zeit des Kunden aber, in der er diese Informationen empfangen und verarbeiten kann, immer konstant bleibt; und es stellt sich heraus, dass mit zunehmender Nachrichtenmenge die einzelnen Botschaften wirkungsloser werden. Also je mehr Kommunikation, desto geringer wird die Wirkung, und desto mehr müssen Unternehmen in Kommunikation investieren. Dieses Paradox lockt aber Unternehmen leicht in eine tödliche Falle. Dagegen empfiehlt Godin Permission Marketing, wobei der Kunde aktiv seine Erlaubnis für einen Kontakt zwischen ihm und dem Unternehmen signalisiert. Dieses positive Prinzip ist im CRM-

---

<sup>31</sup> Schumacher/Meyer (2004), S. 19.

<sup>32</sup> Vgl. Rapp (2005), S. 32 f.

<sup>33</sup> Der ehemalige Marketingleiter von Yahoo.

<sup>34</sup> Vgl. Rapp (2005), S. 33.

Ansatz enthalten. So hat der Kunde bereits eine existierende Beziehung zu einem Unternehmen, die ab dem ersten Kontakt des Kunden mit dem Unternehmen ständig weitergepflegt wird, hat Transaktionen erlebt und hat Informationen über sich an das Unternehmen weitergegeben. Somit kennt das Unternehmen den Kunden und weiß, wie es die Bedürfnisse des Kunden befriedigen und dessen Erwartungen erfüllen kann. Mit diesem Wissen ist es möglich, statt der undifferenzierten, auf Kundenakquisition fokussierten Marketingkommunikation eine zielgerichtete und wertbezogene Kommunikation mit dem Kunden zu realisieren.<sup>35</sup>

Weitere Nachteile des Massenmarketings aus Sicht des Unternehmens sind:

- Sowohl bei der Neukundenwerbung als auch bei den bestehenden Kunden findet keine Unterscheidung zwischen potenziell profitablen und nicht profitablen Kunden statt.
- Die Ursachen für die Abwanderung von Kunden werden nicht untersucht.
- Die Marketingmaßnahmen sind auf die Masse gerichtet und die individuellen Kundenbedürfnisse werden nicht erkannt.<sup>36</sup>

### **3.4 GRUNDLAGEN und ABGRENZUNG DES CRM-ANSATZES**

Setzt man sich mit dem Begriff CRM auseinander, lässt sich beobachten, dass CRM häufig auf seine technologische Komponente reduziert wird. CRM wird hierbei mehr oder weniger mit CRM-Systemen gleichgesetzt, deren Aufgabe in der Sammlung und Auswertung von Kundendaten sowie in der Automatisierung kundenbezogener Prozesse liegt.

Es gibt kein Zweifel daran, dass moderne IT-Systeme das Management von Kundenbeziehungen nachhaltig unterstützen können, trotzdem birgt diese starke IT-Orientierung die Gefahr in sich, die notwendigen Rahmenbedingungen im Unternehmen nicht zu berücksichtigen.<sup>37</sup> Daher ist es sinnvoll, eine strategische Ausrichtung statt IT-getriebenem Aktivismus zu fordern.<sup>38</sup>

Der Ausgangspunkt dieses eher betriebswirtschaftlich orientierten CRM-Verständnisses ist die Überlegung, dass IT-Lösungen nur dann ihre Möglichkeiten ausschöpfen

---

<sup>35</sup> Vgl. Rapp (2005), S. 34.

<sup>36</sup> Vgl. Rapp (2005), S. 26.

<sup>37</sup> Vgl. Englbrecht/Hippner/Wilde (2005), S. 11.

<sup>38</sup> Vgl. Homburg/Sieben (2000), S. 472.

können, wenn die entsprechenden Rahmenbedingungen gegeben sind.

Dennoch weder eine einseitige Konzentration auf CRM-Systeme noch eine ausschließliche Fokussierung auf eine betriebswirtschaftliche CRM-Konzeption versprechen folglich eine erfolgreiche CRM-Umsetzung. „Nur die aufeinander abgestimmte Ausgestaltung von kundenorientierter Strategie und kundenorientierten Informationssystemen kann die Potentiale des CRM-Konzepts ausschöpfen.“<sup>39</sup> Vor diesem Hintergrund kann folgende Definition für CRM vorgeschlagen werden:

„CRM ist die kundenorientierte Unternehmensphilosophie, die mit Hilfe moderner Informations- und Kommunikationstechnologien versucht, auf lange Sicht profitable Kundenbeziehungen durch ganzheitliche und differenzierte Marketing-, Vertriebs- und Servicekonzepte aufzubauen und zu festigen.“<sup>40</sup> Nach dieser Definition lassen sich grundsätzlich zwei zentrale Bereiche von CRM ableiten:

Zum einen erfordert CRM den Einsatz von integrierten Informationssystemen (CRM-Systeme). Die integrierten Informationssysteme umfassen die Zusammenführung aller kundenbezogenen Informationen und die Synchronisation aller Kommunikationskanäle. Sie erlauben eine ganzheitliche Abbildung des Kunden („One Face of the Customer“)<sup>41</sup> und eine differenzierte Kundenansprache („One Face to the Customer“)<sup>42</sup>.

Zum anderen steht CRM auch für eine neue kundenorientierte Unternehmensstrategie.<sup>43</sup> Es definiert die Neuorientierung vom funktionalen, klassischen Marketing, das produktorientiert ist und sich auf die Kundenakquisition konzentriert, hin zum übergreifenden, ganzheitlichen Marketing, das auf die Beziehungen zwischen Unternehmen und Kunden fokussiert ist.<sup>44</sup>

Diese Neuorientierung sämtlicher Geschäftsprozesse und Verantwortlichkeiten muss auf den Kunden hin erfolgen.<sup>45</sup> Also hinter CRM steckt mehr als nur eine reine Softwarelösung. Nach Hagen J. Sexauer können vier Forschungsdisziplinen benannt

---

<sup>39</sup> Englbrecht/Hippner/Wilde (2005), S. 11.

<sup>40</sup> Hippner/Wilde (2003a), S. 6.

<sup>41</sup> Die kundenbezogenen Informationen vermitteln dem Unternehmen eine möglichst präzise Einschätzung des Kunden. Die Kommunikationskanäle müssen zusammengefasst und einheitlich verfügbar sein. Weiters muss jeder Mitarbeiter eines Unternehmens Zugriff auf alle Kundendaten haben, vgl. Homburg/Sieben (2000), S. 472 ff.

<sup>42</sup> Nur die Bündelung und Zentralisierung der betrieblichen Informationen ermöglichen ein einheitliches, professionelles Auftreten des Unternehmens und somit eine einheitliche Kundenansprache, vgl. Homburg/Sieben (2000), S. 472 ff.

<sup>43</sup> Vgl. Hippner/Wilde (2003a), S. 6.

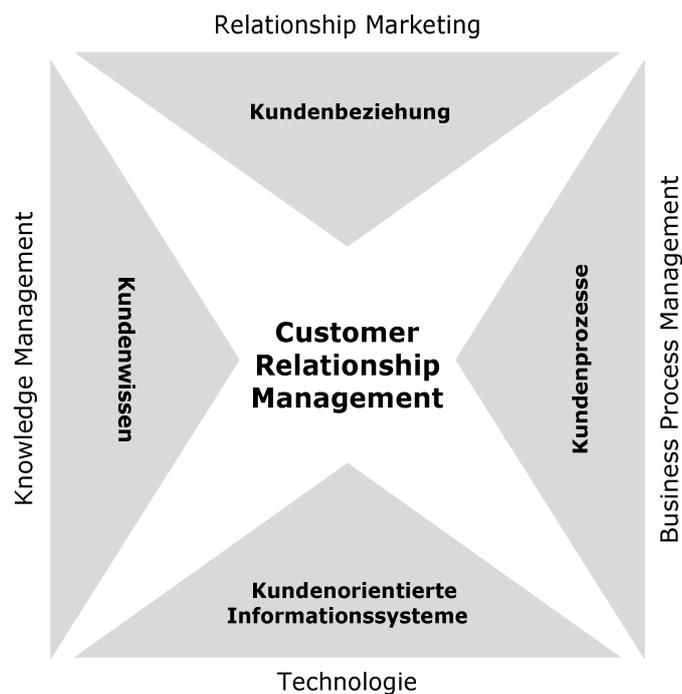
<sup>44</sup> Vgl. Rapp (2005), S. 41.

<sup>45</sup> Vgl. Englbrecht/Hippner/Wilde (2005), S. 11.

werden, die die Entstehung des CRM maßgeblich beeinflusst haben:<sup>46</sup>

- das Relationship Marketing,
- Konzepte der kundenorientierten Informationssysteme,
- das Wissenmanagement,
- das Business Prozess Management (BPM).

Wie auch in der Abbildung 2 leicht zu sehen ist, steht bei dem CRM-Konzept die Fokussierung auf den Kunden im Mittelpunkt. Auf der Grundlage des Kundenwissens, welches durch integrierte kundenorientierte Informationssysteme gesammelt und zur Verfügung gestellt wird, wird es hier angestrebt, langfristig profitable Kundenbeziehungen aufzubauen. Um dies zu ermöglichen, soll die Unternehmensorganisation ganzheitlich auf die Kundenprozesse ausgerichtet werden.



**Abbildung 2: Einflussfaktoren des CRM**

Quelle: Sexauer (2002).

„Customer Relationship Management wurde maßgeblich an den Universitäten von Atlanta, Cranfield und Stockholm entwickelt.“<sup>47</sup> Dieses Konzept ergab sich aus der Tatsache, dass immer mehr Menschen daran zweifelten, ob der traditionelle Marketingansatz hinreichend relevant war. Zur gleichen Zeit gewann langfristige

<sup>46</sup> Vgl. Sexauer (2002).

<sup>47</sup> Rapp (2005), S. 41.

Kundenbeziehungen immer mehr an Bedeutung. Im Zuge der Vorbehalte gegen die Relevanz des klassischen Marketingansatzes wurde eine Anzahl von Unternehmen untersucht und auf Basis der daraus resultierenden Ergebnisse ein umfassender CRM-Ansatz erarbeitet.<sup>48</sup> Der Kerngedanke des Ansatzes ist die Steigerung des Unternehmens- und Kundenwerts durch das systematische Management der bestehenden Kunden.<sup>49</sup> Gleichzeitig bietet dieser Ansatz aber auch die optimale Plattform zur gezielten Akquisition. Das heißt, mittels CRM werden neue Geschäftspotentiale in bestehenden Kunden identifiziert und auf Basis dieser Erkenntnisse können attraktive Neukunden gewonnen werden.<sup>50</sup>

Das Hauptziel des CRM ist die langfristige Sicherung der Existenz der Kunden.<sup>51</sup> Dieses Anliegen kann durch Konzepte wie Interaktion über verschiedene Kanäle, Loyalitätsmaßnahmen, Personalisierung und Dialogmanagement über Call Center oder Internet und Kundenlebenswertbetrachtung im Controlling erreicht werden.<sup>52</sup> CRM fokussiert sich dabei zusätzlich zur Produktdifferenzierung vor allem auf die Prozessdifferenzierung, d. h. wie die Prozesse des Kunden verwaltet bzw. gemanagt werden.<sup>53</sup> „Durch die Analyse der wertschöpfenden Kundenprozesse werden Kunden bei ihrer Wertschöpfung unterstützt.“<sup>54</sup>

Die kundenprozessorientierte Perspektive und der langfristige Unternehmenswert generierende Ansatz sind wichtige Merkmale des CRM-Ansatzes, wodurch sich CRM grundsätzlich von anderen Kundenbindungsmodellen unterscheidet.<sup>55</sup> Die folgenden vier Faktoren geben einen Überblick über den Perspektivenwechsel im Rahmen des Customer Relationship Management und sind weiters bestimmend für den Unterschied zwischen CRM und klassischen Marketing-Ansätzen:

1. *Perspektivenwechsel von der Produkt- bzw. Unternehmenssicht zur Kundensicht:*

Das klassische Marketing hat einen produktorientierten Geschäftsfokus, strebt die Optimierung der internen Unternehmensprozesse an (Führung, Arbeitsabläufe und Kommunikation) und versucht dem Kunden Produkte zum bestmöglichen Preis-/Leistungsverhältnis zu liefern. Die unternehmerische Strategie erfolgt entsprechend von innen nach außen, was die grundlegende Vorgehensweise im

---

<sup>48</sup> Vgl. Rapp (2005), S. 41.

<sup>49</sup> Vgl. Rapp (2005), S. 41.

<sup>50</sup> Rapp/Decker (2000), S. 2.

<sup>51</sup> Vgl. Rapp/Decker (2000), S. 2.

<sup>52</sup> Vgl. Rapp/Decker (2000), S. 2 und Rapp (2005), S. 41

<sup>53</sup> Vgl. Rapp/Decker (2000), S. 2.

<sup>54</sup> Rapp/Decker (2000), S. 2.

<sup>55</sup> Vgl. Rapp (2005), S. 41

Total Quality Management repräsentiert; nämlich vom Unternehmen aus gesehen den Kunden zu erobern. Im Customer Relationship Management wird dagegen von der Fragestellung ausgegangen, wie Unternehmen den Wertschöpfungsprozess des Kunden mit Blick auf eine möglichst lange dauernde Verbindung unterstützen können. Somit wechselt die Perspektive von der unternehmerischen Seite auf die Kundenseite; anders formuliert, man betrachtet jetzt von außen nach innen. Das Ziel des Customer Relationship Management ist nicht mehr wie beim TQM, die Unternehmensprobleme zu lösen, sondern auf die Kundenprobleme einzugehen und darauf Antworten zu geben. Somit legt CRM einen hohen Wert auf die Bedeutung der persönlichen Nutzensvorstellung des Kunden.

## 2. *Flexible Angebote nach Maß für den Kunden:*

Aus Gründen der ständig wachsenden Kundenansprüche wird es immer wichtiger, durch maßgeschneiderte Produkte oder Dienstleistungen individuelle Bedürfnisstrukturen der Kunden anzusprechen. Demgemäß tritt an die Stelle des Massenprodukts die individuelle Maanfertigung (Mass Customization). Piller und Schoder sehen die Varietät von Produktionen als Ziel des Mass Customization, wodurch es für den Kunden möglich ist, exakt das zu finden, was er wünscht und dafür auszugeben bereit ist.<sup>56</sup>

Dieser Begriff wurde 1987 von Davis geprägt: „Mass Customization of markets means that the same large number of customers can be reached as in mass markets of the industrial economy, and simultaneously they can be treated individually as in the customized markets of pre-industrial economies“.<sup>57</sup> Als Beispiel für Mass Customization kann die individuelle PC-Fertigung bei Dell Computer genannt werden. Der vom Kunden auf der Webseite des Unternehmens persönlich konfigurierte PC wird dann genau so zusammengesetzt, dass er exakt diesen Kundenwünschen entspricht.

## 3. *Kunden werden gemäß ihrer Profitabilität behandelt:*

Im traditionellen Marketing, bzw. bei der Denkweise des TQM wird allen Kunden die gleiche Aufmerksamkeit geschenkt. Aber Kunden haben keinesfalls immer dieselbe Attraktivität. Manche Kunden sind wertvoller als andere. Im Rahmen einer Finanzmarktstudie, die von IBM und der Economist Intelligence Unit im Jahr 2006 realisiert wurde, hat es sich ebenso herauskristallisiert, dass rund zwanzig Prozent

---

<sup>56</sup> Piller/Schoder 1999, S. 6.

<sup>57</sup> Davis (1987), S. 169, zit. in: Piller/Schoder 1999, S. 6

der Kunden in der Regel achtzig Prozent oder mehr des Gewinns generieren.<sup>58</sup>

Das Problem besteht darin, dass es für die Unternehmen nicht leicht festzustellen ist, welche Kunden zu den profitablen zwanzig Prozent zählen. Viele Firmen fokussieren sich allzu oft einfach auf jene Kunden, die am häufigsten Geschäfte tätigen oder den höchsten Umsatz generieren, ohne die tatsächliche Rentabilität der einzelnen Kunden zu kennen.<sup>59</sup> Dadurch ist es leicht zu begründen, warum der Verkauf eines Services durch eine kostenintensive persönliche Beratung nur den Kunden vorbehalten sein sollte, die auch nachhaltig zum Unternehmenserfolg beitragen könnten. Demzufolge werden innovative Unternehmen in Zukunft rationalisieren, um sich vermehrt auf die Pflege der Beziehungen zu den profitablen Kunden zu konzentrieren, die auf längere Sicht den höchsten Gesamtertrag generieren.<sup>60</sup>

#### 4. Stellenwert langfristiger Kundenbeziehungen:

„Langfristige Beziehungen zu profitablen Kundengruppen steigern messbar die Gewinne.“<sup>61</sup> Viele Unternehmen kennen leider ihre Kunden nicht. Folglich sind sie weder in der Lage, ihre Profitabilität zu messen, noch verfügen sie über gewisse Kenntnisse, welchen Gewinn das Unternehmen im wirtschaftlichen Sinne durch einen einzelnen Kunden langfristig herausziehen könnte.<sup>62</sup>

Die Ausgestaltung von Geschäftsbeziehungen darf sich nicht nur an der kurzfristigen Profitabilität eines Kunden orientieren, da es sich durchaus rechnen kann, sich auch um solche Kunden zu kümmern, die dem Unternehmen zunächst einmal nur Verluste bringen.<sup>63</sup> Hierfür ist jedoch kritisch, ob der Kunde über ein hohes zukünftiges Potenzial verfügt.

Obwohl z. B. Studenten üblicherweise - kurzfristig betrachtet - eine nur geringe Kaufkraft aufweisen, rechnen Banken bei den Geschäftsbeziehungen zu ihnen immer damit, dass sie sich zukünftig zu einer finanzstarken und somit profitablen Kundengruppe entwickeln können. (z. B. „das Betreiben eines Girokontos für einen Studenten – in der Erwartung einer profitablen Geschäftsbeziehung bei Eintritt in

---

<sup>58</sup> Vgl. Artikel im Internet über die IBM Finanzmarktstudie 2006: „The Trader is Dead, Long Live the Trader“, Author: Jochen Reinhardt, auf URL: <http://www.ibm.com/news/ch/de/2006/08/30.html>, [Abruf: 30-03-2007].

<sup>59</sup> Vgl. ebd.

<sup>60</sup> Vgl. ebd.

<sup>61</sup> Rapp (2002): „CRM – und was hat der Kunde davon?“ - Artikel erschien am 03/2002 in CRM Guide, auf URL: <http://reinholdrapp.com/d/archiveintrag.php?auswahl=8>, [Abruf: 30-11-2007].

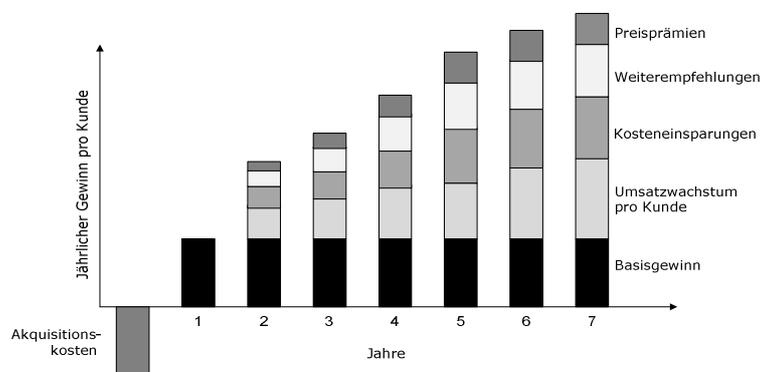
<sup>62</sup> Vgl. Rapp (2005), S. 44.

<sup>63</sup> Vgl. Hippner (2006), S. 25.

das Berufsleben.“<sup>64</sup>)

Das obige Beispiel verdeutlicht speziell die Wichtigkeit der Profitabilität eines Kunden und bezieht sich nicht nur auf die Intensität einer Geschäftsbeziehung, sondern es berücksichtigt auch ihre Dauer.

Nach einer Untersuchung von Reichheld ergab sich, dass mit zunehmender Dauer der Kundenbeziehung auch der daraus resultierende Gewinn ansteigt (siehe Abb. 3). Anfänglich wird es zum Aufbau der Kundenbeziehung viel investiert, und es fallen dann laufend zusätzliche Kosten für deren Erhalt und Ausbau an. Wie die Untersuchung aber graphisch veranschaulicht, rechnen sich schließlich mit zunehmender Dauer all diese mit hohen Kosten verbundenen Bemühungen um den einzelnen Kunden. Deshalb versuchen Unternehmen, durch CRM den Kunden stärker an sich zu binden und weitere Geschäftsabschlüsse zu erzielen. Langfristig decken die erzielten Einsparungen und zusätzlichen Gewinne die Akquisitionskosten ab und führen zu profitablen Kunden. Außerdem lässt sich an der Abbildung erkennen, dass der Nutzen einer Kundenbeziehung nicht nur von der eigentlichen Transaktion (z. B. Umsatzwachstum) abhängt, sondern auch weiche Faktoren, wie z. B. Weiterempfehlungen auch das Nutzenwachstum dieser Kundenbeziehung beeinflussen.<sup>65</sup>



**Abbildung 3: Monetäre Nutzenentwicklung im Verlauf einer Kundenbeziehung**

Quelle: Reichheld (1997), S. 52.

Die Profitmaximierung ist das verfolgte Ziel des CRM-Ansatzes und stellt den zentralen Orientierungspunkt bei der Ermittlung des Kundenwertes aus Unternehmenssicht dar.<sup>66</sup> Vor diesem Hintergrund werden nur solche Kundengruppen sorgsam betreut, mit denen gegenwärtig bzw. zukünftig ein

<sup>64</sup> Rosemann/Rochefort/Behnck (1999), S. 109, zit. in: Hippner (2006), S. 25 f.

<sup>65</sup> Vgl. Hippner (2006), S. 26

<sup>66</sup> Vgl. Hippner (2006), S. 29.

positiver Deckungsbeitrag zu erzielen möglich ist.

Zur Ermittlung längerfristiger Kundenwerte kann das Konzept des Customer Lifetime Value herangezogen werden. „Das Konzept des Customer Lifetime Value (CLV) überträgt Prinzipien der dynamischen Investitionsrechnung auf die Kundenbeziehung“<sup>67</sup>, mit deren Hilfe die Erfolgswahrscheinlichkeit einer Investition in Kundenbeziehungen abgeschätzt werden kann.

$$KW = \sum_{t=0}^n \frac{e_t - a_t}{(1+i)^t} = e_0 - a_0 + \frac{e_1 - a_1}{(1+i)} + \frac{e_2 - a_2}{(1+i)^2} + \dots + \frac{e_n - a_n}{(1+i)^n}$$

$e_t$  = (erwartete) Einzahlungen aus der Geschäftsbeziehung in der Periode t

$a_t$  = (erwartete) Auszahlungen aus der Geschäftsbeziehung in der Periode t

$i$  = Kalkulationszinsfuß zur Abzinsung auf einen einheitlichen Referenzzeitpunkt

$t$  = Periode (t=0, 1, 2, 3, ..., n)

$n$  = Dauer der Geschäftsbeziehung

**Abbildung 4: Berechnung des Kapitalwerts einer Kundenbeziehung**

Quelle: in Anlehnung an Homburg/Daum 1997, S. 402

CLV orientiert sich an der Kapitalwertmethode<sup>68</sup>, wonach der Wert eines Kunden sich aus den diskontierten, dem Kunden direkt zurechenbaren Geldströmen (Einzahlungen und Auszahlungen), die während der gesamten Geschäftsbeziehung erfolgen, errechnet (siehe Abb. 4). „Eine derartige Berechnung erlaubt insbesondere, Neukunden, die anfänglich nur kostendeckend oder gar unter Inkaufnahme von Verlusten gewonnen werden konnten, über die gesamte - prognostizierte und vertraglich gesicherte - Dauer der Verbindung auf ihre Rentabilität hin zu prüfen.“<sup>69</sup>

Der Begriff „CRM“ (Customer Relationship Management) wird in der Literatur häufig mit „Beziehungsmanagement“ (Relationship Management), „Beziehungsmarketing“ (Relationship Marketing) sowie „Kundenbindungsmanagement“ (Customer Retention Management) gleichgesetzt.

Diese Begriffe werden in der Praxis nicht sauber voneinander abgegrenzt. Aus diesem Grund ist es sinnvoll, die Abgrenzung dieser Begriffe zu erläutern sowie ihre Beziehung

<sup>67</sup> Meyer/Weingärtner/Döring (2001), S. 75.

<sup>68</sup> Die Kapitalwertmethode als Verfahren der dynamischen Investitionsrechnung basiert auf dem Prinzip, dass Zahlungen in der Zukunft weniger wert sind als gleich hohe gegenwärtige Zahlungen, zukünftige Ein- und Auszahlungen deshalb mit einem Kalkulationszinsfuß über die Anzahl der betrachteten Perioden abzuzinsen sind, vgl. Schirmeister/Kreuz (2003), S. 347.

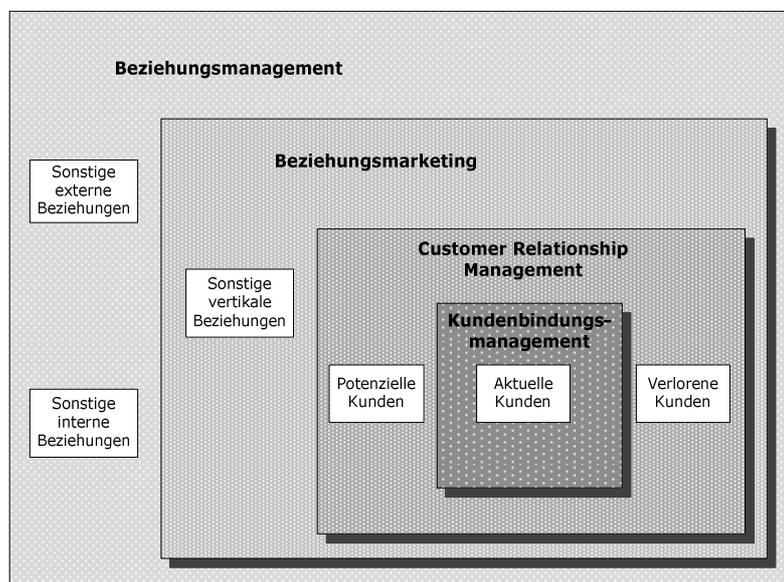
<sup>69</sup> Meyer/Weingärtner/Döring (2001), S. 75.

zueinander darzulegen (siehe auch Abb. 5):

Diller definiert Beziehungsmanagement als „die aktive und systematische Analyse, Selektion, Planung, Gestaltung und Kontrolle von Geschäftsbeziehungen im Sinne eines ganzheitlichen Konzeptes von Zielen, Leitbildern, Einzelaktivitäten und Systemen<sup>70</sup>“. Es beschränkt sich nicht auf die Kundenbeziehungen, sondern wird als ein umfassendes Konzept verstanden, das die horizontalen (Z. B. Vertriebsgemeinschaften), vertikalen (Z. B. Zuliefererbeziehungen), lateralen (Z. B. Beziehungen zu Behörden) oder unternehmensinternen Beziehungen (Z. B. zum Personal) einbezieht.<sup>71</sup>

Berry entwickelte im Jahr 1983 das Konzept des Relationship Marketing und definierte dieses Konzept wie folgt: „Relationship Marketing is attracting, maintaining and (...) enhancing customer relationships.<sup>72</sup>“

Beim Beziehungsmarketing steht die Kundenseite im Vordergrund, aber bezieht die Beziehungen zu den Lieferanten auch mit ein.<sup>73</sup> Das Kundenbindungsmanagement fokussiert sich ausschließlich auf die aktuellen, bereits bestehenden Kundenbeziehungen. Es schließt die Gewinnung von Neukunden sowie die Rückgewinnung abgewanderter Kunden nicht mit ein.<sup>74</sup>



**Abbildung 5: Abgrenzung des CRM von verwandten Begriffen**

Quelle: Hippner (2006), S. 20.

<sup>70</sup> Diller, H. (1995), S. 442.

<sup>71</sup> Vgl. Diller/Kusterer (1988), S. 212.

<sup>72</sup> Berry (1983), S. 25.

<sup>73</sup> Vgl. Berry (1983), S. 25 und Köhler (2001), S. 82.

<sup>74</sup> Vgl. Hippner (2006), S. 19.

Das Kundenbeziehungsmanagement CRM hat sich aus dem Beziehungsmarketing abgeleitet, wodurch es eine große Ähnlichkeit mit dem Beziehungsmarketing aufweist.

Im Gegensatz zum Beziehungsmarketing beschränkt sich CRM jedoch ausschließlich auf die Gestaltung der Beziehung zum Kunden.<sup>75</sup> Ausschlaggebend sind seine kundenprozessorientierte Perspektive und sein langfristig Unternehmenswert generierender Ansatz.<sup>76</sup>

### **3.5 ZIELE DES CUSTOMER RELATIONSHIP MANagements**

Laut Deutschem Dialogmarketing Verband (DDV) ist Customer Relationship Management „...ein ganzheitlicher Ansatz zur Unternehmensführung. Er integriert und optimiert abteilungsübergreifend alle kundenbezogenen Prozesse in Marketing, Vertrieb, Kundendienst sowie Forschung und Entwicklung. Dies geschieht auf der Grundlage einer Datenbank mit einer entsprechenden Software zur Marktbearbeitung und anhand eines vorher definierten Verkaufsprozesses. Die Zielsetzung von CRM ist die Schaffung von Mehrwerten auf Kunden- und Lieferantenseite im Rahmen der Geschäftsbeziehungen.“<sup>77</sup>

Nach Prof. Robert Shaw<sup>78</sup> soll es bei der Zielorientierung – von Unternehmensseite aus betrachtet - versucht werden, eine optimale Balance zwischen den Kosten der aufgebauten Kundenbeziehung und der Erfüllung aller Kundenbedürfnisse zu erreichen, um den maximalen Gewinn zu generieren.<sup>79</sup>

Das Customer Relationship Management verfolgt eine Reihe von unterschiedlichen Zielen durch die Intensivierung der Beziehung zwischen Unternehmen und Kunden.

Der Kerngedanke jedes CRM-Konzepts besteht aber grundsätzlich darin, den Unternehmens- und Kundenwert durch das systematische Management der Kundenbeziehungen zu steigern. Dabei soll die Beziehung zwischen Kunden und Unternehmen zu beiderseitigem Vorteil optimal und vor allem langfristig gestaltet werden.

---

<sup>75</sup> Vgl. Hippner (2006), S. 19.

<sup>76</sup> Vgl. Rapp (2005), S. 41

<sup>77</sup> Offizielle Definition des DDV (Deutscher Dialogmarketing Verband e. V.) im Internet, auf URL: <http://www.ddv.de/index.php?id=100>, [Abruf: 01-10-2008]

<sup>78</sup> Gründer des Shaw Consulting und Author des Buchs „Measuring and Valuing Customer Relationships: How to Develop the Measures That Drive Profitable CRM Strategies“.

<sup>79</sup> Vgl. eigene Übersetzung aus dem englischen Text „How leading experts define CRM?“ im Internet, erschien auf der Webseite von The European Centre for Customer Strategies, o. V., o. J.; auf URL: <http://www.eccs.uk.com/VD25/default.asp?CID=421&PSID=439>, [Abruf: 05-04-2007].

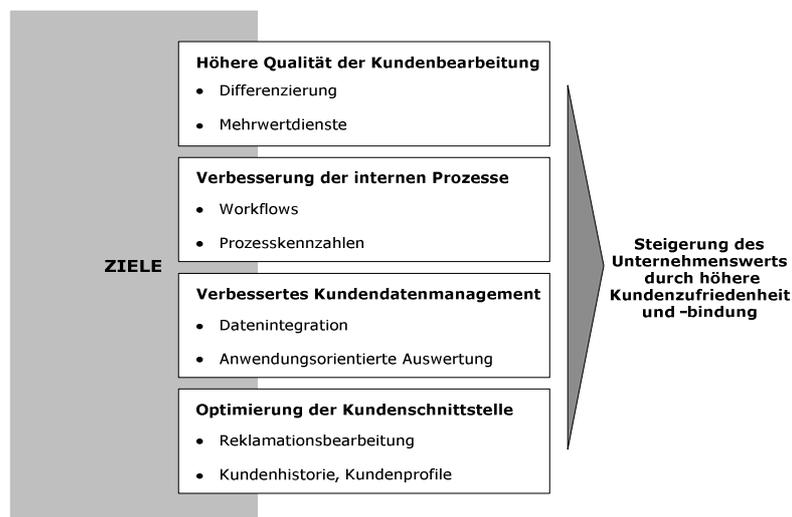
Um den Erfolg von CRM langfristig gewährleisten zu können, sollen folgende strategische Fragenstellungen beantwortet werden<sup>80</sup>:

- Welche Kunden sind am profitabelsten, bezogen auf die Dauer der Kundenbeziehung?
- Welche Leistung muss den Kunden angeboten werden, dass sie langfristig verbunden bleiben?
- Wie können ähnliche, neue, profitable Kunden gewonnen werden, ebenfalls mit dem Ziel einer dauerhaften Verbindung?

Die Ziele, die mit CRM verfolgt werden sollen, lassen sich daraus wie folgt zusammenfassen<sup>81</sup>:

- Die Erhöhung des Share of Wallets<sup>82</sup> des einzelnen Kunden
- Die Optimierung der Kundenbeziehung unter ökonomischen Gesichtspunkten
- Die Gewinnung hochwertiger Neukunden

Um diese genannten Ziele zu erreichen und damit den Unternehmenswert durch höhere Kundenzufriedenheit und -bindung zu steigern, sind folgende untergeordnete Ziele zu verwirklichen (siehe Abb. 6):



**Abbildung 6: Untergeordnete Ziele des CRM-Konzepts**

Quelle: In Anlehnung an Dangelmaier/Helmke/Uebel (2004), S. 5

<sup>80</sup> Vgl. Rapp (2005), S. 45

<sup>81</sup> Vgl. Rapp (2005), S. 45.

<sup>82</sup> Als Share of Wallet wird der Anteil bezeichnet, den ein Kunde von seinen Gesamtausgaben für einen bestimmten Bereich bei einem Anbieter angibt, vgl. Rapp (2005), S. 45.

Zunächst muss die Qualität der Kundenbearbeitung durch eine Differenzierung und durch ein breites Angebot an Mehrwertdiensten für Kunden erhöht werden. Zudem sind die internen und externen Kundenbearbeitungsprozesse so zu optimieren, dass die Schnittstelle zum Kunden optimal gestaltet ist.<sup>83</sup> Ebenso müssen neue Technologien eingesetzt werden, um das Kundendatenmanagement und die Schnittstellen zu den Kunden zu verbessern. Des Weiteren ist eine Neuausrichtung sämtlicher Geschäftsprozesse und Verantwortlichkeiten in den Marketing-, Vertriebs- und Kundenserviceabteilungen auf den Kunden hin zu erfolgen.

Zusammenfassend lassen sich vier grundsätzliche Prinzipien bzw. Zielsetzungen von CRM identifizieren,<sup>84</sup> welche im Folgenden kurz erläutert werden. Diese sind Profitabilität, Differenzierung, Langfristigkeit und Integration.

#### 1. *Profitabilität:*

Das wichtigste Ziel des CRM ist die Erhöhung der Profitabilität von Kundenbeziehungen, weil profitable Kundenbeziehungen unmittelbar auf den Gesamterfolg des Unternehmens wirken. Eine Erhöhung der Profitabilität wird durch eine verstärkte Kundenbindung und die damit einhergehende längerfristige Kundenbeziehung erreicht. Eine hohe Kundenzufriedenheit aufgrund von erbrachtem hochwertigem Service, ein gutes Unternehmensimage und zielgerichtete Marketingmaßnahmen bilden die Grundlage zur Verbesserung der Kundenbeziehung und damit zur Verstärkung der Kundenbindung.

Empirische Studien belegen, dass es mindestens fünfmal teurer ist, einen Neukunden zu gewinnen als einen bestehenden Kunden zu pflegen.<sup>85</sup> Dies macht deutlich, warum die verstärkte Kundenbindung und damit die Lebensdauer einer Kundenbeziehung mehr als die Kundengewinnung im Mittelpunkt unternehmerischen Handelns stehen soll.

Außerdem nimmt die Qualität des Wissens über den Kunden mit zunehmender Dauer einer Geschäftsbeziehung zu, was demzufolge zu einer gezielteren Kundenansprache und besseren Befriedigung der Kundenbedürfnisse führt, wodurch die Wahrscheinlichkeit erneuter Geschäftsabschlüsse steigt.<sup>86</sup> Von befriedigten Kunden geht zudem eine gute Referenzfunktion aus, was wiederum zur Kundengewinnung und damit zur Umsatzsteigerung beiträgt. Ebenso tragen

---

<sup>83</sup> Vgl. Dangelmaier/Helmke/Uebel (2004), S. 5.

<sup>84</sup> Vgl. Kantsperger (2002), S. 1-2. und Hippner/Wilde (2003a), S. 7-12.

<sup>85</sup> Vgl. Töpfer/Bauer/Beckett/Bielecki (1999), S. 59 und Schulze (2002), S. 62.

<sup>86</sup> Vgl. Schulze (2002), S. 62.

Kostensenkungen zur Profitabilitätserhöhung bei.<sup>87</sup> Durch die Automatisierung in den Bereichen Sales und Service werden Transaktionskosten und -zeiten reduziert und Geschäftsabläufe verkürzt. Beispiele von kostensenkenden Impulsen sind die rasche und effiziente Bearbeitung großer Datenmengen, die automatische Abwicklung manueller Tätigkeiten (Z. B. das Selektieren von Adressen für Mailings, die Nachbearbeitung eines Kundenbesuchs), die effiziente Bereitstellung von Unterlagen zur Entscheidungsunterstützung oder die Beseitigung monotoner Routinetätigkeiten.<sup>88</sup>

## 2. *Differenzierung und Personalisierung:*

Die Differenzierung der Kundenbeziehung stellt eine zentrale Forderung des CRM-Konzepts dar. Diese Differenzierung soll sowohl in der Leistungs- als auch in der Kommunikationsebene erfolgen, d. h. um die Kunden möglichst umfassend zu betreuen, müssen sowohl die Produkte und Dienstleistungen als auch der Dialog mit dem Kunden differenziert zugeschnitten werden.<sup>89</sup> Aus Kundensicht bedeutet dies, dass die Lebenssituation und die Eigenheiten des Kunden wie z. B. seine Vorlieben oder Gewohnheiten berücksichtigt werden.<sup>90</sup>

Aus Unternehmenssicht verfolgt die Differenzierung eine ganz pragmatische Zielsetzung. Trotz einer höheren Chance zur langfristigen Kundenbindung, bedeutet sie für das Unternehmen auch eine Abweichung von der Norm und damit zusätzliche Kosten. Es gilt diese zusätzlichen Kosten mit der zu erwartenden Profitabilität des einzelnen Kunden abzugleichen.<sup>91</sup>

Die Maximierung der Kundenorientierung ist zwar ein wichtiger Aspekt des CRM-Konzepts, stellt sie sich jedoch nicht als eine vordringliche Zielsetzung dar; sondern es gilt angesichts der damit verbundenen Kosten, den Grad der Kundenorientierung an den Wert des einzelnen Kunden anzupassen.<sup>92</sup> So kann z. B. eine Bank die intensivere, speziellere Betreuung sehr vermögender Kunden ausschließlich für notwendig halten, während sie für weniger gut situierte Kunden eher einen standardisierten Beratungsservice zur Verfügung stellt, bei dem die Betreuungsintensität des einzelnen Kunden weitaus geringer ausfällt.

---

<sup>87</sup> Vgl. Schulze (2002), S. 63.

<sup>88</sup> Vgl. Schulze (2002), S. 64.

<sup>89</sup> Vgl. Hippner/Wilde (2003a), S. 9.

<sup>90</sup> So besitzt ein 18-jähriger Berufsanfänger sicherlich andere Bedürfnisse und Erwartungen als eine junge Familie oder ein Rentnerhepaar und möchte demzufolge ein spezielles Angebot entgegennehmen, das seinen Wünschen und Forderungen besser entspricht, vgl. Hippner/Wilde (2003a), S. 9.

<sup>91</sup> Vgl. Hippner/Wilde (2003a), S. 9.

<sup>92</sup> Vgl. Hippner/Wilde (2003a), S. 9 und Kreuz/Förster (2001), S. 14.

In Märkten mit einer enormen Anzahl von Kunden ist es aber nicht so leicht und jederzeit möglich, eine echte personalisierte Beziehung zu den Kunden aufzubauen, diese individuell zu pflegen und demgemäß jedem einzelnen Kunden ein individuelles Angebot zu unterbreiten. Um diese Herausforderung anzugehen, werden die Kunden möglichst fein segmentiert, wodurch dann die Individualisierung approximativ erfolgen kann.<sup>93</sup> Hier bilden nicht die individuellen Besonderheiten des einzelnen Kunden, sondern fein differenzierte Kundensegmente die Grundlage für Kommunikations- und Leistungssteuerung. Diese Segmentierung erstreckt sich dann nicht mehr auf nur fünf oder zehn verschiedene Marktsegmente, wie bei der klassischen Marktsegmentierung der Fall ist, sondern – aufbauend auf einer aussagekräftigen Kundendatenbank – auf mehrere hundert unterschiedliche Marktsegmente.

### 3. *Langfristigkeit:*

Eine weitere Zielsetzung von CRM liegt darin, langfristige Kundenbeziehungen aufzubauen und zu festigen.<sup>94</sup> Daraufhin, nicht mehr die Kundenneugewinnung, sondern die kontinuierliche Pflege bereits bestehender Kundenbeziehungen nimmt eine zentrale Rolle ein, um eine dauerhafte Kundenbindung zu erzielen.<sup>95</sup>

Wie bereits erwähnt, es ist ziemlich kostenintensiver und zwar um ein Vielfaches teuer, neue Kunden zu gewinnen, statt bestehende zu halten.<sup>96</sup> Vor diesem Hintergrund erhoffen sich Unternehmen, ihren primären Gewinn aus dem Aufbau einer hohen Kundenbindung im Rahmen eines CRM zu erzielen. So unterstützt die zuvor im Kapitel 3.4 erwähnte empirische Untersuchung von Reichheld Unternehmen, indem sie offensichtlich zeigt, dass sich die hohen Investitionen am Anfang zum Aufbau der Kundenbeziehung sowie die laufenden Kosten für deren Erhalt und Ausbau mit zunehmender Dauer rechnen. Die Gründe hierfür sind:

- Preisprämien: Die Preiselastizität nimmt mit steigender Kundenbindung ab, so dass Kunden eher bereit sind, für ein vertrautes bzw. für gut empfundenes Service aller Art einen höheren Preis zu bezahlen.<sup>97</sup>
- Weiterempfehlungen: Viele Kunden teilen meist ihre Erfahrungen anderen potentiellen Kunden mit. Erfahrungen mit gutem oder

---

<sup>93</sup> Vgl. Hippner/Wilde (2003a), S. 10.

<sup>94</sup> Vgl. Hippner/Wilde (2003a), S. 10.

<sup>95</sup> Vgl. Hippner/Wilde (2003a), S. 10.

<sup>96</sup> Vgl. Stojek (2000), S. 42 und Schulze (2002), S. 62.

<sup>97</sup> Vgl. Hippner/Wilde (2003a), S. 11.

schlechtem Service, mit tollen oder miesen Dienstleistungen, mit großartiger Kundenorientierung oder misshandlungem Standard.

Auf der einen Seite haben diese Weiterempfehlungen, sowohl negativ als auch positiv, einen wichtigen Einfluss auf Kaufentscheidungen von Kunden, denn „91 percent of unhappy customers will never again buy from a company dissatisfying them, and will communicate their dissatisfaction to at least nine other people<sup>98</sup>“.

Auf der anderen Seite neigen loyale, zufriedengestellte Kunden dazu, positive Mund-zu-Mund-Propaganda zu betreiben.<sup>99</sup> Ihr Einfluss auf andere bewirkt ein positives Unternehmensimage und dadurch steigt der Wert des Unternehmens in den Augen der potentiellen Kunden. Nach Reichheld sind gerade solche Kunden, die aufgrund von Empfehlungen an ein Unternehmen herantreten, sind meist wertvoller als Kunden, die auf Werbung oder Preisaktionen reagieren.<sup>100</sup>

- Kosteneinsparungen: Langfristig betrachtet, ist eine kontinuierliche Kundenorientierung kostensparend als eine einmalig bzw. kurzfristig geplante Transaktionsorientierung.<sup>101</sup>

Die über Jahre lang gesammelten Kundeninformationen können effizienter bearbeitet werden und zu einer Reduzierung von Streuverlusten bei Marketingaktionen beitragen. Aus den gewonnenen Informationen können zielgerichtete Kundenprofile gemustert werden, die eine differenzierte Kundenansprache ermöglichen und somit die Kundenzufriedenheit erhöhen.

- Umsatzwachstum: Neben der Kosteneinsparung bzw. -senkung können Unternehmen bei Anwendung des CRM auch ihren Umsatz steigern. Dieser steigt, wenn ein Kunde, der mit einem Produkt oder einer Dienstleistung sowie mit den damit verbundenen Zusatzleistungen schon zufrieden ist, über die Zeit mit dem Unternehmen zusätzliche Geschäftsabschlüsse tätigt.

Diese Wieder- und Folgekäufe, wobei sich *Cross/Up Selling*-

---

<sup>98</sup> Vavra (1995), S. 14, zit. in: Schulze (2002), S. 60.

<sup>99</sup> Vgl. Diller (1996), S. 82.

<sup>100</sup> Vgl. Reichheld (1997). S. 65.

<sup>101</sup> Vgl. Hippner/Wilde (2003a), S. 11.

Potentiale<sup>102</sup> leichter ausschöpfen lassen, tragen zur Umsatzsteigerung bei, indem sie Kunden dauerhaft an das Unternehmen binden.

#### 4. *Integration:*

Jeder Kunde hinterlässt an den unterschiedlichsten Stellen im Unternehmen seine Spuren. Die Informationen über einen Kunden werden meistens in mehreren verschiedenen Datenbanken gespeichert. Diese dezentrale Verteilung von gespeicherten Kundendaten führt aber an den einzelnen Kontaktpunkten zwischen dem Kunden und Unternehmen (Customer Touch Points) zu einer unzureichenden Kenntnis über die spezifischen Bedürfnisse und Erwartungen des Kunden. Aus diesem Grund muss alle kundenrelevante Informationen aus den Bereichen Marketing, Vertrieb, Service usw. integriert werden, um dadurch ein möglichst scharfes und umfassendes Bild vom Kunden und seiner Geschäftsbeziehung zu erhalten (One Face of the Customer).<sup>103</sup>

Die Voraussetzung hierfür ist das Vorhandensein einer Kundendatenbank als zentrale Integrationsplattform für die Gesamtheit der kundenorientierten Informationssysteme und der übrigen Touch Points.<sup>104</sup> Die zentrale Datenbank integriert sämtliche Kundenkontakte, da sie alle relevanten Informationen, die beim Kundenkontakt im Marketing-, Verkaufs- und Serviceprozess entstehen, sammelt und wiederum allen Prozessen zur Verfügung stellt.

Diese Integration der einzelnen Customer Touch Points ist umso wichtiger, als der Kunde häufig nacheinander bzw. parallel mehrere Sales Cycles mit einem Unternehmen unterhält (Folgekäufe, Cross/Up Selling). Bei Versicherungsgesellschaften z. B. besitzen Kunden üblicherweise mehrere ganz unterschiedliche Versicherungen. Sind solche Unternehmen bspw. nach einzelnen Sparten organisiert, wird der Kunde unter Umständen mehrfach erfasst und spartenspezifisch bearbeitet, wobei zwischen den einzelnen Sparten kein Abgleich der Kundeninformationen stattfindet. Dies kann dann dazu führen, dass ein Kunde nach Abschluss einer Versicherung als Neukunde begrüßt wird, obwohl er bereits seit längerer Zeit eine andere Versicherung desselben Unternehmens besitzt. „Um

---

<sup>102</sup> Beim Cross Selling wird ermittelt, welche der Kunden, die Produkt A gekauft haben, auch für den Kauf von Produkt B in Frage kommen. Wie z. B. in der Versicherungsbranche häufig der Fall ist, schließen Kunden nicht nur eine, sondern mehrere Versicherungen bei einem Unternehmen ab. Beim Up Selling wird recherchiert, inwieweit die Möglichkeit besteht, einem Kunden ausgehend von seinem derzeitigen Produkt ein höherwertiges zu verkaufen. So wird z. B. in der Automobilbranche versucht, Neukunden mit Einstiegsmodellen zu gewinnen und dann kontinuierlich in der Modellpalette nach oben zu führen, vgl. Hippner/Wilde (2003a), S. 11.

<sup>103</sup> Vgl. Hippner/Wilde (2003a), S. 12.

<sup>104</sup> Vgl. Link (2001), S. 15.

ein solch unwissendes Auftreten dem Kunden gegenüber zu vermeiden, ist es wichtig, nach außen ‚One Face to the Customer‘ zu präsentieren.“<sup>105</sup>

Die Grundlage hierfür bildet die integrierte Steuerung aller Kundenkontakte an Customer Touch Points.<sup>106</sup> Nebenbei ist auch die erfolgreiche Integration zwischen Front und Back Office-Bereich von besonderer Bedeutung, damit es eine rasche und fehlerlose Datenübermittlung erfolgen kann, die dem Kunden eine hohe Auskunftsbereitschaft des Unternehmens bietet und auch Schnelligkeit und Kostengünstigkeit der Abwicklungsprozesse sicherstellt.<sup>107</sup>

### **3.6 KOMPONENTEN und AUFGABENBEREICHE EINES CRM-SYSTEMS**

Die Zusammenführung aller Informationen über den einzelnen Kunden durch spezialisierte CRM-Systeme bildet die Grundlage für eine erfolgreiche Umsetzung der CRM-Philosophie.

Unter CRM-Systemen können sowohl Informations- als auch Kommunikationssysteme<sup>108</sup> verstanden werden, die zur Unterstützung des CRM eines Unternehmens eingesetzt sind.

Bei verteilten Architekturformen werden diese sogenannten IuK-Systeme von mehreren unterschiedlichen Systemen und Anwendungen gemeinsam genutzt. So sind z. B. Kommunikationssysteme, wie E-Mail, Telefon oder Fax häufig in mehrere Anwendungen eingebunden und können sowohl für die Kommunikation innerhalb eines Unternehmens (interne Kommunikation) als auch für dessen Kommunikation mit seiner Umwelt (externe Kommunikation) eingesetzt werden. Ebenso Funktionalitäten, wie z. B. Internetanwendungen, Datenbanken oder Workflowmanagement-Systeme kommen in mehreren Anwendungen mehrmals zum Einsatz. Vor diesem Hintergrund können CRM-Systeme von sonstigen unternehmerischen Systemen am besten durch die inhaltliche Ausrichtung und den Einsatzbereich der einzelnen Systeme abgegrenzt werden.

CRM-Systeme werden besonders in den Bereichen<sup>109</sup>

---

<sup>105</sup> Hippner/Wilde (2003a), S. 12.

<sup>106</sup> Vgl. Hippner/Wilde (2003a), S. 12.

<sup>107</sup> Vgl. Link (2001), S. 15.

<sup>108</sup> Hier lässt sich folgende grundsätzliche Charakterisierung vornehmen: *Kommunikationssysteme* sind primär auf Austausch, Übermittlung und Weiterleitung von Informationen und Daten zwischen mehreren Teilnehmern ausgerichtet. *Informationssysteme* sind primär auf Verarbeitung, Speicherung, Generierung und Darstellung von Informationen und Daten ausgerichtet, vgl. Amberg/Schuhmacher (2002), S. 22.

<sup>109</sup> Vgl. Amberg/Schuhmacher (2002), S. 22.

- Kundenkommunikation und Multi Channel Management<sup>110</sup>,
- Marketing,
- Vertrieb,
- Service und in der
- Verwaltung von Kundeninformationen eingesetzt.

„Die Spezifikation eines Systems als CRM-System erfolgt demnach anhand der prozessualen Zugehörigkeit bzw. der inhaltlichen, logischen Ausrichtung auf das CRM.“<sup>111</sup> Dies beinhaltet sowohl speziell für CRM-Bereich entwickelte und eingesetzte Systeme, als auch rein thematisch integrierte; d. h., es wird keine spezielle Nutzung einzelner Systeme vorausgesetzt.<sup>112</sup> So sind IuK-Systeme in die CRM-Systemumgebung eines Unternehmens integriert und werden darüber hinaus von anderen Systembereichen verwendet.<sup>113</sup>

In vielen Unternehmen ist bereits Software zur Unterstützung einzelner Funktionsbereiche vorhanden, es stellt sich aber die Problematik, dass die historisch gewachsenen IT-Lösungen (z. B. Computer Aided Selling, Helpdesks, Call Center, Analysesysteme usw.) untereinander eine große Heterogenität aufweisen, d. h., dass die isolierten Insellösungen keine einheitliche Sicht auf die im Unternehmen vorhandenen Kundendaten bieten.<sup>114</sup> Inkonsistente, veraltete oder gar falsche Daten sind oft das Resultat einer heterogenen Systemlandschaft.<sup>115</sup>

Informationstechnologisch betrachtet gibt es zwei Möglichkeiten, CRM-Systeme einzuführen.<sup>116</sup> Die eine ist der sogenannte integrierte Ansatz: Man versucht, durch die integrierte Lösung eines Softwareanbieters alle CRM-Anwendungen aus Marketing, Vertrieb und Service, wie auch aus den Bereichen Internet, Call Center u. a. in einer Systemlandschaft zu vereinen. Ein solcher integrierter Ansatz ist immer ein gewisser Kompromiss, weil in der Regel kein Hersteller in allen Segmenten die besten Lösungen

---

<sup>110</sup> Im Rahmen eines *Multi Channel Management* konzentriert man verstärkt auf die Ausgestaltung sowie die Koordination der Kanäle, da sie als zentrales Bindeglied zwischen dem Kunden und Unternehmen angesehen werden, vgl. Hippner (2006), S. 36. Die Aufgabe des Multi Channel Management ist es, für den Kunden denjenigen Kontaktmix, d. h., die optimale Allokation der Produkte, Serviceleistungen und Kanäle bereitzustellen, der von dem Kunden gewünscht wird, gleichzeitig die Kostenstruktur des Unternehmens jedoch so wenig wie nötig belastet, vgl. Hippner (2006), S. 36.

<sup>111</sup> Amberg/Schumacher (2002), S. 22.

<sup>112</sup> Vgl. Amberg/Schumacher (2002), S. 22.

<sup>113</sup> Vgl. Amberg/Schumacher (2002), S. 22.

<sup>114</sup> Vgl. Hippner/ Rentzmann/Wilde (2006), S. 47.

<sup>115</sup> Vgl. Hippner/ Rentzmann/Wilde (2006), S. 47.

<sup>116</sup> Vgl. Hippner/Hoffmann et al. (2006), S. 77 ff.

bietet. Solche integrierte CRM-Systeme sind auf dem derzeitigen Softwaremarkt als CRM-Komplettsystem (Standardsystem bzw. -software) erhältlich, welches Komponenten für Marketing, Vertrieb und Service zur Verfügung stellt und somit alle CRM-Funktionalitäten abbildet.

Die zweite Alternative ist die sogenannte Best of Breed-Lösung, d. h., man verwendet einen sogenannten „ERP-Backbone“ als Basis oder Rückgrad der Gesamtlösung und koppelt dann im Sinn von Satelliten „Best of Breed-Komponenten“, also die jeweils beste (die beste oder bereits vorhandene) Lösung eines anderen Herstellers an. Somit wird eine unternehmensindividuelle Lösung durch die Kombination spezialisierter CRM-Funktionsmodule konfiguriert. Eine Kombination von verschiedenen, isoliert betrachteten guten CRM-Einzellösungen mag funktionale Vorteile bieten, birgt jedoch bei der Implementierung erhebliche Risiken. Die Nachteile sind, dass man eine Reihe von Schnittstellenproblemen lösen muss und das Thema Datenintegration unter Umständen schwierig ist.

IT-Spezialisten und CRM-Experten sind sich in der Frage nicht einig, welche der beiden Alternativen zur IT-Unterstützung von CRM vorzuziehen ist, denn eine klare, gültige Antwort dafür kann nur unternehmensindividuell festgelegt werden. Die Einigkeit herrscht jedoch bei der Bewertung von isolierten Insellösungen, die keine einheitliche Sicht auf den Kunden (One Face of the Customer) erlauben und weiterhin eine Barriere für ein einheitliches Auftreten gegenüber dem Kunden (One Face to the Customer) darstellen.

Die Abbildung 7 stellt eine typische Klassifizierung der zurzeit auf dem Markt vorhandenen CRM-Systeme dar.

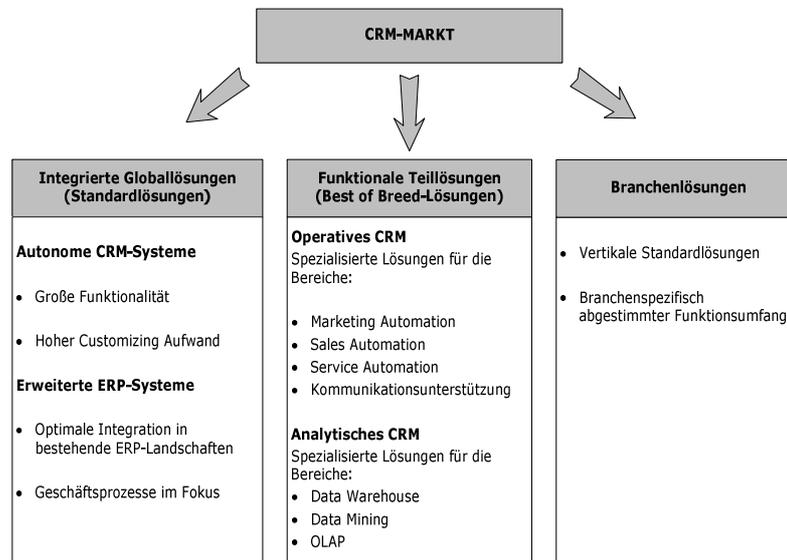
Integrierte Globallösungen (Standardlösungen), dazu gehören autonome CRM-Systeme, welche eine auf die spezifischen Anforderungen von CRM zugeschnittene technische Architektur bereitstellen und um CRM-Funktionalitäten erweiterte ERP-Systeme<sup>117</sup>, decken (nahezu) alle geforderten CRM-Funktionalitäten ab.

Trotz des hohen Aufwands des Customizing und der damit verbundenen Kosten bringen solche Systeme eine solide Funktionalität. Darüber hinaus bieten integrierte Globallösungen eine Vielzahl von weiteren Vorteilen, wie z. B. die internationale und schnelle Verfügbarkeit, der umfassende Support des Softwareanbieters, das effiziente Verarbeiten großer Datenbestände und das breite Funktionspektrum, ohne

---

<sup>117</sup> Ein ERP-System ist „eine standardisierte Anwendungssoftware zur Unterstützung der Prozesse und Funktionen in den betrieblichen Funktionsbereichen“, Leßweng/Lanninger/Thome (2004), S. 221, zit. in: Hippner/ Hoffmann/Rimmelspacher/Wilde (2006), S. 78.

zusätzlichen Integrationsaufwand. Daneben existieren auch CRM-Systeme, welche sich auf funktionalen Teillösungen (Best of Breed-Lösungen), d. h., auf ausgewählte Funktionalitäten einer CRM-Architektur konzentrieren. Diese einzelnen Lösungen müssen aber über entsprechende Schnittstellen verfügen, um mit anderen Teillösungen kommunizieren zu können.



**Abbildung 7: Klassifizierung von CRM-Systemen**

Quelle: Hippner /Hoffmann/Rimmelspacher/Wilde (2006), S. 79

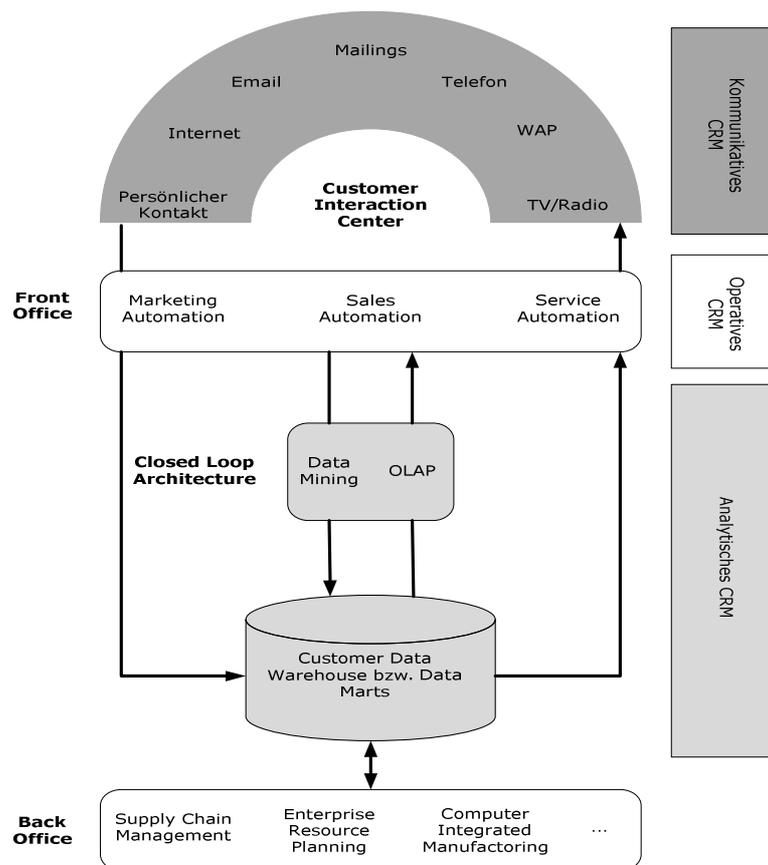
Vielfach werden viele Komponenten und Funktionalitäten einer Standardlösung in manchen Branchen nicht voll ausgeschöpft oder gar nicht genutzt und machen das ganze System somit unnötig, unübersichtlich und vor allem sehr teuer. So wird bspw. in der Pharmaindustrie bedeutend mehr Wert auf eine optimierte Unterstützung des Außendienstes gelegt, während im Versandhandel der Fokus eher auf einer umfassenden Contact/Call Center-Unterstützung liegt. Vor diesem Hintergrund entwickelt sich das Bedürfnis in Unternehmen, je nach Branche einzelne CRM-Funktionalitäten nur eingeschränkt oder aber verstärkt benötigen zu können. Durch Anpassung der CRM-Systeme an die spezifischen Anforderungen verschiedener Branchen entstehen letztlich auch branchenspezifische Standardlösungen. Egal ob Standardsoftware oder Best of Breed-Lösungen, CRM-Systeme zeichnen sich durch<sup>118</sup>

- eine Synchronisation und operative Unterstützung der funktionalen Bereiche Marketing, Vertrieb und Service (operatives CRM),
- eine Steuerung und Integration aller Kommunikationskanäle zwischen Kunde und Unternehmen (kommunikatives/kollaboratives CRM),

<sup>118</sup> Vgl. Hippner/Wilde (2003a), S. 13.

- sowie eine Zusammenführung und Auswertung aller Kundeninformationen zur bestmöglichen Kundenwertanalyse und zur Optimierung der Kundenbeziehungen durch das Wissen über die Präferenzen eines Kunden (analytisches CRM)

aus und stellen also den „Technology Enabler“ dar, um letztendlich die Prozess- und Bearbeitungsqualität im Kundenmanagement zu erhöhen.<sup>119</sup> Die Abbildung 8 zeigt deutlich das Zusammenspiel zwischen operativem, kommunikativem und analytischem CRM sowie den Komponenten der einzelnen Bereiche.



**Abbildung 8: Komponenten einer CRM-Lösung**

Quelle: Hippner/Wilde (2003), S. 15

### 3.6.1 Operatives CRM (oCRM)

IT-Lösungen, die in den Bereichen Marketing, Service und Vertrieb vorkommen, sind unter dem Begriff operatives CRM zusammengefasst. Das operative CRM umfasst alle Anwendungen, die im direkten Kontakt mit dem Kunden stehen (Front Office) und verfügt über Lösungen zur Marketing-, Sales- und Service-Automation (bspw. Systeme zum Kampagnenmanagement, zum Computer Aided Selling, zum Sales Force

<sup>119</sup> Vgl. Helmke, Stefan/Uebel Matthias 2001, S. 2.

Automation und zum Beschwerdemanagement), um den Dialog zwischen Kunden und Unternehmen sowie die dazu erforderlichen Geschäftsprozesse zu unterstützen.<sup>120</sup>

Die Steuerung der Kommunikationskanäle wird durch das oCRM übernommen, welches auch auf im Rahmen des aCRM (analytisches CRM) gewonnene Erkenntnisse zurückgreift.<sup>121</sup> Die einzelnen Automation-Bereiche decken somit administrative, analytische und kontaktunterstützende Aufgaben ab.<sup>122</sup> Weiters muss das operative CRM an vorhandene Back Office-Lösungen (Enterprise Resource Planning, Supply Chain Management, Computer Integrated Manufacturing, u. a.) angekoppelt werden, um den Kunden verlässliche Aussagen über Liefertermin, Verfügbarkeit, Produktionsschritt, u. a. zu machen.<sup>123</sup> Im Rahmen des operativen CRM wird letztendlich versucht, Informationen über einzelne Kunden möglichst umfassend zu bearbeiten, um die Kunden optimal ansprechen und bestmöglich betreuen zu können. Hippner und Wilde unterscheiden hierbei den aktionsorientierten Ansatz und den kundenorientierten Ansatz als zwei verschiedene Vorgehensweisen:<sup>124</sup>

Der aktionsorientierte Ansatz zeichnet sich durch vorgegebene Marketingaktionen aus, die auf diejenigen Kunden bzw. Kundensegmente angewendet werden, die aufgrund ihrer Kundenmerkmale und Kontakthistorie die beste Response erwarten lassen. Die Problematik des aktionsorientierten Ansatzes ist, dass die Mailing-Aktion für alle Kunden des ausgewählten Kundensegments gleichzeitig initiiert wird.<sup>125</sup>

Dagegen wird beim kundenorientierten Ansatz für jeden einzelnen Kunden die seinen Bedürfnissen entsprechend individuell zugeschnittene Mailing-Aktion bestimmt und danach wird der Kunde mit dieser festgelegten Aktion zu einem optimalen Zeitpunkt angesprochen. Der Ausgangspunkt ist hier also der Kunde und nicht die Marketingaktion. Der kundenorientierte Ansatz bildet den eigentlichen Kern des Kampagnenmanagements.

### **3.6.1.1 Marketing-Automation**

Die Marketing-Automation dient der Steuerung und Unterstützung der kundenbezogenen Geschäftsprozesse im Marketing. Das Ziel ist dabei, die Kundenkontakte komplett so zu gestalten, dass sie miteinander logisch abgestimmt

---

<sup>120</sup> Vgl. Hippner/Wilde (2003a), S. 14 und Hippner/Rentzmann/Wilde (2006), S. 48.

<sup>121</sup> Vgl. Hippner/Wilde (2003a), S. 20.

<sup>122</sup> Vgl. ebd., S. 20.

<sup>123</sup> Vgl. ebd., S. 14.

<sup>124</sup> Vgl. ebd., S. 21.

<sup>125</sup> Dies führt dazu, dass generell nur eine kleine Anzahl von Kunden zu einem aus ihrer Sicht angemessenen Zeitpunkt kontaktiert wird.

sind. Die zentrale Aufgabe dieses Bereichs von operativem CRM ist das Kampagnenmanagement (administrative Aufgabe der Marketing-Automation), das dem richtigen Kunden das richtige Informations- und Leistungsangebot im richtigen Kommunikationsstil über den richtigen Kommunikationskanal zum richtigen Zeitpunkt vermittelt.

Die im Rahmen des Kampagnenmanagements entwickelte Marketingaktionen (Kampagnen) basieren auf kontinuierlichen Analysen der Kundencharakteristika und des Kundenverhaltens. Von besonderer Bedeutung für ein erfolgreiches Kampagnenmanagement ist die Synchronisation aller Customer Touch Points im Hinblick auf die Forderung nach One Face to the Customer, unter Einbezug aller dialogorientierten Werbemaßnahmen (z. B. Direct Mailings, E-Mail, Wurfsendungen, usw.) sowie der Kundenkontakte über Service und Vertrieb.

Zum Kampagnenmanagement gehören Funktionen zur Planung, Steuerung und abschließenden Auswertung von Marketingmaßnahmen.

Die Kampagnenplanung ist der erste Schritt innerhalb des Kampagnenmanagements. Zunächst erfolgt eine nähere Definition der verfolgten Ziele wie z. B. Kundenrückgewinnung oder Umsatzsteigerung. Hier sollte besonders auf die Operationalität dieser Ziele und damit auf die Messbarkeit Wert gelegt werden.<sup>126</sup> Darüber hinaus werden die Marketingchancen auf Zielmärkten analysiert und demzufolge wird eine Marketingstrategie bzw. ein Marketingprogramm entwickelt. Im Weiteren wird die Zielgruppe zur Durchführung der Kampagne bestimmt. Die Segmentierung der in sich heterogenen Kundschaft im Kampagnenmanagement erweist sich als zentraler Ausgangspunkt für eine zielgruppengerechte Kundenbearbeitung. Dabei gewinnen neben den klassischen Kriterien wie Alter, Wohnort, Geschlecht, Familienstand oder Einkommen auch verhaltensbezogene und psychografische<sup>127</sup> Merkmale an Bedeutung. Im Anschluss werden für den Dialogfluss bestens geeignete Kommunikationskanäle vorbereitet.<sup>128</sup> Für die Wahl des passenden Kommunikationskanals berücksichtigt man die Präferenzen auf der einen, sowie die Kosten-Nutzen-Relation für das Unternehmen auf der anderen Seite.<sup>129</sup> Nach diesem Schritt kann die Kampagne kalkuliert und freigegeben werden. In der Phase Prozessdefinition fließen alle zuvor getroffenen Überlegungen hinsichtlich der

---

<sup>126</sup> Vgl. Hafner (2003).

<sup>127</sup> Sozio-psychographische Merkmale sind qualitative Merkmale zur Beschreibung einer Zielgruppe, wie z. B. Werthaltungen und Einstellungen.

<sup>128</sup> Das Kampagnenmanagement strebt die Umsetzung integrierter Kontaktketten an, die aus dem kombinierten Einsatz der einzelnen Kommunikationskanäle bestehen (Multichannel Integration).

<sup>129</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 55.

Zielsetzung der Kampagne, Zielgruppe, Medienwahl sowie der Zeit- und Budgetplanung ein und werden schließlich in einem Kampagnenmanagementsystem visualisiert.<sup>130</sup> Es ist durchaus sinnvoll, vorher Abbruchkriterien im Kampagnenmanagementsystem festzulegen. Als Abbruchkriterien können bspw. ein zu geringer Rücklauf oder andere kritische Ereignisse definiert werden, die eine Fortführung der Kampagne nicht rentabel erscheinen lassen.

Die Kundenanalysen, aus denen konkrete Marketingaktionen bzw. Kampagnen abgeleitet werden (analytische Aufgabe der Marketing-Automation), bilden die Grundlage für das Kampagnenmanagement. Diese unter Einsatz von OLAP und Data Mining durchgeführten Analysen sind:

- *Kundensegmentierungsanalyse*: Unter Segmentierung versteht man die Bildung von Kundengruppen, die hinsichtlich ihrer Produkt-, Dienstleistungs- und Kommunikationsbedürfnisse möglichst homogen sein sollten. Die Gruppierung kann z. B. nach Kaufkraft, Kaufverhalten und Präferenzen erfolgen. Danach können diese Gruppen entsprechend ihrer jeweiligen Bedürfnisstruktur effektiv bearbeitet werden.
- *Kundenbewertungsanalyse*: Kundenbewertungsmodelle ermitteln für jeden Kunden anhand ausgewählter Merkmale seinen Wert für das Unternehmen (Kundenscoring). Als Kundenwert wird hier der vom Anbieter wahrgenommene, bewertete Beitrag eines Kunden bzw. des gesamten Kundenstamms zur Erreichung der monetären und nicht-monetären Ziele verstanden.<sup>131</sup> Es werden in den einzelnen Segmenten die profitabelsten Kunden herausgefiltert, und für diese Zielsetzung verwendet man als Orientierungsgröße meistens den Customer Life Time Value<sup>132</sup>. Mit Hilfe dieser Kundenbewertung kann das zukünftige Potential des einzelnen Kunden abgeschätzt und somit unnötige Marketingaktionen vermieden werden.
- *Assoziationsanalyse*: „Assoziationsregeln beschreiben Korrelationen zwischen gemeinsam auftretenden Dingen.“<sup>133</sup> Der Zweck einer Assoziationsanalyse besteht also darin, Items (Elemente einer Menge, wie z. B. einzelne Artikel eines Warenkorb) zu ermitteln, die das Auftreten anderer Items innerhalb

---

<sup>130</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 56.

<sup>131</sup> Vgl. Cornelsen (2000), S. 38.

<sup>132</sup> Neben der Berechnung des Customer Life Time Value gibt es auch andere Methoden und Verfahren, die bei der Kundenwertanalyse angewendet werden können. Einige von denen sind ABC-Kundenanalyse, Punktbewertungsverfahren, Erstellung von Kundenportfolios, Rentabilitätsanalysen wie z. B. Kundendeckungsbeitragsrechnungen usw.

<sup>133</sup> Bollinger (1996), S. 257.

einer Transaktion implizieren.<sup>134</sup> Eine solcherart aufgedeckte Beziehung zwischen zwei oder mehr Items kann dann als Regel der Form „Wenn Item (Menge) A, dann Item (Menge) B“ bzw.  $A \rightarrow B$  dargestellt werden.

Wenn man an große Datenmengen denkt, die als Transaktionsdaten im Einzelhandel (Scanner-Kassen) und E-Commerce und als Navigationsdaten im Web (Suchmaschinen, digitale Bibliotheken, Wikis, u. a.) jede Sekunde gesammelt werden, kommt der Analyse von Abhängigkeiten oder Assoziationen zwischen Datensätzen große Bedeutung zu. So lässt sich z. B. mit Hilfe moderner Scanner-Kassen in Supermärkten leicht analysieren, welche Produkte häufig im Zusammenhang gekauft werden und welche Marken und Produkte die Käufer bevorzugen. Waren im Supermarkt können aufgrund dieses gewonnenen Wissens besser platziert und somit kann deren bestehende Kaufwahrscheinlichkeit erhöht werden. Zudem ist es auch möglich, Kunden mit noch gezielterer Werbung und Promotionen zu versorgen.

Ein weiteres Beispiel für die Anwendung von Assoziationsregeln ist die Warenkorbanalyse. Die Analyse der Warenkorbinhalte kann Aufschluss über die Produktpräferenzen von Kunden, sowie über mögliche Cross und Up Selling-Potentiale liefern. Deswegen ist diese Analyse besonders im Zusammenhang mit einer Cross und Up Selling-Analyse interessant, welche zur Ermittlung des Cross und Up Selling-Potenzials und damit zur Prognose des zukünftigen Kaufverhaltens herangezogen wird.

- *Kundenabwanderungsanalyse (Churnanalyse)*: Die Kunden werden hinsichtlich ihrer Abwanderungsneigung differenziert. Bei der Analyse wird es versucht, die Gründe für die Kundenabwanderung zu ermitteln, um rechtzeitig präventive Marketingmaßnahmen einzuleiten und so abwanderungsgefährdete Kunden vom drohenden Wechsel zu einem Konkurrenten abzuhalten.<sup>135</sup>
- *Webanalyse*: Seine Aufgabe liegt darin, den Inhalt und die Struktur von Websites sowie das Kundenverhalten auf einer Website (Dauer, aufgesuchte Seiten) zu analysieren. Es wird versucht, zwischen Surfen und eventuellem Kauf des Kunden eine Korrelation zu finden. Dieses Wissen wird dann für Online Marketing-Aktionen verwendet. Mit der Webanalyse erkennt man bspw. wie viele Nutzer eine Website hat, woher sie kommen, welche Seiten ihre Aufmerksamkeit erregen, ob sie auf Werbung oder Sonderangebote reagieren,

---

<sup>134</sup> Cabena et al. (1998), S. 80, zit. in: Hettich/Hippner (2001), S. 459.

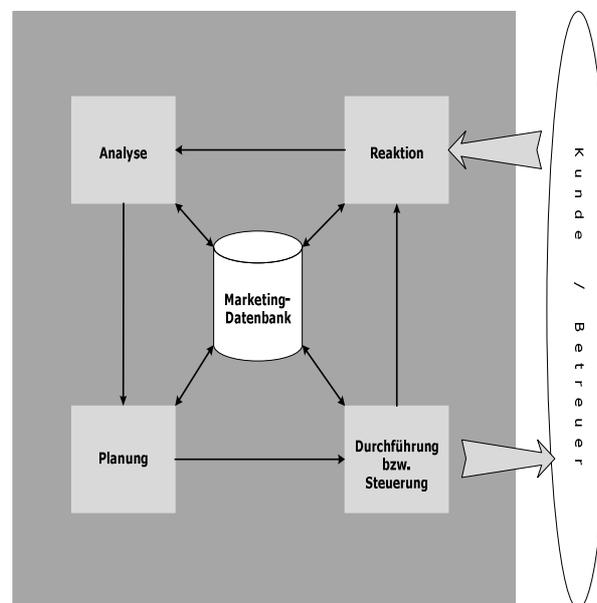
<sup>135</sup> Vgl. Hippner/Wilde (2003a), S. 24 und Gawlik/Kellner/Selfert (2002), S. 49.

was sie gekauft haben und anschließend wann und wo sie die Website verlassen haben.

Das Kampagnenmanagement ist als ein Kreislauf zu verstehen (siehe Abb. 9): Die Analyse des Kundenstamms führt zur Kampagnenplanung, Kampagnendurchführung und zu einer Reaktion des angesprochenen Kunden.

Im Kampagnenmanagement wird es allgemein davon ausgegangen, dass immer ein Respons des Kunden auf eine Aktion des Unternehmens folgt. Dabei ist auch eine Ablehnung des Angebots oder eine Nichtreaktion (dieser negative Respons zeigt jedoch nicht, dass sich der Kunde für angebotene Leistungen überhaupt nicht interessiert, denn er könnte ja auch verhindert sein) wertvoll für die Analyse.

Im Verlauf der Kampagne wird jede erzielte Kundenreaktion, egal ob positiv oder negativ, in der Marketing-Datenbank erfasst und mit vordefinierten Werten für das Auslösen einer Folgeaktion des Unternehmens abgeglichen.<sup>136</sup> Die Reaktion des Unternehmens auf ein bestimmtes Kundenverhalten erfolgt auf Basis der zuvor in der Kampagnenplanung aufgestellten Kommunikationsregeln (z. B. Wenn Kunden nicht auf E-Mail reagiert, dann kontaktiere ihn per Telefon).<sup>137</sup> Aufgrund der Auslösung von Marketingmaßnahmen nach dem Eintreten bestimmter Ereignisse wird diese „wenn/dann“ Vorgehensweise auch eventgetriggertes Marketing genannt.<sup>138</sup>



**Abbildung 9: Kreislauf des Kampagnenmanagements**

Quelle: In Anlehnung an Nils Hafner (2003).

<sup>136</sup> Vgl. Hippner/Wilde (2003a), S. 23 und Gawlik/Kellner/Selfert (2002), S. 47.

<sup>137</sup> Vgl. Gawlik/Kellner/Selfert (2002), S. 47.

<sup>138</sup> Vgl. Hippner/Wilde (2003a), S. 23 und Gawlik/Kellner/Selfert (2002), S. 47.

Nach Abschluss der Kampagnendurchführung erfolgt die Kampagnenanalyse (Wirkungsanalyse): Die Reaktionen der Kunden auf einzelne Kampagnen werden kontinuierlich der Marketing-Datenbank zugeführt. Mit Hilfe der Kampagnenanalyse werden diese aktualisierten Daten ausgewertet und daraus handlungsrelevante Informationen für den weiteren Verlauf der Kampagne gewonnen, die dann in die Planung von folgenden Aktionen einbezogen werden müssen.<sup>139</sup> So werden wesentliche Entscheidungshilfen für weitere zukünftige Marketingaktionen bzw. -kampagnen generiert (Closed Loop-Prinzip). Im Rahmen des Data Mining kann z. B. das Reaktionsverhalten verschiedener Kundensegmente in Anhängigkeit von der jeweiligen Kauf- und Kontakthistorie analysiert und zur Steuerung nachfolgender Aktionen genutzt werden.<sup>140</sup> Die Marketing-Datenbank spielt hier eine besondere Rolle. Sie ist in alle Phasen des Kampagnenmanagements einbezogen.<sup>141</sup> Im Gegensatz zu klassischem Databasemarketing, wobei die Datenbank nur die Kundendaten für die Ermittlung der idealen Zielgruppe für eine Marketingaktion liefert, werden im Kampagnenmanagement die Kampagnenplanung und -durchführung komplett in der Marketing-Datenbank abgebildet, um so aus kleinsten Veränderungen der Durchführung später Rückschlüsse ziehen zu können.<sup>142</sup>

Ein wichtiger Bestandteil der Kampagnenanalyse ist das Kampagnencontrolling. Dem Kampagnencontrolling kommt in den Zeiten knapper Budgets immer eine besondere Rolle zu, denn das Ergebnis einer Kampagne wirkt sich auch auf nachfolgende Kampagnen aus. Die Erfolgskontrolle entscheidet über die Durchführung weiterer Kampagnen. In einer laufenden Kampagne werden die Kampagnenergebnisse permanent mit den in der Planungsphase bereits definierten Kontrollparametern sowie Abbruchkriterien verglichen. Außerdem wird in regelmäßigen Zyklen überwacht, ob es notwendig ist, die Kampagne in ihrer Durchführung an die realen Kundenbedürfnisse und Reaktionen anzupassen. Nach Abschluss einer Kampagne erfolgt die Analyse des gesamten Kampagnenerfolgs. Dabei werden Kriterien wie erwarteter Nutzen, Responsquote, Akzeptanz der gewählten Kommunikations- und Vertriebskanäle und Cross Selling-Potentialität den Kampagnenkosten gegenübergestellt. Die Ergebnisse des Kampagnencontrollings werden für weitere Analysen historisiert und abgespeichert und laufen wieder in die Planung weiterer Kampagnen ein.

Eine weitere Aufgabe (kontaktunterstützende Aufgabe) der Marketing-Automation ist die Erstellung, Verwaltung und Bereitstellung von Marketingmaterialien, um damit den

---

<sup>139</sup> Vgl. Hippner/Wilde (2003a), S. 23.

<sup>140</sup> Vgl. Hippner/Wilde (2003a), S. 23.

<sup>141</sup> Vgl. Hafner (2003).

<sup>142</sup> Vgl. Hafner (2003).

Kundenkontakt an den Customer Touch Points zu unterstützen.<sup>143</sup> Hierfür wurden früher Marketing-Enzyklopädie-Systeme (MES) eingesetzt, welche den Zugriff auf sämtliche im Unternehmen verfügbare Marketingmaterialien wie Produktinformationen, Wettbewerberdaten, Präsentationen oder Werbebroschüren ermöglichte. Heute wird aber die Aufgabe von MES von einem Content Management System (CMS) übernommen, das zusätzlich die Mitarbeiter im Marketing bei der Erstellung der Marketinginhalte unterstützt. Unter einem Content Management System versteht man ein computergestütztes Erstellungs-, Verwaltungs- und Archivierungssystem für den Content in unterschiedlichsten digitalen Formaten.<sup>144</sup>

In Anlehnung an Schramm kann Content (Inhalt) als von Menschen erzeugte und in medienspezifischer Form präsentierte digitale, computerverwertbare Information unterschiedlichster Art bezeichnet werden.<sup>145</sup> Sie können als Dateien vorliegen, die einzeln verarbeitet werden wie statische Bilder, Grafiken (JPEG, GIF, etc.), Text-Dateien (Word, Excel, PDF, etc.) und dynamische Audio/Video-Dateien oder auch als zusammenhängende, miteinander verknüpfte Dateisysteme, wie z. B. Webseiten, oder in strukturierter Form, wie z.B. in Datenbanken. Die Inhalte können sowohl intern durch Mitarbeiter im Unternehmen als auch extern durch den Kunden direkt genutzt werden. Die Kernaufgabe eines CMS besteht diesbezüglich darin, alle vorhandenen Contents im Unternehmen zu integrieren und diese für die Geschäftsprozesse im Marketing-, Sales- und Service-Bereich bereitzustellen. Den Mitarbeitern in diesen Bereichen wird dadurch ein einheitliches Marketing- und Produktmaterial zur Verfügung gestellt, welches regelmäßig aktualisiert wird. So können Situationen im Kontakt mit den Kunden, in denen Kunden aufgrund vorangegangener Kampagnen besser informiert sind als die Mitarbeiter sicherlich vermieden werden.

### **3.6.1.2 Sales-Automation**

Da die intensivste Beziehung zum Kunden im Vertrieb aufgebaut wird, soll dieser Bereich des operativen CRM, welche die engste Schnittstelle zwischen Kunden und Unternehmen darstellt, intensiv unterstützt werden. Wie Marketing-Automation deckt Sales-Automation administrative, analytische und kontaktunterstützende Aufgaben ab.

Bei der Vertriebsunterstützung unterscheiden sich jedoch CRM-Systeme in ihren ausgewählten vertriebsunterstützenden Funktionen, je nachdem, ob die CRM-Lösung auf B2B- oder B2C-Sektor ausgerichtet ist.

---

<sup>143</sup> Vgl. Hippner/Wilde (2003a), S. 24.

<sup>144</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 69.

<sup>145</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 69.

Unternehmen, die bspw. über Versandhandel im Internet nur Endkonsumenten bedienen (B2C), zeichnen sich einerseits durch eine enorm hohe Kundenanzahl, andererseits durch wesentlich geringe Erträge bei einzelnen Kunden aus. Dementsprechend sind Vertriebsanwendungen wie Key Account Management oder Field Service bei relativ geringen Umsätzen der Top-Kunden sinnlos und sogar kostspielig.

Daher liegt der Schwerpunkt der B2C-Vertriebsunterstützung generell in dem Service und einer automatisierten Angebotserstellung, Preisfindung und Auftragserfassung, wogegen es im gewinnträchtigen B2B-Bereich mehr auf Key Account Management und Field Service fokussiert ist.

Heutige CRM-Systeme unterstützen umfangreich die Routine- und Administrationsaufgaben des Vertriebs, welche ihn im Verkaufsablauf so schwer belasten, dass er sich nicht voll und ganz auf seine Kernaufgaben konzentrieren kann. Dies umfasst z. B.:<sup>146</sup>

- Termin- und Routenplanung
- Spesenabrechnung
- Besuchsberichterfassung
- Kunden- und Ansprechpartnerverwaltung (Account Management)
- Verwaltung von Kundenkontakten (Contact Management)
- Aktivitätensteuerung (Activity Management)
- Unterstützung bei der Angebotserstellung, Preisfindung und Auftragserfassung
- Unterstützung bei der Zielplanung und Budgetierung
- Informationen zu Verkaufsvorgängen und Produkten

Eine Besuchsberichterfassung ermöglicht dem Außendienst ein schnelles Einarbeiten in Gespräche und Verhandlungen vergangener Kundenbesuche. Mittels Account und Contact Management werden auf der einen Seite Angaben zum Kundenunternehmen und zu Ansprechpartnern im Kundenunternehmen festgehalten, auf der anderen Seite Informationen zu Kundenkontakten wie Zeitpunkt, Inhalt und Ergebnis des Kontakts

---

<sup>146</sup> Vgl. Link/Hildebrand (1993), S. 93 ff und Schulze (2002), S. 28 f.

dokumentiert. Zudem sind auch persönliche und private Informationen über Kunden sehr hilfreich, um sie individuell bedienen zu können und so ein persönliches Verhältnis zum Kunden aufzubauen.

Activity Management dient zur Planung und Dokumentation der Tätigkeiten von Mitarbeitern im Unternehmen und zur Weiterleitung von Aufgaben an andere Personen als Workflow-Ansatz. Um den Kunden jederzeit Auskunft über die einzelnen Verkaufsvorgänge geben zu können, müssen dem Vertriebsmitarbeiter Angaben zu Anfragen, Angeboten, Aufträgen, aber auch zu Serviceleistungen und Reklamationen zur Verfügung gestellt werden.

Neben der Unterstützung im Angebots- und Auftragsprozess hat Sales-Automation auch die Zielplanung und Budgetierung zu unterstützen, die dazu dienen, den künftigen Umsatz bzw. Absatz für Verkaufsgebiete oder Mitarbeiter zu planen und demgemäß das Marketing-Budget zu bestimmen.

Weiters werden in Sales-Automation sämtliche Analysen durchgeführt, um die Potentiale des einzelnen Kunden auszuwerten und dadurch entsprechende Marketingmaßnahmen einzusetzen. Es werden folgende Analysen ausgeführt<sup>147</sup>:

- *Lost-Order-Analyse*: Dabei werden alle Angebote, die nicht zu einem Auftrag führen, dahingehend analysiert, weshalb es zu keinem Abschluss gekommen ist. Sie liefert Erkenntnisse über die Wettbewerbsfähigkeit des Unternehmens und Ansätze zur Verbesserung der strategischen Vorgehensweise gegenüber dem Kunden.
- *Sales-Cycle-Analyse*: Diese Analyse liefert den nächsten voraussichtlichen Wiederbeschaffungszeitpunkt, so dass zum richtigen Zeitpunkt dem Kunden ein neues Angebot gemacht wird. Der Vorteil für den Verkäufer liegt darin, dass er den Kunden frühzeitig ansprechen kann, bevor er zur Konkurrenz abwandert.
- *Chancen-Verfolgung-Analyse (Opportunity Management)*: Darunter versteht man die mehrstufige Erfassung, Pflege und Qualifizierung jedes Kundenkontakts – von dem anonymen Neukunden bis hin zum letztendlichen Vertragsabschluss. So entsteht für den Vertriebsmitarbeiter die Möglichkeit, zusätzliche Verkaufschancen (sogenannte Opportunities) wahrzunehmen. Dabei wird er durch graphische Pipeline-Analyse-Tools unterstützt. Mit Hilfe dieser Tools kann er jederzeit den Status eines Kontakts bzw. Angebots abfragen, um

---

<sup>147</sup> Vgl. Schulze (2002), S. 28 f und Hippner/Wilde (2003a), S. 26.

sich einen aktuellen Gesamtüberblick über bestehende Verkaufschancen mit Abschlusswahrscheinlichkeiten, erwarteten Abschlussbeträgen und -terminen zu verschaffen.

- *Vertriebsanalyse und Forecasting*: Dienen zur Erstellung von Prognosen auf Basis der noch nicht abgeschlossenen Verkaufsaktivitäten und zur Ableitung von Trends aus einer vergangenheitsbezogenen Analyse der Verkäufe.

Während des Verkaufsgesprächs kommt es vielmehr darauf an, dem einzelnen Kunden ein individuell zugeschnittenes und einzigartiges Kontakt- und Verkaufserlebnis zu bieten. Hier finden sogenannte Interactive Selling Systeme (ISS) ihren Einsatz, welche als eine Ergänzung zu früheren CAS (Computer Aided Selling)-Systeme gesehen werden können und speziell für die Unterstützung des Vertriebsmitarbeiters beim Verkaufsgespräch entwickelt worden sind.<sup>148</sup> Sie enthalten elektronische Produktkataloge, Produktkonfiguratoren sowie die bereits erwähnten Marketing-Enzyklopädien und liefern dem Vertriebsmitarbeiter alle notwendigen Informationen, die er zur Unterstützung seiner Argumentation beim Kunden benötigt.

Eine automatische Angebotserstellung, Preisfindung und Auftragserfassung, wie in herkömmlichen Online Shops der Fall ist, bildet den Kern der B2C-Vertriebsunterstützung. Kunden können sich im Internet über sämtliche Produkte und Leistungen informieren und online direkt kaufen. Die Bestellung wird automatisch an das ERP-System weitergeleitet. Durch die Anbindung des operativen CRM-Systems an das ERP-System ist eine Online-Auftragserfassung möglich. So sind Informationen über Verfügbarkeit, Produktionsschritt und geschätzte Liefertermin von Produkten und Leistungen jederzeit abrufbar, was wiederum dazu verhilft, den Automatisierungsgrad von zeit- bzw. geldintensiven Aufgaben des Vertriebs zu erhöhen und somit Prozesskosten gering zu halten. Besteht ein Produkt aus mehreren Komponenten, so kann der Kunde durch einen sogenannten Produktkonfigurator sein Produkt individuell nach den gewünschten Anforderungen zusammenstellen. Die Firma DELL z. B. stellt einen solchen Konfigurator in ihren Online-Kauf-Seiten zur Verfügung, wo der Kunde einen Computer mit verschiedenen Komponenten ausstatten kann. Der Produktkonfigurator benützt dabei Konfigurationsregeln und aktuelle Preisangaben und ermittelt automatisch den Preis der Produktkonfiguration für den Kunden. Im Sinne des Key Account Management kann der Kunde zudem noch eine individuelle Kondition zur Finanzierung des Produktes wahrnehmen.

---

<sup>148</sup> Vgl. Hippner/Wilde (2003a), S. 26.

### 3.6.1.3 Service-Automation

Service-Automation stellt Funktionen bereit, die zur Lösung von Kundenproblemen und Beschwerden in der After-Sales-Phase benötigt werden. Wichtige Funktionen für den Servicebereich sind Problemlösungsmanagement (bzw. Beschwerdemanagement), Management von Serviceverträgen, Serviceanalysen, Call Center Management, Internet Self Service und Field Service.<sup>149</sup>

Der Servicebereich eines Unternehmens umfasst den Serviceinnendienst und den Serviceaußendienst.

Der Serviceinnendienst nimmt die vom Kunden initiierten Kontakte von außen an und bearbeitet sie (Kontaktunterstützung der Service-Automation). Der Servicemitarbeiter des Außendienstes wird durch die Service-Automation schwerpunktmäßig bei seinen administrativen Aufgaben unterstützt. Viele Aufgabenstellungen eines CRM-Systems zur Serviceaußendienstunterstützung ähneln denen des Vertriebsaußendienstes, so dass entsprechende Funktionen von Sales-Automation (z. B. Angebotserstellung, Spesenverwaltung Besuchsberichterfassung, Routenplanung u. a.) auch im Servicebereich angewendet werden können.<sup>150</sup>

Das Management der Serviceverträge verwaltet die vertraglichen Vereinbarungen und Geschäftsdetails mit dem Kunden und dient zur Identifizierung der zu erbringenden Serviceleistungen (z. B. Garantieleistungen, Reparaturen, u. a.), so dass die reibungslose Abwicklung der Geschäftsbeziehung zwischen dem Kunden und Unternehmen garantiert wird.<sup>151</sup> Mit Hilfe von Serviceanalysen können die Qualität der erbrachten Serviceaktivitäten ausgewertet werden.

Wie bereits im Marketing- und Sales-Bereich bietet sich auch im Service die Möglichkeit an, aus den umfangreichen Kundeninformationen wichtige Erkenntnisse abzuleiten. So klassifiziert der weltberühmte Automobil-Hersteller Daimler-Chrysler gemeldete Pannen und Probleme von PKWs nach Ausstattung, Modell und Baujahr, analysiert sie und versucht, gravierende technische Störungen vorzeitig zu erkennen und sie bereits bei der nächsten planmäßigen Inspektion präventiv zu beheben.<sup>152</sup> Auch im Bereich des Beschwerdemanagements werden verstärkt Analysen durchgeführt (z. B. für die Zahl und Art der eingegangenen Beschwerden, zur Messung der durchschnittlichen Erreichbarkeit des Mitarbeiters und der

---

<sup>149</sup> Vgl. Schulze (2002), S. 31.

<sup>150</sup> Vgl. Hippner/Wilde (2003a), S. 27.

<sup>151</sup> Vgl. Schulze (2002), S. 31.

<sup>152</sup> Vgl. Hippner/Wilde (2003a), S. 28.

Beschwerdebearbeitungsdauer, u. a.), welche wertvolle Erkenntnisse über das Serviceniveau, Verbesserungs- und Einsparungspotentiale liefern. Durch Text Mining können E-Mails von Kunden anhand ihrer Inhalte und Dringlichkeit klassifiziert und automatisch an den entsprechenden Sachbearbeiter weitergeleitet werden.

Das Problemlösungsmanagement enthält Standardschritte zur Diagnose und Bearbeitung von Servicefällen. Dazu wird eine Problemlösungsdatenbank verwendet, in der eine Fallbasis erstellt wird. Mittels fall-basierten Schließen werden zur Lösung eines Problems die Fälle aus der Datenbasis gesucht, welche große Ähnlichkeiten mit dem neu auftretenden Problemfall aufweisen, und diese dann zur Bewertung des Problemfalles herangezogen. So können zukünftige Entscheidungen aus den Erfahrungen der Vergangenheit abgeleitet werden.

Oft kontaktiert der Kunde das Unternehmen erst dann wieder, wenn er ein Problem mit der Leistung des Unternehmens hat. In einer solchen Situation ist es erforderlich, schnell zu reagieren, Beschwerden anzunehmen und dem Kunden eine Soforthilfe über einen Kommunikationskanal anzubieten (Kontaktunterstützung). Durch gezieltes Beschwerdemanagement<sup>153</sup> soll aus dieser Situation ein zufriedengestellter Kunde hervorgehen. Deswegen ist es kritisch, die eingehenden Beschwerden als eine Chance zu sehen, die Erwartungen des Kunden bei der Beschwerdebehandlung möglichst gut zu erfüllen und durch adäquate Reaktion des Unternehmens den Kunden wieder an das Unternehmen zu binden, und diesmal sogar mit einer noch stärkeren Loyalität als zuvor.

Das Field Service dient der Unterstützung des technischen Außendienstes bei der Bearbeitung von Störungsmeldungen und der Behebung der Störungen.<sup>154</sup> Die Erfassung der erledigten Arbeiten und der Schadensreporte, Arbeitszeiterfassung und Spesenabrechnung gehören zu seinen Aufgaben, und zusätzlich enthält er auch Funktionen zum Management der Serviceverträge und Funktionen des Verkaufsprozesses.

Das Call Center spielt eine sehr wichtige Rolle bei der Behandlung von Servicefällen. Bei technischen Fragen an das Service steht ein sogenannter Help Desk dem Call Center-Agent zur Verfügung und bietet ihm eine gezielte Unterstützung zur Lösung anfallender Servicefälle an.

---

<sup>153</sup> Alle Beschwerden müssen systematisch in einer Beschwerdedatenbank erfasst werden. Zu jedem einzelnen Fall wird eine Beschwerdennummer vergeben und die weiterführenden Aktionen festgelegt, die eine angemessene Behandlung begründeter Beschwerden ermöglichen. Können die Beschwerden trotzdem nicht behoben werden, so werden sie an eine übergeordnete Abteilung weitergeleitet.

<sup>154</sup> Vgl. Schulze (2002), S. 31.

Ein Help Desk ist ein wissensbasiertes Datenbanksystem, das für die Aufnahme von Störungsfällen, für die Beantwortung von Benutzeranfragen und die Weiterleitung nicht beantworteter Fälle zuständig ist.<sup>155</sup> Dabei stellt der Call Center-Agent dem anrufenden Kunden im System vordefinierte Fragen, um das bestehende und zu behebende Problem zu identifizieren. Handelt es sich dabei um ein bereits bekanntes Problem, so wird dem Call Center-Agent vom System die dazu entsprechende optimale Lösung vorgeschlagen. Die Service-Automation bietet dem einzelnen Kunden aber auch die Möglichkeit, Probleme selbst zu lösen, indem sie Supportinformationen über Produkte und Leistungen auf der Webseite zur Verfügung stellt. Mit dem sogenannten Internet Self Service können Kunden jederzeit via E-Mails Serviceanfragen stellen und bereits existierende Anfragen über den WWW-Browser abrufen.

Weiters ist es möglich, dem Kunden Statusinformationen über seine Auftragsdurchführung mitzuteilen. Durch die Anbindung an das ERP-System liefert das sogenannte Order Tracking System dem Serviceinnendienst Informationen über den aktuellen Stand der Auftragsbearbeitung. Ein besonders gutes Beispiel hierfür ist das Order Tracking System vom PC-Hersteller Dell. Nach einer Bestellung erhält der Kunde sein persönliches Passwort und eine Trackingnummer per Email zugesandt, mit denen er sich weltweit auf der Webseite von Dell einloggen und sich über den Stand des Produktionsprozesses bis hin zum Versandtermin erkundigen kann.

### **3.6.2 Analytisches CRM (aCRM)**

Die Integration und Bündelung aller kundenbezogenen Daten<sup>156</sup>, die dann analysiert und vom operativen CRM weiterverarbeitet werden können, bildet die Grundlage für eine erfolgreiche Umsetzung der CRM-Strategie.

Die zentrale Aufgabe des analytischen CRM besteht darin, sämtliche Kundendaten systematisch zu bewahren und zu analysieren.

Dabei ist jedoch zu beachten, erstens, dass Unmengen an kundenrelevanten Daten für eine adäquate Analyse Unternehmen zur Verfügung stehen<sup>157</sup> und zweitens, dass die IT-Landschaft in den meisten Unternehmen durch zahlreiche historisch gewachsene Insellösungen geprägt ist<sup>158</sup>. Es finden sich z. B. Computer Aided Selling, Help Desks, Call Centers, Analysesysteme, Webanwendungen u. a. welche den Unternehmen ganz unterschiedliche Informationen über den Kunden übermitteln. Diese Situation

---

<sup>155</sup> Vgl. Hippner/Wilde (2003a), S. 29.

<sup>156</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 57.

<sup>157</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 57.

<sup>158</sup> Vgl. Gawlik/Kellner/Selfert (2002), S. 40. und Hippner/Rentzmann/Wilde (2006), S. 47.

verhindert natürlich eine einheitliche Sicht auf die im Unternehmen vorhandenen Kundendaten, was oft zu inkonsistenten, veralteten oder falschen bzw. unvollständigen Informationen über den Kunden führen kann.<sup>159</sup> Um jedoch eine kundenzentrierte Unternehmensstrategie erfolgreich zu verfolgen, ist es für Unternehmen unerlässlich, sich ein ganzheitliches Bild von seinen Kunden zu verschaffen. Dies ist nur mit der gezielten Zusammenführung der einzelnen Insellösungen und Funktionsbereichen - wie z. B. Marketing, Vertrieb und Service - möglich<sup>160</sup>, damit alle Unternehmensbereiche nur noch auf eine zentrale Kundendatenbank zugreifen<sup>161</sup>. Somit kann der Kundenbestand segmentiert, können Analysen über Kundenzufriedenheit, -loyalität, -profitabilität durchgeführt und Kenntnisse über typische Lebenszyklen innerhalb der Kundenbasis erlangt werden.<sup>162</sup> Die diesen Analysen zugrunde liegenden Prozesse und Softwarelösungen bezeichnet man als analytisches CRM.<sup>163</sup>

Im Mittelpunkt des analytischen CRM steht der Analyse-Aspekt, und diese Analyse erfolgt durch ein sogenanntes Customer Data Warehouse, dessen Aufgabe darin besteht, Kundendaten aus den unterschiedlichsten dezentralen Quellen (sowohl aus den internen transaktionsorientierten Systemen des Unternehmens wie z. B. operativen CRM-Systemen, ERP-Systemen als auch aus externen Quellen) in eine logische, einheitliche, konsistente Systemumgebung zu integrieren. Dazu werden die relevanten Informationen selektiert und aufbereitet sowie für Analysen zur Entscheidungsunterstützung bereitgestellt.<sup>164</sup>

Typische Informationen in einem Customer Data Warehouse sind:<sup>165</sup>

- *Stammdaten* von Kunden und Interessenten: (z. B. Name, Adresse, Soziodemographie, Kaufverhaltensmerkmale, Lifestyle Typ). Das sind jene Daten, die längerfristig gleich bleiben und produktunabhängig sind.
- *Kaufhistorien*: Wann und wie oft wurde was gekauft? Kaufhistorien liefern produkt- und zeitpunktbezogene Anhaltspunkte für das kundenindividuelle Nachfragevolumen (z. B. Bedarfszeitpunkte, nächster Wartungstermin, Position

---

<sup>159</sup> Vgl. Gawlik/Kellner/Selfert (2002), S. 40. und Hippner/Rentzmann/Wilde (2006), S. 47.

<sup>160</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 47.

<sup>161</sup> Vgl. Gawlik/Kellner/Selfert (2002), S. 40. und Hippner/Wilde (2003a), S. 15.

<sup>162</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 57.

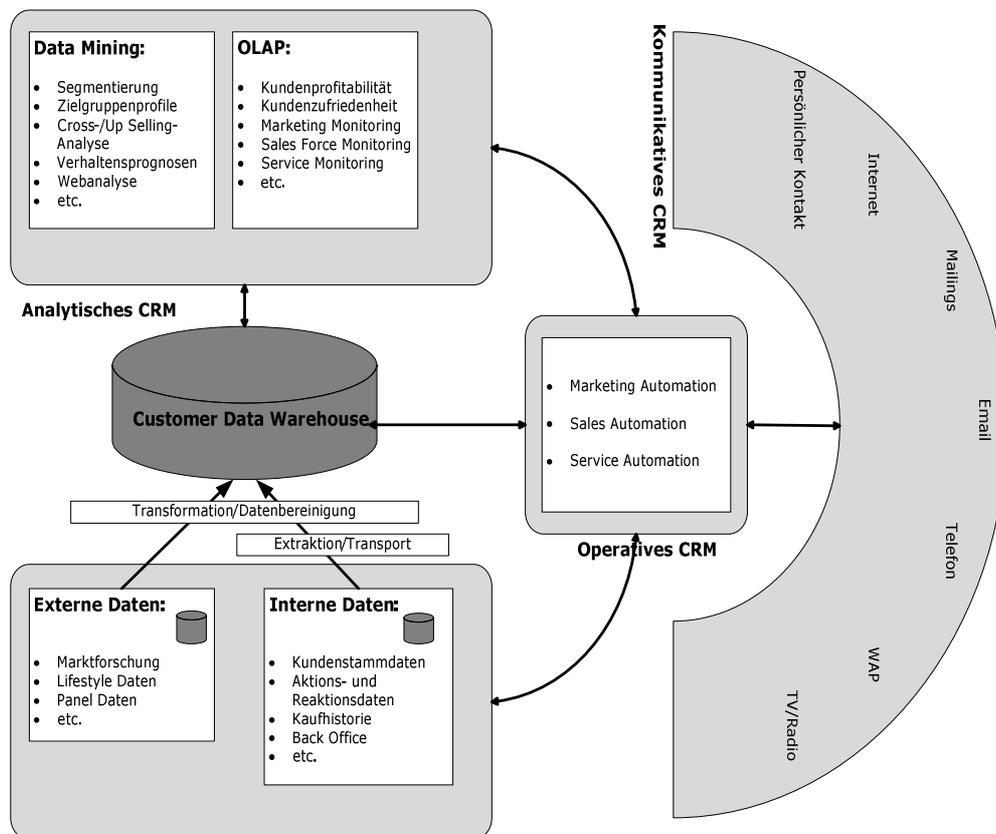
<sup>163</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 58.

<sup>164</sup> Die Daten des Data Warehouse (DWH) werden in der Regel nicht direkt zu Analysezwecken verwendet. Stattdessen werden die zur Analyse relevanten Daten aus dem DWH entnommen, denormalisiert und in sogenannten Data Marts gespeichert. Data Marts können in diesem Zusammenhang als kleine DWHs angesehen werden, die anwendungsbezogene Daten stark verdichtet bereitstellen.

<sup>165</sup> Vgl. Hippner/Wilde (2003a), S. 15, Gawlik/Kellner/Selfert (2002), 43. und Schulze (2002), S. 35.

im Kundenportfolio, Kundenklassifizierung)

- *Aktionsdaten*: Wer hat wen, wann und wie kontaktiert? Das sind Informationen über Art/Kanal, Intensität, Umfang, Häufigkeit, Zeitpunkte, Inhalte von Unternehmensaktivitäten bzgl. des Kunden (z. B. Werbebriefe, Angebotserstellung, Kundenbesuche)
- *Reaktionsdaten*: Wer hat wie auf einen Kontakt reagiert? Hat es irgendwelche Beschwerden gegeben? Also Informationen über Anfragen, Einstellungen, Beschwerden von Kunden, Dauer der Kundenbeziehung, Loyalitätsgrad des Kunden.



**Abbildung 10: Zentrale Bedeutung des Customer Data Warehouse im CRM-Konzept**

Quelle: Hippner/Wilde (2003), S. 16.

In Wirklichkeit ist das Customer Data Warehouse eine von den operativen Datenbanken getrennte Analysedatenbank. Die Entkopplung der Datenanalyse von den operativen Systemen stellt sicher, dass die Tagesabwicklung von Geschäftsfällen nicht von rechenintensiven Analyseanwendungen beeinträchtigt wird.<sup>166</sup>

Mit dem Aufbau eines Customer Data Warehouse erstellt man jedoch nur die Basis des analytischen CRM. Um aus diesen im Data Warehouse gespeicherten Informationen

<sup>166</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 49.

ein umfangreiches Wissen für das Unternehmen generieren zu lassen, werden spezielle Analyseinstrumente und -verfahren wie Online Analytical Processing (OLAP) und Data Mining benötigt. Durch den erfolgreichen Einsatz dieser beiden genannten Analyseansätze gewinnt man wertvolle Erkenntnisse und Erfahrungen, die zur kontinuierlichen Optimierung der kundenbezogenen Geschäftsprozesse ausgewertet werden können. Das CRM wird dadurch zu einem lernenden System (Closed Loop Architecture), in dem Kundenreaktionen gezielt genutzt werden, um die Abstimmung von Kundenkommunikation, Produkten und Dienstleistungen auf differenzierte Kundenbedürfnisse kontinuierlich zu verbessern (siehe Abbildungen 8 und 10).<sup>167</sup>

### **3.6.2.1 Online Analytical Processing**

OLAP ist eine multidimensionale Datenbanktechnologie und stellt die Abkürzung für **On-Line Analytical Processing** dar. Dieser Begriff wurde 1993 von Codd et al.<sup>168</sup> geprägt und umfasst Konzepte für die Datenversorgung des Managements und der Fachabteilungen, insbesondere im Bereich der Entscheidungsunterstützung und der Datenanalyse.

Ein OLAP-System ermöglicht dem jeweiligen Anwender, die im Data Warehouse aufbereiteten, umfangreichen und in mehrdimensionalen Daten verborgenen, geschäftsrelevanten Zusammenhänge bzw. Fakten aufzudecken. Im Gegensatz zu einem Online Transaction Processing System<sup>169</sup> (OLTP-System), steht hier die Durchführung komplexer Analysevorhaben im Vordergrund. Während bei OLTP typischerweise nur aktuelle Dateninhalte gespeichert und satzweise abgefragt oder mittels Transaktionen - d. h., Lese- und Schreiboperationen auf kurzfristig veränderlichen Datenbeständen - aktualisiert werden, finden bei OLAP aggregierte und historische - über einen längeren Zeitraum gesammelte - Daten Verwendung, welche die Durchführung von komplexen und Ad-hoc-Abfragen ermöglichen. Bei Datenbankabfragen, die große Mengen an Daten einschließen und durch OLTP-Systeme durchgeführt werden, besteht die Gefahr, dass erhebliche Performance Probleme auftreten, welche die Abwicklung von Transaktionen aus dem operativen Geschäftsbereich hemmen können.<sup>170</sup> Das OLAP-Konzept schafft Abhilfe für diese Probleme.

Die von E. F. Codd eingeführten Regeln, welche die Anforderungen an ein OLAP-

---

<sup>167</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 49.

<sup>168</sup> Vgl. Codd et al. (1993).

<sup>169</sup> OLTP-Systeme werden typisch für die Transaktionsverarbeitung im operativen Bereich des Unternehmens eingesetzt. SAP R/3 ist ein Beispiel eines OLTP-Systems.

<sup>170</sup> Vgl. Knox/Maklan/Ryals (2000), S. 65.

Konzept darstellen, waren sehr ausführlich. Pendse und Creeth definierten deshalb 1995 OLAP mit der Abkürzung FASMI (**F**ast **A**nalysis of **S**hared **M**ultidimensional **I**nformation).<sup>171</sup> Die neue Definition dient zur Erleichterung, da sie kürzer und einfacher zu merken ist.<sup>172</sup> Der Begriff FASMI setzt sich aus den Anfangsbuchstaben der Wörter Fast, Analysis, Shared, Multidimensional und Information zusammen:<sup>173</sup>

- *Fast*: Schnell. Damit ist gemeint, dass die Antwortzeiten bei einfachen Anwenderabfragen nicht länger als 1-2 Sekunden sein sollten und bei komplexeren bis zu maximal 20 Sekunden dauern dürfen.
- *Analysis*: OLAP sollte mit Geschäftslogik und statistischen Analysen umgehen können, die für den Endbenutzer im Rahmen von Datenanalysen relevant sind. Für die Analyse des Datenmaterials ist es notwendig, dass das OLAP-System dem Anwender ermöglicht, neue ad hoc-Berechnungen zu definieren und durchzuführen sowie deren Ergebnisse auf verschiedenen Wegen zu präsentieren. Typische Analyseformen sind Zeitreihenvergleiche, Kostenverteilung, das Ausnahmeberichtswesen, u. dgl.
- *Shared*: Ein OLAP-System muss von mehreren Benutzern gleichzeitig genutzt werden können. Außerdem soll das OLAP-System über Zugriffsrechte und dazu gehörende Sperrmechanismen verfügen.
- *Multidimensional*: Als Hauptkriterium fordern Pendse und Creeth eine mehrdimensionale Strukturierung der Daten mit voller Unterstützung der Dimensionshierarchien. Mit OLAP soll es möglich sein, dass der Benutzer eine multidimensionale Sicht auf die Daten bekommt und beliebige Dimensionen in einer Anfrage kombiniert.
- *Information*: Bei der Analyse sollen einem Anwender alle benötigten Daten (Basisdaten und Berichtsdaten) zur Verfügung stehen, unabhängig von der Datenmenge und -herkunft.

Die FASMI-Regeln befassen sich eher mit Benutzeranforderungen und gehen wenig auf technische Anforderungen ein. Insgesamt sind sie allerdings weniger spezifisch als die Regeln nach Codd. Deswegen können nach dieser Definition wesentlich mehr Systeme dem OLAP zugeordnet werden.

---

<sup>171</sup> Vgl. Pendse/Creeth (1995).

<sup>172</sup> Vgl. Pendse (2005).

<sup>173</sup> Vgl. Pendse (2005) und Chamoni (1998), S. 237.

OLAP-Systeme bilden betriebswirtschaftlich relevante Maßgrößen, wie Absatz, Umsatz, Deckungsbeitrag durch einen mehrdimensionalen Datenwürfel (sog. OLAP-Cube) ab.<sup>174</sup> Die Dimensionen des Würfels entsprechen betriebswirtschaftlichen Kriterien, wie z. B. Kundengruppen, Produktgruppen, Verkaufsgebiete, usw.<sup>175</sup> Je nach der spezifischen Fragestellung können entlang dieser Dimensionen die betriebswirtschaftlichen Maßzahlen aufgebrochen bzw. verfeinert (sog. Drill down) oder aggregiert (sog. Roll up) werden.<sup>176</sup> Ergänzend kann der Würfel mittels bestimmter Navigationsfunktionen gekippt und gedreht (sog. Dice) oder in einzelne Scheiben zerlegt (sog. Slice) werden<sup>177</sup>, um eine bessere und detaillierte Sicht auf Daten zu erhalten. Dadurch kann z. B. analysiert werden, welche Stückzahl einer Produktgruppe in welcher Filiale in welchem Jahr verkauft wurde. Weitere Fragestellungen, die mit OLAP beantwortet werden können, sind: Wie hoch ist der Anteil der Umsätze in allen Verkaufsgebieten außer bspw. der Verkaufsgebiete B und D am Gesamtumsatz? Wie hoch sind die Abweichungen zu den geplanten Geschäftszahlen? Wie gut sind die neuen Produkte im Vergleich zu den alten Produkten? usw.

Bei OLAP handelt es sich mehr oder weniger um eine mehrdimensionale Analyse mit differenzierten Datenbankabfragen<sup>178</sup>, wobei die mehrdimensionale Sichtweise der OLAP-Systeme im Vergleich zur zweidimensionalen Abbildung in relationalen Systemen eine problemadäquatere Darstellung des naturgemäß mehrdimensionalen Unternehmensumfeldes ermöglicht<sup>179</sup>. So ist der Anwender in der Lage, sein Unternehmen aus jedem geschäftlichen Blickwinkel zu betrachten und zu analysieren. Die unterschiedlichen Sichtweisen von OLAP-Systemen sind auf die Kunden, auf die Produkte, auf die Zeit, auf geographische Gebiete oder Märkte ausgerichtet. All diese sind die normalen geschäftlichen Dimensionen eines Unternehmens und mit OLAP lassen sie sich gleichzeitig und gleichberechtigt untersuchen, was viele Möglichkeiten und Potentiale für Unternehmen aufdeckt.

Heutige OLAP-Tools zeichnen sich durch eine einfache und leicht zu erlernende Benutzeroberfläche aus, die dem Management einen direkten Zugang zur Datenanalyse gibt.<sup>180</sup> Allerdings gibt es eine Reihe von Einschränkungen<sup>181</sup>, die man in Betracht ziehen sollte, wenn es um OLAP-Systeme geht: OLAP liefert sinnvolle

---

<sup>174</sup> Vgl. Chamoni (1998), S. 233.

<sup>175</sup> Vgl. Chamoni (1998), S. 223.

<sup>176</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 49.

<sup>177</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 49.

<sup>178</sup> Vgl. Hippner/Wilde (2003a), S. 17.

<sup>179</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 50.

<sup>180</sup> Vgl. Hippner/Wilde (2003a), S. 18.

<sup>181</sup> Vgl. Hippner/Wilde (2003b), S. 215f.

Informationen nur dann, wenn differenzierte Fragestellungen (a priori-Hypothesen) vom Anwender konkret bzw. scharf formuliert worden sind. Die konsequente Erfassung der in den Daten verborgenen geschäftsrelevanten Merkmale und interessanten Zusammenhänge erfordert aber darüber hinaus die automatische Abarbeitung unscharfer a priori-Hypothesen, bei denen lediglich ein zu erklärendes Zielmerkmal (Z. B. Kundenwert) vorgegeben ist, wogegen die Bestimmung der Art des Zusammenhanges und der Merkmale aus einer Vielzahl von Möglichkeiten den eigentlichen Kern der Datenanalyse bildet. Die analytische Bearbeitung solcher unscharfen Fragestellungen ist mit sehr viel Zeit und Aufwand verbunden – da die Komplexität der möglichen Zusammenhänge innerhalb der zu untersuchenden Daten enorm steigt<sup>182</sup> – und erscheint so aus unternehmerischen Gründen nicht rational. Allein schon im Fall von scharfen a priori-Hypothesen überschreitet die deskriptive Darstellung der komplexen Zusammenhänge die kognitiven Grenzen der Anwender (auch aufgrund fehlenden Wissens über betriebswirtschaftliche Zusammenhänge) und die darstellungstechnischen Grenzen der OLAP-Tools. So stellte sich im Laufe der Zeit die Forderung nach einer *automatisierten* Suche nach unscharfen Beziehungen heraus, um die oben genannten Schwachstellen der eingeschränkten *manuellen* Suche von OLAP nach scharfen Zusammenhängen zur Gänze zu beseitigen. Demzufolge wurde sogenanntes Data Mining entwickelt.

### 3.6.2.2 Data Mining

Der Begriff Data Mining stammt aus dem Bergbau (Mining), wo mit großem technologischen Aufwand enorme Gesteinsmengen maschinell abgebaut und aufbereitet werden, um Edelmetalle und Edelsteine ans Tageslicht zu fördern.<sup>183</sup> In informationstechnischem Kontext kann man unter Data Mining die Anwendung von (statistisch-mathematischen) Methoden auf einen Datenbestand mit dem Ziel der Mustererkennung verstehen.

Mittlerweile findet man in der Literatur eine Vielzahl von Definitionen für Data Mining (siehe u. a. Fayyad/Shapiro/Smyth 1996 a,b,c, Berry/Linoff 1997, Janetzko/Steinhöfel 1997). Einige von ihnen werden in Folgendem erläutert:

Berry und Linoff sehen Data-Mining als Erforschung und Analyse großer Datenmengen mit automatischen oder halbautomatischen Werkzeugen, um bedeutungsvolle Muster

---

<sup>182</sup> Bei scharf formulierten Fragestellungen ist der Suchraum durch die vom Anwender vorgegebenen Dimensionen eingeschränkt. Im Gegensatz dazu erhöht sich bei unscharfen Fragestellungen die Analysekomplexität erheblich, da hier die für die Analyse relevanten Dimensionen erst entdeckt werden müssen, vgl. Hippner/Wilde (2003a), S. 18.

<sup>183</sup> Vgl. Hippner/Wilde (2003b), S. 216.

und Regeln aufzufinden.<sup>184</sup> Laut dieser Definition wird der Anwendungsbereich von Data Mining auf große Datenmengen eingeschränkt. Dieses Merkmal trifft schon logisch auf viele Anwendungen zu. Da aber auch in kleinen Datenmengen bedeutungsvolle Muster und Regeln existieren können, ist die Definition nach diesem Merkmal nicht als allgemeingültig anzusehen.

Nach Decker und Focardy ist Data Mining als eine Methodik zur Problemlösung zu verstehen, um logische oder mathematische, zum Teil komplexe Beschreibungen von Mustern und Regelmäßigkeiten in Datensätzen zu entdecken.<sup>185</sup>

Eine weitere allgemeine Definition liefern Fayyad, Piatetsky-Shapiro und Smyth, in der Data Mining als Teilschritt des übergeordneten *KDD-Prozesses*<sup>186</sup> angesehen wird, der die Anwendung von spezifischen Mining-Methoden und Analysealgorithmen umfasst und zur Interpretation von Mustern, die mittels Algorithmen aus den Daten gewonnen wurden, dient.<sup>187</sup> „Knowledge Discovery in Databases is the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.“<sup>188</sup> Da keine spezifizierten Einschränkungen gefordert werden, erreicht diese Variante der Definition das höchste Maß an Allgemeingültigkeit, dennoch lässt sie sich nicht klar von verwandten Forschungsgebieten, wie z. B. der explorativen Statistik abgrenzen.

Aus den vorgestellten Definitionen kann man zusammenfassend unter dem Begriff Data Mining die Anwendung von Algorithmen auf Daten mit der Zielsetzung, Muster aus den Daten zu extrahieren, verstehen.

Während Fayyad, Shapiro und Smyth zwischen KDD und Data Mining als einer Phase innerhalb des KDD-Prozesses differenzieren, werden die Begriffe heute synonym verwendet.<sup>189</sup> So wird der *iterative und interaktive KDD-Prozess*, verbunden mit dem Terminus *Data Mining*, zum Data Mining-Prozess.<sup>190</sup> „Data Mining is the process of extracting previously unknown, valid, and actionable information from large databases and then using the information to make crucial business decisions.“<sup>191</sup>

---

<sup>184</sup> Vgl. Berry/Linoff (1997), S. 5.

<sup>185</sup> Vgl. Decker/Focardi (1995), S. 3.

<sup>186</sup> Wissenentdeckungsprozess, die Abkürzung KDD steht für Knowledge Discovery in Databases.

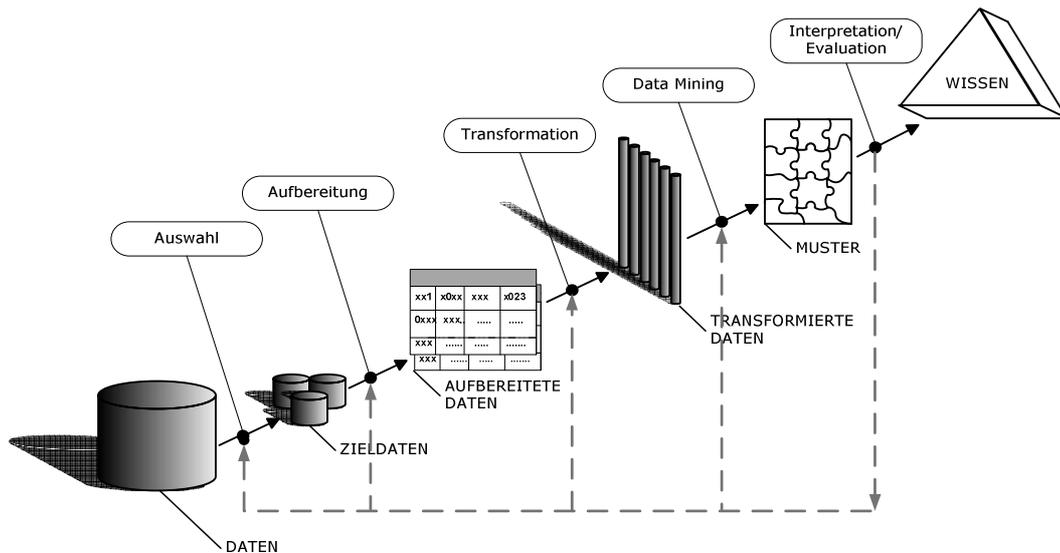
<sup>187</sup> Vgl. (eigene Übersetzung) Fayyad/Shapiro/Smyth, (1996a), S. 4.

<sup>188</sup> Fayyad/Shapiro/Smyth, (1996b), S. 40 f.

<sup>189</sup> Vgl. Hippner/Wilde (2003b), S. 217.

<sup>190</sup> Vgl. Hippner/Wilde (2003b), S. 217.

<sup>191</sup> Cabena/Hadjinian/Standler/Verhees/Zanasi (1998), S. 12, zit. in: Hippner/Wilde (2003b), S. 217.



**Abbildung 11: Stufen des Data Mining-Prozesses**

Quelle: in Anlehnung an Fayyad/Shapiro/Smyth, (1996a), S. 9.

Der Data Mining-Prozess umfasst somit alle erforderlichen Schritte von der Datenbeschaffung über die effiziente Methodenanwendung bis hin zur Abbildung verschiedener Muster und deren Präsentation (siehe Abb. 11). Hier ist das fachliche Wissen der Anwender sehr wichtig, da es ohne dieses kaum möglich ist, gefundene Erkenntnisse nutzbringend auf profitable betriebswirtschaftliche Anwendungsmöglichkeiten zu prüfen und entsprechend umzusetzen. „Wer mit gängigen Abfragesprachen das Wesentliche herausfischen will, muss schon vorher wissen, wonach er sucht.“<sup>192</sup> Data Mining setzt sich zum Ziel, effiziente Konzepte zur automatischen Datenanalyse zu entwickeln, um in den Daten verborgene implizite Wissen aufzufinden und explizit darzustellen. Um dieses Ziel zu realisieren, verwendet es Methoden und Verfahren, die meist auf Statistik beruhen (Klassifikation, Clustering, Regression u. a.); befasst sich aber auch mit den Aspekten des Maschinellen Lernens (Maschine Learning), der Mustererkennung (Pattern Recognition) und der Künstlichen Intelligenz (Artificial Intelligence).<sup>193</sup> Data Mining befindet sich so im Schnittpunkt dieser verschiedenen Wissenschaftsdisziplinen<sup>194</sup> und profitiert nicht nur von den Fortschritten der einzelnen Disziplinen, sondern insbesondere von kombinierten Einsatzmöglichkeiten dieser verschiedenen Wissenschaften.

Die prozessartige Vorgehensweise des Data Mining, die dabei angewendeten Techniken bzw. Methoden und wie diese im Customer Relationship Management ihren Einsatz finden, werden in den folgenden Kapiteln näher erläutert.

<sup>192</sup> Janetzko, D./Steinhöfel, K. (1997), S. 294 zit. in: Hippner/Wilde (2003b), S. 216.

<sup>193</sup> Vgl. (eigene Übersetzung) Fayyad/Shapiro/Smyth, (1996b), S. 39-40.

<sup>194</sup> Vgl. (eigene Übersetzung) Fayyad/Shapiro/Smyth, (1996b), S. 39 f.

### 3.6.3 Kommunikatives CRM (kCRM)

Die Ausschöpfung der geschäftlichen Potentiale des Customer Relationship Management wird erst durch die Nutzung verschiedener Kanäle und Medien ermöglicht, da die Unternehmen und ihre Kunden über die Kanäle und Medien in Kontakt treten und Leistungen austauschen. Die Kommunikationskanäle und –medien werden so eingesetzt, dass sie eine Kommunikation in zwei Richtungen (bidirektional) zwischen Kunden und Unternehmen ermöglichen.

Das kommunikative CRM (oder kollaboratives CRM) befasst sich mit der gesamten Steuerung und Unterstützung sowie der Synchronisation dieser Kommunikationskanäle bzw. –medien (Telefonie, Internet, Email, Mailings, etc.) zum Kunden.<sup>195</sup>

Die einzelnen Kommunikationskanäle weisen ganz unterschiedliche Eigenschaften auf (besitzen spezifische Stärken und Schwächen). Aufgrund dieser Tatsache eignen sie sich für verschiedene Anwendungsgebiete. Außerdem stellen die Kunden spezifische Forderungen an die Kanäle, wie beispielsweise geringe Transaktionskosten, kompetente Information und Kaufgeschwindigkeit, die sich jedoch situativ verändern.<sup>196</sup> Deshalb setzen Unternehmen die Kanäle meist kombiniert ein, um verschiedene Wettbewerbsvorteile zu erlangen. Sie verfolgen damit eine Mehrkanal- bzw. Multikanalstrategie.<sup>197</sup>

Durch die Verfolgung einer solchen Strategie erreichen Unternehmen neue Kundengruppen und stellen ihnen mehr Interaktionsmöglichkeiten zur Verfügung. Außerdem versuchen sie durch das Anbieten zusätzlicher Kanäle zu bereits bestehenden Kunden die Vertriebskosten zu senken.

Durch den Einsatz zusätzlicher Kanäle wird aber das Konfliktpotential zwischen den Kanälen erhöht und deren Steuerung erschwert. Wenn die Kunden über mehrere Kommunikationskanäle mit dem Unternehmen in Kontakt treten können, ist es sicherlich möglich, dass verschiedene Kanäle miteinander konkurrieren. Daraus können Steuerungsprobleme entstehen, wenn sie unabhängig voneinander operieren und damit ihre Kooperation erschwert wird.<sup>198</sup> Diese Probleme führen dazu, dass Kunden während der Kommunikation mit dem Unternehmen durch Verlust der Transparenz überfordert werden. Deswegen ist für jedes Unternehmen von besonderer Bedeutung, genau festzulegen, welche Kanäle es überhaupt bedienen möchte und wie

---

<sup>195</sup> Vgl. Gawlik/Kellner/Selfert (2002), S. 55 ff und Hippner/Wilde (2003a), S. 14 f.

<sup>196</sup> Vgl. Schulze (2002), S. 42.

<sup>197</sup> Vgl. ebd., S. 42.

<sup>198</sup> Vgl. ebd., S. 42.

diese im Unternehmen verwaltet werden. Hier kommt das sogenannte *Multi Channel Management* zur Anwendung.

Das Multi Channel Management im CRM stellt sicher, dass einerseits die Kunden über alle Kanäle eine einheitliche Sicht auf das Unternehmen (One Face to the Customer) erhalten, andererseits aber auch das Unternehmen eine einheitliche Sicht auf den Kunden bekommt (One Face of the Customer).<sup>199</sup>

Wenn z. B. ein Kunde über Call Center das Unternehmen telefonisch kontaktiert hat, ermöglicht das Multi Channel Management dem Unternehmen, auch seine Außendienst- und Filialemitarbeiter über diesen telefonischen Kontakt des Kunden mit dem Unternehmen zu informieren.<sup>200</sup> Unabhängig davon, welche Kanäle der Kunde für seine Anfrage vorzieht, kann ihm somit eine schnelle und kompetente Reaktion auf seine Wünsche sichergestellt werden.<sup>201</sup>

Da die Kontaktaufnahme heutzutage oft über Telefon und Internet erfolgt, wird im Folgenden auf das Customer Interaction Center (entwickelt aus dem früheren Call Center) und das Internet näher eingegangen.

### **3.6.3.1 Customer Interaction Center**

Der CRM-Ansatz bezieht explizit alle Kommunikationskanäle des Unternehmens zu Kunden mit ein.<sup>202</sup> Während bisher die einzelnen Kanäle noch überwiegend isoliert voneinander organisiert wurden, werden sie heute zunehmend in ein Customer Interaction Center (CIC) integriert, und zwar durch eine technische Anbindung verschiedener Kanäle an eine zentrale Kundendatenbank, z. B. auf Basis einer EAI-Implementierung.<sup>203</sup>

CICs sind die gewünschte Weiterentwicklung herkömmlicher Call Centers, deren Fokus auf der Telephonie liegt. Im Gegensatz dazu unterstützen CICs zusätzlich Internet, Email, SMS, WAP (Wireless Application Protocol) u. a.

Durch die Nutzung von CICs kann das vielzitierte Ziel, die integrierte Steuerung aller Kommunikationskanäle zur Individualisierung der Kundenansprache, realisiert werden.

Um als ein Customer Interaction Center bezeichnet zu werden, soll ein Call Center

---

<sup>199</sup> Vgl. Schulze (2002), S. 43 und Hippner/Rentzmann/Wilde (2006), S. 65.

<sup>200</sup> Vgl. Schulze (2002), S. 43.

<sup>201</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 65.

<sup>202</sup> Vgl. Hippner/Wilde (2003a), S. 29.

<sup>203</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 65. und Schulze (2002), S. 44.

über folgende Fähigkeiten zur weitgehenden Automatisierung verfügen:<sup>204</sup>

- *Outbound-Funktionalitäten*: Sie ermöglichen die Kontaktaufnahme zum Kunden, z. B. für Telemarketing-Aktivitäten.
- *Automatic Call Distribution*: Eingehende Anrufe werden automatisch und gleichmäßig auf die einzelnen Servicemitarbeiter verteilt.
- *Computer Telephony Integration (CTI)*: Die Computer werden mit der Telekommunikationsanlage verbunden. Dies ermöglicht dem Servicemitarbeiter, gleich beim Anruf den Kunden anhand seiner im System gespeicherten Telefonnummer über das CTI-System zu identifizieren. Die wichtigste Funktionalität des CTI-Systems ist, dass dieses System sofort alle kundenrelevanten Informationen auf dem Bildschirm des Call Center-Agents erscheinen lässt. Alle Kundendaten stehen so dem Agenten zur Verfügung und ermöglichen ihm daher eine schnelle Aufnahme und Bearbeitung des Kundenwunsches.
- *Interactive Voice Response (IVR)*: Diese Technologie dient dazu, die geeigneten Servicemitarbeiter für den Kunden zu ermitteln und/oder standardisierte Kundenanfragen automatisch zu bearbeiten. Dabei wird der Anrufer von einer digitalisierten Audio-Aufzeichnung begrüßt und kann seinen Input entweder verbal oder über die Tastatur seines Telefons geben. So kann der Kunde bzw. sein Anliegen qualifiziert und dem dafür zuständigen Mitarbeiter zugeordnet werden - noch bevor der Kunde mit einem Call Center-Agent in persönlichen Kontakt tritt.<sup>205</sup>
- *Skill Based Routing*: Anhand dieser Technologie werden Anrufe an die Mitarbeiter entsprechend deren Fähigkeiten weitergeleitet, um den Kunden schnell zufrieden zu stellen und das Potential der Mitarbeiter besser auszuschöpfen. So wird beispielsweise ein Anruf aus der Türkei in einem deutschsprachigen Call Center an einen Mitarbeiter weitergeleitet, der türkisch sprechen kann, oder für ein spezifisches Anliegen des Kunden verbindet der Call Center-Agent den Kunden mit einem Experten. Auf diese Weise kann vermieden werden, dass der Kunde immer wieder weitervermittelt werden muss, bis er sich endlich mit der richtigen Person in Verbindung setzt.

Um die optimale Kommunikation zwischen dem CIC und anderen

---

<sup>204</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 65 f und Hippner/Wilde (2003a), S. 29 ff.

<sup>205</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 66.

Unternehmensbereichen wie z. B. Marketing, Vertrieb, Service zu ermöglichen, müssen Workflow-Systeme implementiert werden, welche von Kunden angestoßenen operativen Geschäftsprozesse kontrolliert ablaufen lassen.<sup>206</sup> Somit wird ermöglicht, gegenüber den Kunden Gewährleistungen und Zusagen über z. B. Liefer- und Besuchstermine, E-Mail-Benachrichtigung, Sonderkonditionen u. a. einzuhalten. Der Call Center-Agent hat z. B. Zugriff auf den Terminkalender des Außendienstes, und wenn der Kunde über das Call Center einen Termin vereinbart, werden diese Vereinbarung und alle dazugehörenden Informationen automatisch an den betreffenden Außendienstmitarbeiter weitergeleitet.

Bei Workflow-Systemen liegt der Schwerpunkt auf kundenbezogenen Geschäftsprozessen und durch einen virtuellen Workflow-Manager wird der Status bzw. Fortschritt jedes einzelnen Prozesses kontrolliert, den auch der Call Center-Agent kontinuierlich verfolgen kann.<sup>207</sup> Diese Tracking-Funktion soll gewährleisten, dass die Kundenwünsche bzw. -anforderungen so schnell wie möglich bearbeitet werden. Bei der Überschreitung von vordefinierten Zeitlimits für die Bearbeitung einer Kundenanfrage wird die Aufgabe automatisch an übergeordnete Stellen weitergeleitet (Eskalation).<sup>208</sup> Durch solche Systeme werden die Reaktionszeiten auf Kundenanfragen verkürzt, und damit auch eine hohe Servicequalität garantiert.<sup>209</sup>

### **3.6.3.2 Internet**

Das Internet wird immer populärer, weil es damit möglich ist, Unternehmen 24/7 virtuell zu präsentieren. Angesichts der stürmischen Entwicklung des Internets und damit auch des E-Commerce hat sich die Notwendigkeit ergeben, Internet als Kommunikationskanal in das CIC aufzunehmen. Somit wird den Kunden die Möglichkeit angeboten, mit dem Unternehmen jederzeit, sogar außerhalb der Betriebszeiten in Kontakt zu treten und ihre Geschäfte abzuwickeln.

Außerdem kann durch die erfolgversprechende Unternehmenspräsenz im Internet der gesamte Customer Buying Cycle<sup>210</sup>, welcher die Pre-Sales-, Sales- und After-Sales-

---

<sup>206</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 66.

<sup>207</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 66.

<sup>208</sup> Vgl. Hettich/Hippner/Wilde (2000), S. 1364 und Hippner/Rentzmann/Wilde (2006), S. 66.

<sup>209</sup> Vgl. Hippner/Rentzmann/Wilde (2006), S. 66.

<sup>210</sup> Der Customer Buying Cycle zeigt das Zusammenspiel zwischen dem Kunden und dem Unternehmen, aufgeteilt in verschiedene Phasen. Er hilft, den Prozeß des Kunden zu verstehen bzw. zu strukturieren, seine Bedürfnisse zu erkennen und mögliche Formen der IT-Unterstützung zu finden. Weiters dient das CBC-Modell dazu, eine systematische Betrachtung aller möglichen Berührungspunkte zwischen dem Anbieter und dem Kunden im Zusammenhang mit dem Kauf, Besitz und der Entsorgung einer Marktleistung darzustellen. Es ist ein geeignetes Instrument zur Visualisierung und Strukturierung der Schnittstellen zum Kunden, vgl. Muther (2001), S. 15 f.

Phase umfasst, abgedeckt werden.<sup>211</sup>

Vor dem Kauf (Pre-Sales-Phase) kann sich der Kunde durch eine sogenannte Marketing-Enzyklopädie über die angebotenen Produkte des Unternehmens informieren (siehe Kapitel 3.6.1.2 Sales-Automation). Amazon.com und noch viele weitere online Buchhändler stellen z. B. Kurzbeschreibungen über den Inhalt der Bücher und Reviews zur Verfügung. In der Sales-Phase kommt der bereits erwähnte Produktkonfigurator zum Einsatz. Es ist damit für den Kunden besonders einfach, sein gewünschtes Produkt individuell zusammenzustellen. Die von Kunden erteilten Aufträge werden dann an die entsprechenden Transaktionssysteme (Z. B. ERP-Systeme) weitergeleitet und dort bearbeitet. Nach dem Kauf (After-Sales-Phase) bekommt der Kunde ein Online Support zur Selbsthilfe sowie für eine Diagnose von Produktfehlern. Das erfolgt entweder durch die Einbindung eines Help Desk-Systems in das Internet, durch die Bereitstellung sogenannter FAQs (Frequently Asked Questions) oder durch die Funktion eines Call Back Button.<sup>212</sup> Über Call Back Button kann der Kunde darum bitten, dass ihn ein Call Center-Agent zurückruft.

Call Back Button ist ein einfaches Tool, um schnell den Kontakt zwischen dem Kunden und dem Berater herzustellen. Dieser Button ist auf der Unternehmenswebseite fest eingebunden. Der Kunde gibt seine Telefonnummer an und klickt einfach auf den Button. Seine Nummer wird durch eine Software automatisch an den entsprechenden Service-Mitarbeiter weitergeleitet, der dann so schnell wie möglich zu diesem Kunden Kontakt aufnimmt. Es ist möglich, diese Call Back-Funktionalität ebenfalls in den vorgelagerten Phasen einzusetzen.<sup>213</sup>

Außer Help Desk Systems, FAQs und CBBs bietet der VoIP-Kanal (Voice over Internet Protocol) eine interessante Anwendung, nämlich das sogenannte Shared Browsing. Shared Browsing erlaubt dem Call Center-Agenten, mit der Annahme des Webkontaktes den kontaktierenden Kunden zu führen. Beide synchronisieren ihre Bildschirme, und der Agent führt den Kunden durch die Homepage, wobei beide gleichzeitig über die weitere ISDN-Leitung miteinander sprechen oder im Falle einer analogen Leitung miteinander chatten können.

Alle diese Möglichkeiten der Kundeninteraktion im Rahmen des kommunikativen CRM ergibt die Auswertung der Kundenprofile, denn der Kunde hinterlässt bei jedem internetbasierten Kontakt mit dem Unternehmen aussagekräftige Informationen, die

---

<sup>211</sup> Vgl. Muther (2001), S. 13, 67-69.

<sup>212</sup> Vgl. Hippner/Wilde (2003a), S. 32.

<sup>213</sup> Vgl. Hippner/Wilde (2003a), S. 32.

über Cookies, Log Files gesammelt werden. Das Web Mining macht es möglich, daraus (Z. B. Kaufpräferenzen, individuelle Interessensgebiete, bevorzugte Zahlungsweisen u. a.) individualisierte Kundenprofile zu erstellen, aus denen das Unternehmen seine Geschäftsaktionen auf Kunden abstimmen ableiten kann.<sup>214</sup> Dies erfolgt z. B. durch die Gestaltung personalisierter Homepages basierend auf gewonnenen Erkenntnissen aus der Analyse des Surf- und Chatverhaltens von Kunden. Die persönliche Ansprache führt schließlich zu einer Erhöhung der Kundenzufriedenheit und Verbesserung der Kundenbindung.<sup>215</sup> Mittels derartiger personalisierter Homepages können außerdem jedem einzelnen Kunden Produkte und Dienstleistungen angezeigt werden, die zu denen passen, die er kaufen will oder gekauft hat (Cross Selling).

### **3.7 WETTBEWERBSVORTEILE DURCH CRM**

Wie schon in vorigen Kapiteln mehrfach erwähnt, unterstützt CRM Unternehmen bei der Sicherung von Wettbewerbsvorteilen durch eine verstärkte Kundenbindung und durch eine Erhöhung der Kundenprofitabilität. Dabei wird CRM unternehmensweit über alle Funktionen hinweg durch moderne Informations- und Kommunikationstechnologien unterstützt.

Die Abbildung 12 zeigt, wie sich die Vorteile der Informationstechnologie und die damit einhergehende Kundenorientierung für die Unternehmen lohnen.

Ein Vorteil liegt darin, besser auf die differenzierten Kundenbedürfnisse einzugehen. Somit kann jeder einzelne Kunde persönlich angesprochen werden.

Zunächst kommt es zu einer schnelleren Angebotserstellung, Auftragserfassung und –bearbeitung sowie einer Früherkennung von Marktchancen bzw. -trends durch die Analyse der Kundeninformationen. Alle die Faktoren, das Eingehen auf Kundenwünsche, die individuelle Ansprache und höhere Beratungskompetenz versprechen dem Kunden einen höheren Nutzen sowie ein geringeres Kaufrisiko.<sup>216</sup> Durch die individuelle Kundenbetreuung und Kundenbindung sowie ein frühzeitiges Erkennen von Ersatzbedürfnissen kann das Unternehmen Verständnis bei seinen Kunden für höhere Preise erzielen und neue Kunden gewinnen.<sup>217</sup> Mit zunehmender Dauer der Geschäftsbeziehung nimmt die Qualität des Wissens über den Kunden zu, was zu einer gezielten Kundenansprache führt und wodurch die Wahrscheinlichkeit erneuter Geschäftsabschlüsse steigt. Somit wird es dem Unternehmen möglich, seine

---

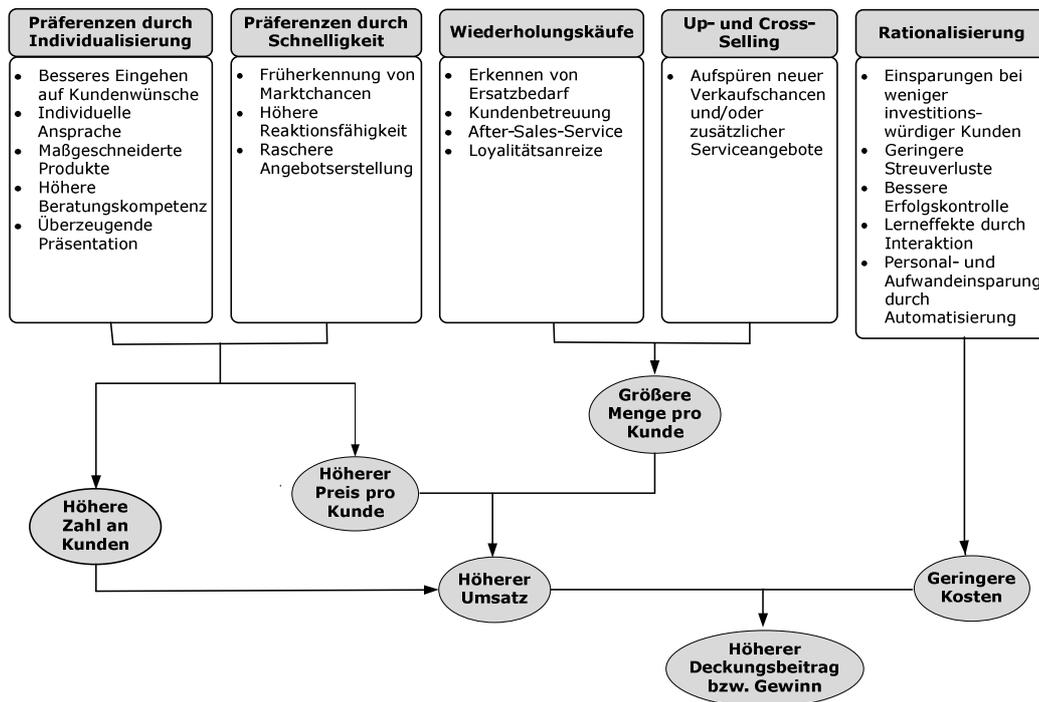
<sup>214</sup> Vgl. Hippner/Wilde (2003a), S. 33.

<sup>215</sup> Vgl. Muther (2001), S. 69.

<sup>216</sup> Vgl. Link/Hildebrand (1995), S. 18.

<sup>217</sup> Vgl. Link/Hildebrand (1995), S. 18.

Kunden zu Wiederholungskäufen und Cross Buying anzuregen, welche dazu beitragen, die Absatzmenge und den Umsatz pro Kunde zu steigern.<sup>218</sup>



**Abbildung 12: Kosten- und Nutzensvorteile durch CRM**

Quelle: Link/Hildebrand, (1995), S. 18.

Letztlich kommt es zu einer Rationalisierung. Durch die Identifikation weniger rentabler Kunden werden die Kosten gesenkt. So können z. B. die Marketingkosten (z. B. Erstellung teurer Kataloge, Außendienstbesuche) eingespart werden, die bisher für diese Kunden aufgewandt wurden. Die gezielte Direktansprache erfolgversprechender Kunden verringert wiederum Streuverluste von Werbemaßnahmen. Durch die zunehmende Interaktion zwischen Kunden und Unternehmen entstehen Lerneffekte, die zu einer Senkung der Kundenbetreuungskosten führen.<sup>219</sup> Schließlich entlastet die Automatisierung der administrativen Prozesse sowohl im Verkauf als auch in der Verwaltung das Verkaufspersonal von Routinearbeiten sowie manuellen Tätigkeiten (z. B. Selektieren von Adressen für Mailings) und trägt so zur Kostensenkung bei.<sup>220</sup>

Mit diesen oben erläuterten Kosten- und Nutzensvorteilen verschafft der CRM-Ansatz den Unternehmen einen enormen Wettbewerbsvorsprung. Weitere Wettbewerbsvorteile, die für die wirtschaftliche Anwendung von CRM-Systemen den Ausschlag geben können, sind die Effizienz- und Effektivitätssteigerung.<sup>221</sup>

<sup>218</sup> Vgl. Link/Hildebrand (1995), S. 18.

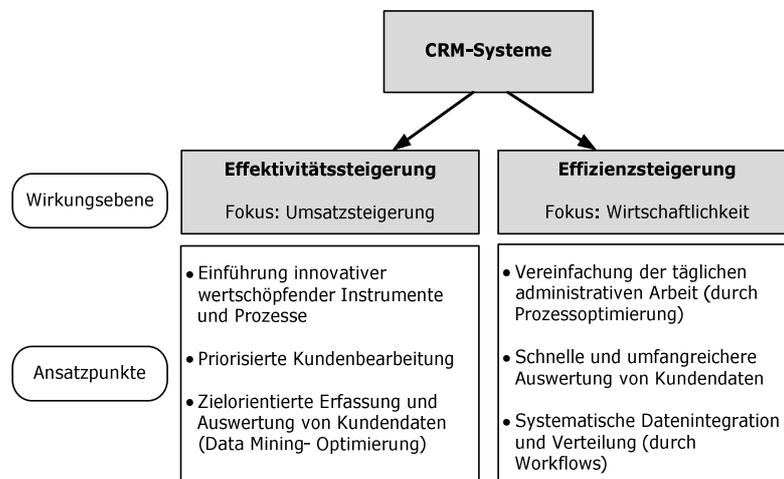
<sup>219</sup> Vgl. ebd., S. 18.

<sup>220</sup> Vgl. ebd., S. 18.

<sup>221</sup> Vgl. Holland (2004a), S. 181 und Dangelmaier/Helmke/Uebel (2001): „Trends im Customer Relationship Management (CRM)“ - Artikel erschien auf der Webseite Competence SITE, [Abruf: 14-11-2007].

Mittels CRM-Systeme werden Informationen über Kunden effizienter in der Unternehmensorganisation verteilt, und sie werden zur Gestaltung der Kundenbeziehungen genutzt. Weiterhin unterstützen CRM-Systeme die ganzheitliche Bearbeitung der Beziehungen des Unternehmens zu den Kunden.<sup>222</sup> Dabei soll zum einen eine höhere Effizienz in den entsprechenden Prozessen erzielt und zum anderen die Effektivität der Kundenbearbeitung verbessert werden.<sup>223</sup>

Die Effizienz- und Effektivitätssteigerungen resultieren beispielsweise aus einer Vereinfachung der täglichen administrativen Arbeit durch Prozessoptimierungen, aus einer systematischen Datenintegration und -verteilung oder einer gezielten Analyse dieser Daten<sup>224</sup> (siehe Abb. 13).



**Abbildung 13: Wirkungsebenen und Ansatzpunkte von CRM-Systemen**

Quelle: In Anlehnung an Dangelmaier/Helmke/Uebel (2001).

Die Optimierung der internen Back Office-Bearbeitungsprozesse, bspw. durch Workflowsfunktionalitäten zur automatisierten Verteilung von Informationen, ermöglicht den Servicemitarbeitern, im Front-Office-Bereich eine optimale, auf die Bedürfnisse des einzelnen Kunden zugeschnittene Leistung anzubieten.<sup>225</sup> Im Allgemeinen ist es möglich, hinsichtlich der Kundendatenbank Effektivitätssteigerungen zu erzielen, wenn es eine Konzentration auf wesentliche Informationen vorhanden ist und die Erfassung dieser Informationen auf Basis vordefinierter Ziele, wie z. B. die Bedeutung einer Information für die Kundenzufriedenheit oder die Aktualität der Daten, erfolgt.<sup>226</sup> Zuletzt ist die ganzheitliche Ausrichtung der Kundenbearbeitung an den Bedürfnissen der Kunden ein

<sup>222</sup> Vgl. Dangelmaier/Helmke/Uebel (2001): „Trends im Customer Relationship Management (CRM)“.

<sup>223</sup> Vgl. Dangelmaier/Helmke/Uebel (2001): „Trends im Customer Relationship Management (CRM)“.

<sup>224</sup> Vgl. Holland (2004a), S. 181.

<sup>225</sup> Vgl. ebd., S. 181.

<sup>226</sup> Vgl. ebd., S. 181.

wichtiger Punkt für die Effizienz- und Effektivitätssteigerung.<sup>227</sup>

### **3.8 ERFOLGSFAKTOREN IM CRM**

Für eine gute Kundenbetreuung ist die Einstellung des ganzen Unternehmens im Sinne einer durchgängigen Kundenorientierung und deren Umsetzung mit Hilfe geeigneter Abläufe und Systeme besonders wichtig.

Customer Relationship Management stellt hier einen ganzheitlichen Ansatz zur Unternehmensführung. Es integriert und optimiert auf der Grundlage einer Kundendatenbank und einer Software zur Marktbearbeitung sowie eines definierten Verkaufsprozesses abteilungsübergreifend alle kundenbezogenen Prozesse in Marketing, Vertrieb, Service und anderen Abteilungen. Eine erfolgreiche CRM-Einführung verbessert die Wettbewerbsposition des Unternehmens gegenüber Konkurrenten, indem es das Umsatz- und Ertragspotential erweitert.

Angesichts der zu erwartenden Potentiale und Nutzung von CRM-Systemen verwundert es nicht, dass mit jährlich steigenden Wachstumsraten in diese Systeme investiert wird. Schätzungen von Forschungsinstitution Forrester zufolge haben Unternehmen im Jahr 2005 über 3 Milliarden US-Dollar für neue Softwarelizenzen im Bereich CRM ausgegeben, und die Gesamtausgaben für CRM dürften sich sogar auf über 12 Milliarden US-Dollar belaufen haben.<sup>228</sup> Forrester und die Marktforscher von Gartner gehen davon aus, dass Unternehmen zur Verbesserung ihrer kundenbezogenen Prozesse auch in Zukunft viel Geld in CRM-Technologie investieren werden.<sup>229</sup> „Nach unrealistischen Erwartungen, gefolgt von extremem Pessimismus, wird CRM nun zu einer Kernaufgabe im Unternehmen, da Firmen so mit ihren Kunden besser in Kontakt treten können“, erläutern die Analysten von Forrester.<sup>230</sup>

Trotz all dieser großen Investitionsvorhaben in CRM ist es aber sehr erstaunlich, dass viele CRM-Projekte in der nahen Vergangenheit nicht so erfolgreich verliefen, wie es die betroffenen Unternehmen erwartet hatten. So wurden nach Einschätzung der Unternehmensberatung Booz Allen Hamilton mindestens 70% aller CRM-Projekte als Misserfolge verbucht. Zu diesem Ergebnis gelangte BAH in ihrer Studie, die sie im Jahr 2002 gemeinsam mit der Kellogg School of Management at Northwestern University durchgeführt hatte. Weiters erforschte die Forrester Group im Jahr 1999, dass nur

---

<sup>227</sup> Vgl. Holland (2004a), S. 181.

<sup>228</sup> Vgl. „CRM-Trends 2006 - ein Blick in die Kristallkugel“ - Artikel erschien auf der Webseite SAP Info am 06.03.2006, Autor: Barbara Gengler, [Abruf: 14-11-2007].

<sup>229</sup> Vgl. „CRM-Trends 2006 - ein Blick in die Kristallkugel“.

<sup>230</sup> Laut der Artikel „CRM-Trends 2006 - ein Blick in die Kristallkugel“ sind diese und andere Details in dem von Forrester-Marktforschern zusammengefassten Bericht „Trends 2006: CRM“ zu finden.

14% der CRM-Projekte die Erwartungen der Unternehmen erfüllten. Schließlich zeigten die Untersuchungen von Gartner Group, dass rund 60% aller CRM-Projekte scheiterten. Die Ursache für das mögliche Scheitern von CRM-Projekten liegt vor allem daran, dass die Einführung komplexer CRM-Systeme neben einer kundenorientierten Ausrichtung des Unternehmens, gegebenenfalls die Neuorientierung verschiedener organisatorischer und technologischer Prozesse erfordert.<sup>231</sup> So sind kundenorientierte Reorganisationsmaßnahmen im Unternehmen durchzuführen, welche die Optimierung kundenbezogener Geschäftsprozesse, deren Unterstützung durch CRM-Systeme sowie die systematische Planung und Kontrolle der Veränderungen durch ein Projektmanagement und ein Change Management umfassen.<sup>232</sup>

In der Literatur werden zahlreiche Problembereiche bei der Implementierung von CRM-Systemen benannt.<sup>233</sup> Im Folgenden werden einige von diesen Problembereichen kurz erläutert:

- *Datenqualität*: Die Datenqualität ist eine entscheidende Komponente für den erfolgreichen Einsatz einer CRM-Lösung. Oft sind Kundendaten über mehrere Systeme innerhalb einer Firma verstreut. In einer Umfrage von IBM Global Data Management bei 600 führenden Unternehmen berichteten 75% der befragten Unternehmen von beträchtlichen Problemen aufgrund fehlerhafter Daten.<sup>234</sup> Gute Datenqualität heißt, dass die Daten aktuell, korrekt und redundanzfrei sind. Nur wenn solche Daten als Input zur Verfügung stehen, können sinnvolle Analysen und Prognosen basierend auf diesen Daten erstellt werden.<sup>235</sup> Die Datenqualität sicherzustellen, ist die Hauptherausforderung des CRM.<sup>236</sup>
- *Integration der CRM-Systeme*: Ein Schwachpunkt bisher betriebener CRM-Systeme ist deren fehlende Integration in die vorhandene Systemumgebung der Unternehmen.<sup>237</sup> Viele CRM-Systeme laufen daher bislang isoliert und können somit das vorhandene Datenmaterial anderer Unternehmensbereiche nicht nutzen.<sup>238</sup> „Im CRM muss neben der horizontalen Integration mit anderen

---

<sup>231</sup> Vgl. Holland (2004), S. 208.

<sup>232</sup> Vgl. Hippner (2006), S. 21.

<sup>233</sup> Mögliche Probleme sind z. B. eine falsche CRM-Strategie und CRM-Philosophie, Fehler bei der Implementation und organisatorischen Verankerung von CRM-Systemen und die Verwendung falscher CRM-Software, vgl. Schaller/Stotko/Piller 2004, S.125 und Dangelmaier/Helmke/Uebel 2004, S. 12 f.

<sup>234</sup> Vgl. o. V. (2007), „Die acht Grundlagen von CRM: Ein Handbuch für Führungskräfte zu den acht unabdingbaren Elementen jeder erfolgreichen CRM-Initiative“ - Whitepaper (2007) von Salesforce.com.

<sup>235</sup> Vgl. Zellner (2002), S. 7 f.

<sup>236</sup> Vgl. ebd., S. 7 f.

<sup>237</sup> Vgl. ebd., S. 8.

<sup>238</sup> Vgl. ebd., S. 8.

operativen Applikationen auch die Dimension der vertikalen Integration berücksichtigt werden. Während die horizontale Integration primär die Integration von Applikationen entlang des Kundenprozesses umfasst, adressiert die vertikale Integration die Verbindung und Verdichtung von Daten z. B. in einem Data Warehouse zu Analysezwecken.“<sup>239</sup>

- *Nutzerakzeptanz*: CRM-Projekte scheitern wegen geringer Akzeptanz unter den Benutzern. Schließlich ist die CRM-Technologie nur so gut wie seine Benutzer, die sie verwenden. „Die mit der Einführung eines CRM-Systems i. d. R. verbundenen Veränderungen im Unternehmen, wie beispielsweise die Notwendigkeit neuer Arbeitstechniken, können zu Widerständen und Ängsten, z. B. vor Rationalisierungsmaßnahmen bzw. Arbeitsplatzverlust, auf Seiten der Mitarbeiter führen.“<sup>240</sup> Für eine erfolgreiche Einführung ist jedoch deren Akzeptanz maßgeblich.<sup>241</sup> Ein den Einführungsprozess begleitendes Change Management, welches darauf zielt, eventuell entstehende Ängste und Widerstände abzubauen, die Mitarbeiter zu motivieren und zu schulen, ist daher unabdingbar.<sup>242</sup>
- *Einbindung der Mitarbeiter*: Einer der wichtigsten Erfolgsfaktoren bei der CRM-Einführung ist die frühzeitige Einbindung der Mitarbeiter, die am Ende mit den Ergebnissen arbeiten müssen. „Die CRM-Aktivitäten werden von den Mitarbeitern getragen, denn sie stellen die Schnittstelle zum Kunden dar und transportieren den CRM-Gedanken“<sup>243</sup> und „entscheiden letztlich über den Erfolg einer CRM-Einführung“<sup>244</sup>. Es ist wichtig, klar darlegen zu können, welchen Nutzen das CRM-System für den einzelnen Mitarbeiter darstellt.<sup>245</sup> Denn, nur wenn durch das CRM-System der Arbeitsaufwand reduziert oder das Arbeitsergebnis verbessert wird, sind die Mitarbeiter dazu bereit, den mit der Pflege des CRM-Systems verbundenen Mehraufwand in Kauf zu nehmen. Es ist allerdings nicht selten, dass die Mitarbeiter bei der CRM-Einführung nicht ausreichend eingebunden sind. Deshalb ist es sinnvoll, sicherzustellen, dass alle Betroffenen von Beginn an in die Projektarbeit und die Konzeptionsphase

---

<sup>239</sup> Vgl. Puschmann/Alt (2002), S. 28.

<sup>240</sup> Vgl. Holland (2004a), S. 209.

<sup>241</sup> Vgl. ebd., S. 211.

<sup>242</sup> Vgl. ebd., S. 209.

<sup>243</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“ - Artikel, veröffentlicht am 11/2006 auf der Webseite Competence SITE, [Abruf: 10-11-2007].

<sup>244</sup> Vgl. Winkelmann (2002), zitiert nach: „Zehn Bausteine für den CRM-Erfolg“ - Artikel, erschien am 07.11.2003 auf der Webseite von Computerwoche, [Abruf: 15-11-2007].

<sup>245</sup> Vgl. Schwetz (2007): „Kundenorientierung ist Chefsache“ - Artikel, erschien am 03/2007 auf der Webseite von Computerwoche, [Abruf: 05-11-2007]

einbezogen sind.<sup>246</sup>

- *Unterstützung durch das Top-Management:* Die Einführung eines CRM-Projektes beeinflusst auf der einen Seite die Geschäftsprozesse und die IT-Infrastruktur und auf der anderen die Unternehmenskultur und die Interaktion zwischen Kunden und Mitarbeitern.<sup>247</sup> Daher bedingen sie eine umfangreiche organisatorisch-kulturelle Umstrukturierung innerhalb eines Unternehmens. Da diese Umstrukturierung oft eine unternehmensweite Umstellung der Aufbau- und Ablauforganisation erfordert<sup>248</sup>, muss das obere Management CRM-Projekte voll unterstützen.<sup>249</sup>

In diesem Zusammenhang braucht jede CRM-Einführung einen Sponsor aus der oberen Geschäftsführung.<sup>250</sup> Er vertritt den CRM-Gedanken im Unternehmen<sup>251</sup> und sorgt dafür, dass sich sämtliche betroffenen Unternehmensbereiche für CRM engagieren.<sup>252</sup>

- *Reduzierung der Komplexität der CRM-Systeme:* Zur Erhöhung der Nutzerakzeptanz ist die Komplexität der Funktionalitäten von CRM-Systemen zu vereinfachen.<sup>253</sup> Das Ziel ist nicht, möglichst viele Funktionalitäten zusammenzustellen, sondern – entsprechend dem Grundsatz *technology follows function* – solche Funktionalitäten auszuwählen, die einen möglichst großen Anwendungsnutzen für das Unternehmen stiften.<sup>254</sup> Zudem sollte sich die Einführung zunächst auf wenige wesentliche Funktionalitäten beschränken, um den Erfolg nicht durch Überkomplexität oder zu großen Implementierungsaufwand zu riskieren (*think big, start small*)<sup>255</sup> Erst nachdem die Anwender das System in der Praxis kennen gelernt haben, können die Funktionalitäten je nach Bedürfnis erweitert werden.
- *Der Wille alles auf einmal einzuführen:* Dagegen spricht, dass eine Einführung in kleinen Schritten erfolgversprechend ist.<sup>256</sup> Zunächst sollte eine solide CRM-

---

<sup>246</sup> Vgl. Schwetz (o. J.): „CRM - Der Weg zum kundenorientierten Unternehmen“ - Artikel im Internet, [Abruf: 05-11-2007].

<sup>247</sup> Vgl. Kehl/Rudolph (2001), S. 260.

<sup>248</sup> Vgl. Ematinger/Sommer/Stengl (2001), S. 49 und Puschmann/Alt (2002), S. 16.

<sup>249</sup> Vgl. Kehl/Rudolph (2001), S. 260 und Kreuz/Förster (2001), S. 17 f.

<sup>250</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

<sup>251</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

<sup>252</sup> Vgl. Holland (2004a), S. 208.

<sup>253</sup> Vgl. Zellner (2002), S. 8.

<sup>254</sup> Vgl. Holland (2004a), S. 210.

<sup>255</sup> Vgl. Holland (2004a), S. 211.

<sup>256</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

Basis durch die Akzeptanz der CRM-Software, der neuen Abläufe und CRM-Ausrichtung geschaffen werden,<sup>257</sup> und die daraus gewonnenen Erfahrungen der aktiven Anwender bilden dann die Grundlage für die weiteren CRM-Maßnahmen. Eine sukzessive Einführung des CRM-Systems, begleitet von Schulungen der Anwender, gewährleistet die problemlose Integration in das Tagesgeschäft.<sup>258</sup>

- *Zu hohe Erwartungen:* Die Erwartungen, die mit der Einführung von CRM-Systemen verbunden sind, werden meistens zu hoch eingeschätzt.<sup>259</sup> Was an erster Stelle bei einer CRM-Einführung zu erwarten ist, sind Effizienzverbesserungen in Vertriebs- und Kundenprozessen sowie eine rasche Amortisation der Kosten für die technologischen Komponenten.<sup>260</sup>
- *Keine klaren Zielvorstellungen:* Der Nutzen und die Chancen von CRM und einer Kundenorientierungsstrategie für das eigene Unternehmen werden nicht gut genug erkannt, da es beim Management oft an Zukunftsvisionen und klaren Zielen fehlt.<sup>261</sup> Vor der Einführung eines CRM-Systems müssen aber die mit CRM verfolgten, messbaren Ziele des Unternehmens klar definiert sein.<sup>262</sup> Das Management überlässt dieses Thema meist der IT-Abteilung.<sup>263</sup> Hier ist es aber zu beachten, dass eine CRM-Einführung kein IT-Projekt ist.<sup>264</sup> Jede CRM-Einführung muss als ein strategisches Projekt verstanden werden, welches weit mehr ist als nur die Einführung einer Kundenmanagement-Software.<sup>265</sup> Deshalb sollte das CRM-Projekt neben IT-Fachleuten auch Experten anderer Fachabteilungen und potenzielle Anwender mit einbeziehen.<sup>266</sup> Ferner sollten die Verantwortlichkeiten klar zugeordnet und eine zielorientierte Projektplanung festgelegt werden.<sup>267</sup>

In weiterer Folge können sechs kritische Erfolgsfaktoren genannt werden, die beim Umgang mit den oben angeführten Problemen bei der Einführung von CRM-Systemen

---

<sup>257</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

<sup>258</sup> Vgl. Holland (2004a), S. 211.

<sup>259</sup> Vgl. Zellner (2002), S. 8 und Kehl/Rudolph (2001), S. 261.

<sup>260</sup> Vgl. Kehl/Rudolph (2001), S. 261.

<sup>261</sup> Vgl. Schwetz (2007): „Kundenorientierung ist Chefsache“.

<sup>262</sup> Vgl. Schwetz (o. J.): „CRM - Der Weg zum kundenorientierten Unternehmen“.

<sup>263</sup> Vgl. Schwetz (2007), „Kundenorientierung ist Chefsache“.

<sup>264</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“ und Schwetz (2007): „Kundenorientierung ist Chefsache“.

<sup>265</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

<sup>266</sup> Vgl. Holland (2004b), S. 21 und Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“

<sup>267</sup> Vgl. Holland (2004a), S. 209.

zu beachten sind.

Diese Erfolgsfaktoren sind auf der nächsten Seite tabellarisch dargestellt.

Erfolgsfaktor	Beschreibung
<b>(1) Organisation</b>	Die Anpassung der Organisation an die neue Zielsetzung, an neue Vorgehensweisen und Technologien ist notwendig. <sup>268</sup> Die Aufbau- und Ablauforganisationen richten sich an den neuen kundenorientierten Prozessen wie z. B. in den Bereichen Marketing, Vertrieb und Service aus.
<b>(2) CRM-Prozess</b>	„Die Integration des Kundenprozesses über alle Kundenkontaktpunkte und alle Kanäle hinweg ist der Zweck von CRM. Eine einheitliche Sicht auf den Kunden ist nur dann möglich, wenn alle Kundenkontaktpunkte über denselben Informationsstand bezüglich des Kunden verfügen. Prozessintegration endet daher nicht in der Integration von Front-Office Prozessen in Marketing, Vertrieb und Service, sondern schließt die Integration nachgelagerter Supply Chain Prozesse mit ein.“ <sup>269</sup>
<b>(3) Systemarchitektur</b>	Der Integration der Prozesse folgt die vollständige Integration der Systeme, um ein unternehmensweites CRM-System zu erreichen. <sup>270</sup> Die Kundendatenbank spielt hier eine wichtige Rolle, indem sie ermöglicht, eine homogene Sicht auf einzelne Kundendaten zu erstellen. <sup>271</sup>
<b>(4) Projektmanagement</b>	„Bei der CRM-Realisierung handelt es sich um einen komplexen Prozess der Organisationsentwicklung, in dessen Verlauf verschiedenen Kernkompetenzen pünktlich integriert werden müssen. Der Bewertung von CRM-Projekten im Rahmen eines CRM-Controlling kommt eine zentrale Rolle im CRM-Projektmanagement zu.“ <sup>272</sup>
<b>(5) Change-management</b>	Die Umsetzung einer CRM-Strategie und die daraus resultierenden strukturellen Veränderungen im Unternehmen erfordert neben dem Commitment der oberen Geschäftsführung auch das volle Engagement aller Mitarbeiter im Unternehmen. <sup>273</sup> Hierfür ist ein den CRM-Einführungsprozess begleitendes adäquates Change Management nötig, welches darauf zielt, eventuell entstehende Ängste und Widerstände der Mitarbeiter abzubauen, sie zu motivieren und zu schulen. <sup>274</sup> „Um das Change Management erfolgreich umsetzen zu können, müssen alle betroffenen Abteilungen frühzeitig in das Projekt integriert werden.“ <sup>275</sup>
<b>(6) Kultur</b>	Wenn eine Unternehmenskultur bereits vorhanden ist, können dann auch Geschäftsprozesse und Mitarbeiter im Unternehmen erfolgreich auf Kundenorientierung ausgerichtet werden. <sup>276</sup> So „bedarf die Unternehmenskultur einer konsequenten Ausrichtung auf den Kunden. Kundenorientierung als Maxime für alle Mitarbeiter ist einer der entscheidenden Aspekte für den Erfolg eines Unternehmens. Der Nutzen einer offenen und ganzheitlichen Kundenbetrachtung muss transparent bzw. für jeden Mitarbeiter nachvollziehbar sein.“ <sup>277</sup> Vor diesem Hintergrund ist eine bewusste Konfrontation und Kommunikation mit den Mitarbeitern wünschenswert, da die Mitarbeiter schließlich der Motor des CRM-Erfolgs sind. <sup>278</sup>

**Tabelle 2: Erfolgsfaktoren bei der Einführung von CRM-Systemen**

Quelle: eigene Darstellung

<sup>268</sup> Vgl. Puschmann/Alt (2002), S 35.<sup>269</sup> Vgl. Puschmann/Alt (2002), S 35 f.<sup>270</sup> Vgl. Zellner (2002), S. 8.<sup>271</sup> Vgl. Puschmann/Alt (2002), S 36.<sup>272</sup> Vgl. Hippner (2006), S. 34.<sup>273</sup> Vgl. Hippner (2006), S. 34.<sup>274</sup> Vgl. Holland (2004b), S. 23.<sup>275</sup> Vgl. Holland (2004b), S. 21.<sup>276</sup> Vgl. Schwetz (o. J.): „CRM - Der Weg zum kundenorientierten Unternehmen“.<sup>277</sup> Vgl. Kehl/Rudolph (2001), S. 261.<sup>278</sup> Vgl. Schneider (2004): „Die häufigsten Stolpersteine bei der CRM-Einführung“.

# KAPITEL 4

---

## KNOWLEDGE MANAGEMENT (KM)

### 4.1 EINFÜHRUNG

*„The essence of management is to make knowledge productive.“*

*[Peter F. Drucker]*

*„For knowledge too, is itself  
power.“*

*[Francis Bacon, englischer  
Philosoph]*

Die steigende Intensität des Wettbewerbs, verbunden mit der zunehmenden Globalisierung, Internationalisierung sowie Liberalisierung der Wirtschaft stellen Unternehmen ständig vor neue und große Herausforderungen. Die Fähigkeit, zentrale Wettbewerbsinformationen und erfolgskritische Unternehmensdaten sowie deren Ursache-Wirkungsbeziehungen im Blick zu haben, wird zur Grundvoraussetzung für den Erhalt nachhaltiger Wettbewerbsvorteile. Die richtigen Informationen zur richtigen Zeit am richtigen Ort bestimmen den Unternehmenserfolg, indem sie zum spezifischen, unternehmerischen Know-how werden.

Wissen - als ein kritischer, entscheidender Wettbewerbsfaktor für den Unternehmenserfolg - muss ebenso wie die herkömmlichen Produktionsfaktoren Rohstoff, Kapital und Arbeit bewirtschaftet werden, um die Wachstums- und Kostensenkungspotentiale effizient ausschöpfen zu können. Somit wird Wissen zum vierten Produktionsfaktor.<sup>279</sup> Es wird als einzig verbleibender Wettbewerbsvorteil der

---

<sup>279</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 242.

entwickelten Länder im globalen Wettbewerb erkannt<sup>280</sup> oder sogar als Basis für ein neues Paradigma des Managements gesehen<sup>281</sup>.

Für eine wissensintensive Wertschöpfung ist es besonders wichtig, Information und Wissen als strategische Ressourcen im Geschäftsprozess, im Produkt und als Produkt zu nutzen.<sup>282</sup> „Wissen wird zum Motor und zur entscheidenden Größe im Wertschöpfungsprozess.“<sup>283</sup> Aus diesem Grund ist es für Unternehmen sehr wichtig, dass sie all ihre vorhandenen Informationen intelligent beherrschen und aus denen für das Unternehmensgeschäft relevantes Wissen generieren. Das intellektuelle Potential des Unternehmens muss gezielt und effektiv genutzt und gestaltet werden, um Wettbewerbs- und Marktchancen zu erreichen. Demzufolge kann unter dem Begriff *Knowledge Management (KM)* - der bewusste Umgang mit dem Rohstoff Wissen (Knowledge) und dessen zielgerichtetem Einsatz im Unternehmen verstanden werden.<sup>284</sup>

Um ein theoretisches Standbein für diese Arbeit zu schaffen, wird in diesem Kapitel erst der Wissensbegriff anhand verschiedener Definitionen und seiner im Vergleich zu anderen Produktionsfaktoren besonderen Eigenschaften näher spezifiziert und von verwandten Begriffen abgegrenzt. Danach wird das Knowledge Management methodisch eingeführt und anschließend detaillierter auf die *Business Intelligence (BI)* - welche die technische Gestaltungsdimension des Knowledge Management darstellt - sowie die im CRM eingesetzten BI-Systeme und -Konzepte (wie z. B. Data Marts und Data Warehouse) eingegangen.

## **4.2 DEFINITIONEN und EIGENSCHAFTEN DES WISSENS**

### **4.2.1 Was ist Wissen?**

Zunächst muss genau geklärt werden, was mit dem Begriff Wissen gemeint ist. Die Definition des Begriffes Wissen dient als Grundlage und Ausgangspunkt für jede Auseinandersetzung mit dem Thema Wissensmanagement.

Ein einheitliches Verständnis über die Bedeutung dieses Begriffs lässt sich jedoch in der Literatur nicht leicht ausmachen. Je nach Autor und dessen speziellem Fokus gibt es ganz unterschiedliche Definitionen.

---

<sup>280</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 242.

<sup>281</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 242.

<sup>282</sup> Vgl. Bullinger/Wagner/Oldhausen (2000), S. 77.

<sup>283</sup> Bullinger/Wagner/Oldhausen (2000), S. 77.

<sup>284</sup> Vgl. Dittmar/Gluchowski (2002), S. 27.

In ihrem Buch „Wissen managen“ definieren Probst, Raub und Romhardt Wissen folgendermaßen:

„Wissen bezeichnet die Gesamtheit der Kenntnisse und Fähigkeiten, die Individuen zur Lösung von Problemen einsetzen. Dies umfasst sowohl theoretische Kenntnisse als auch praktische Alltagsregeln und Handlungsanweisungen. Wissen stützt sich auf Daten und Informationen, ist im Gegensatz zu diesen jedoch immer an Personen gebunden. Es wird von Individuen konstruiert und repräsentiert deren Erwartungen über Ursache-Wirkungs-Zusammenhänge.“<sup>285</sup>

Von Davenport und Prusak wird Wissen wie folgt definiert:

„Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms.“<sup>286</sup>

Auf Informationen und somit auch Wissen, insbesondere Kundenwissen, trifft eine Reihe von Charakteristika zu, die klassische Unternehmensressourcen bzw. Produktionsfaktoren nicht aufweisen:

- „Knowledge is intangible and difficult to measure,
- Knowledge is volatile, that is, it can ‚disappear‘ overnight,
- Knowledge is, most of the time, embodied in agents with wills,
- Knowledge is not ‚consumed‘ in a process, it sometimes increases through use,
- Knowledge has wide ranging impacts in organizations (e. g. ‚knowledge is power‘),
- Knowledge can not be bought on the market at any time, it often has long lead times,
- Knowledge is ‚non-rival‘, it can be used by different processes at the same time.“<sup>287</sup>

---

<sup>285</sup> Probst/Raub/Romhardt (1999), S. 48.

<sup>286</sup> Davenport/Prusak (2000), S. 5.

<sup>287</sup> Wiig/de Hoog/van der Spek (1997), S. 16.

Nach Krogh und Roos lässt sich Wissen im Weiteren abstrakt folgendermaßen charakterisieren<sup>288</sup>:

- *Knowing is distinction making*: Wissen beschreibt die Welt und unterteilt in Bekanntes und Unbekanntes, Negatives und Positives, Erwartetes und Unerwartetes usw. Diese Grundcharakteristik des Wissens als individuelle Unterscheidung weist darauf hin, dass Wissen stets individuell und kontextabhängig ist.<sup>289</sup>
- *Knowing is caring*: Es reicht nicht, die Welt nur zu beobachten und zu beschreiben; um sie zu verstehen, sind jedoch Interesse, Aufmerksamkeit und Interpretation wesentliche Voraussetzungen. Demnach ist Wissen nicht reine Information, sondern deren zweckgebundene Vernetzung.<sup>290</sup>
- *Knowing is languaging*: Grundlage menschlicher Denkprozesse sind modellhafte Vorstellungen über reale Zusammenhänge, sogenannte mentale Modelle, die wiederum einen Wissensbestand darstellen.<sup>291</sup> Diese haben aufgrund unterschiedlicher Erfahrungen und Interpretationen individuell verschiedene Ausprägungen, so dass ein Begriff mit individuell unterschiedlichen Konnotationen versehen sein kann, was Kommunikation und Wissenstransfer erschwert. Wissen beinhaltet somit auch eine sprachschöpferische Komponente, was Wortprägungen wie „Infotainment“ oder „Datenautobahn“ oder alle Arten von unternehmensspezifischen Jargons beispielhaft belegen.
- *Knowing is Shaping the Future*: Die Möglichkeit, unter Heranziehung aktuellen Wissens alternative Zukünfte zu entwerfen und danach ausgerichtete Handlungsalternativen zu bewerten, bildet die Grundlage für die Gestaltung der Zukunft.

Die obigen Definitionen führen letztlich zu einer Zusammenfassung der zentralen Punkte von Wissen:

- Wissen stellt eine unerschöpfliche Ressource dar und kann physisch nicht abgenutzt werden.<sup>292</sup>
- Wissen wird durch Nutzung und die daraus gewonnenen Erfahrungen immer

---

<sup>288</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 245 f.

<sup>289</sup> Vgl. ebd., S. 245.

<sup>290</sup> Vgl. ebd., S. 245.

<sup>291</sup> Vgl. ebd., S. 246.

<sup>292</sup> Vgl. Hannig (2002), S. 15.

größer.<sup>293</sup>

- Wissen besteht aus vernetzten, kontextualisierten Informationen.
- Wissen ist personenbezogen.
- Wissen ist zweckorientiert.
- Wissen entsteht und findet Anwendung in den Köpfen der Wissensträger.
- In Organisationen wird es nicht nur in Dokumenten und Repositorien, sondern auch in organisatorischen Routinen, Prozessen, Praktiken und Normen verankert.

Um einen nachhaltigen Wettbewerbsvorteil gegenüber anderen Konkurrenten sichern zu können, müssen die einzusetzenden Unternehmensressourcen über spezielle Eigenschaften, wie z. B. Wertgenerierung, Knappheit, geringe Abnutzbarkeit und Transferierbarkeit verfügen. Außerdem dürfen sich die Ressourcen nicht einfach imitieren bzw. substituieren lassen. Der Aspekt der Imitier- bzw. Substituierbarkeit spielt bei Ressourcen, die auf den Faktormärkten schwer zu erhalten sind, eine überragende Rolle - insbesondere wenn von Unternehmen zum Schutz ihrer eigenen Erfolgsressourcen „Imitationsbarrieren“ errichtet worden sind. Ein großer Anteil dieser Imitationshindernisse hängt direkt oder indirekt mit Wissen bzw. der Entwicklung von Wissen zusammen. Dies, sowie die Tatsache, dass bezüglich aller oben angeführten Aspekte intangible Ressourcen (und damit auch Wissen) tangiblen Ressourcen überlegen sind, deuten auf die Rolle von Wissen bei der Erzielung nachhaltiger Wettbewerbsvorteile hin.

Die immaterielle Natur des Wissens begründet zum Teil die Schwierigkeiten des Wissensmanagements in der Praxis. Was immateriell ist, kann schließlich schwer gemessen und gemanagt werden.

#### **4.2.2 Wie entsteht Wissen?**

Ein Ansatz, der im Informationsmanagement für die Abgrenzung von Wissen zu verwandten Begriffen hilfreich sein kann, geht von einer hierarchischen Konzeption aus (siehe Abb. 14).

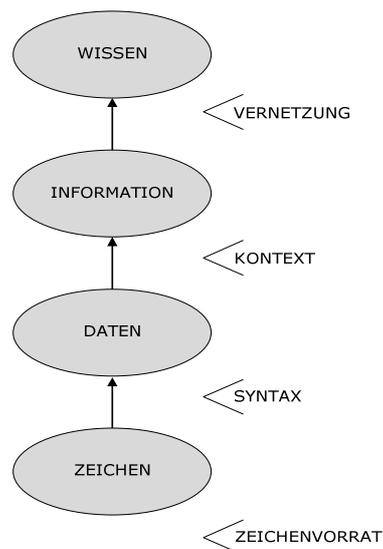
Auf der untersten Ebene sind die Zeichen angesiedelt. Ein Zeichen ist das kleinste Datenelement, auf das bei einer Programmausführung zugegriffen wird. Zeichen

---

<sup>293</sup> Vgl. Hannig (2002), S. 15.

können z. B. Buchstaben eines Alphabets, numerische Ziffern oder Sonderzeichen sein. Die Menge aller zur Verfügung stehenden Zeichen wird als Zeichenvorrat bezeichnet. Aus diesem Zeichenvorrat können einzelne Zeichen oder Sonderzeichen herausgenommen werden, die dann jedoch alleinstehend sind und daher in keinem Zusammenhang zueinander stehen.<sup>294</sup> Zeichen werden durch Syntaxen<sup>295</sup>, d. h. einer Ordnungsregel, zu Daten.

Eine Stufe darüber liegen die Daten. Sie werden durch Zeichen repräsentiert, die zur Verarbeitung dargestellt werden.<sup>296</sup> „Daten sind das Gegebene zur Verwendung ohne Verwendungshinweise.“<sup>297</sup> Sie können sowohl aus einzelnen Zeichen bestehen als auch aus einer Folge von Zeichen, die in einem sinnvollen (bekannten oder unterstellten) Zusammenhang zueinander stehen.<sup>298</sup> Trotzdem kann man auf dieser Stufe der Begriffshierarchie keine Aussage über ihre Bedeutsamkeit bzw. Belanglosigkeit treffen.<sup>299</sup> Daten entstehen im Unternehmenskontext als strukturierte Beschreibungen bei Vorgängen, Ereignissen oder Transaktionen<sup>300</sup>, und sie werden zum Großteil in technischen Systemen gesammelt und verwaltet. Somit stellen sie das Rohmaterial für Informationen dar.



**Abbildung 14: Ebenen der Begriffshierarchie**

Quelle: in Anlehnung an Rehäuser/Krcmar (1996), S. 6.

„Informationen sind als in den Kontext eines Problemzusammenhangs gestellte Daten

<sup>294</sup> Vgl. Rehäuser/Krcmar (1996), S. 4.

<sup>295</sup> Die Syntax gibt die Muster vor, nach denen man einzelne Zeichen kombinieren kann.

<sup>296</sup> Vgl. Rehäuser/Krcmar (1996), S. 4.

<sup>297</sup> Vgl. ebd., S. 4.

<sup>298</sup> Vgl. ebd., S. 4.

<sup>299</sup> Vgl. Rehäuser/Krcmar (1996), S. 4 und Ackerschott (2001), S. 10.

<sup>300</sup> Vgl. Ackerschott (2001), S. 10.

zu betrachten.<sup>301</sup> Daten werden zu Informationen, sobald sie in einen Problembezug eingeordnet und für die Erreichung eines Zieles verwendet werden.<sup>302</sup> Dies geschieht über die Transformation von Daten in Informationen durch das Individuum, z. B. innerhalb einer Kommunikation. Es ist durchaus klar, dass eine Kommunikation zwei Individuen, einen Sender und Empfänger einer Nachricht voraussetzt. Der Sender schickt so dem Empfänger eine Nachricht (message), und der Empfänger entscheidet dann, ob es sich bei den in Form einer Nachricht zu ihm gesendeten Daten für ihn tatsächlich um eine Information handelt. Für eine korrekte Aufwertung von Daten zu Informationen sind aber folgende Punkte zu berücksichtigen:<sup>303</sup>

- *Kontextualisierung*: Dem Empfänger ist der Zweck, für den die Daten vorliegen, ebenso bekannt wie das Ziel, das mit ihrer Bereitstellung verfolgt wird.
- *Kategorisierung*: Die Komponenten des Datenmaterials und ihr Zusammenhang sind dem Empfänger bekannt.
- *Berechnung*: Das Datenmaterial wurde durch exakte Beobachtung, (mathematische) Analyse oder (statistische) Auswertung ermittelt.
- *Korrektur*: Daten sollten bereinigt sein, d. h., fehlerhafte Daten wurden aus der Gesamtheit eliminiert.
- *Zusammenfassung*: Einzelne Daten wurden in einer bedeutungsreicheren Form verdichtet.

„Wissen entsteht, wenn Menschen Informationen aufnehmen und diese mit bereits verfügbaren Informationen unter Einbeziehung ihrer Einstellungen, Fertigkeiten und Erfahrungen vernetzen.“<sup>304</sup> Steinmüller bezeichnet Wissen als die „zweckorientierte Vernetzung von Informationen“.<sup>305</sup> Es wird als eine Anhäufung von Informationen gesehen. Informationen haben einen punktuellen Charakter, und erst wenn neue Informationen mit anderen oder früher gespeicherten Informationen vernetzt werden können, werden sie für das Individuum wertvoll, da sie sein Wissen erweitern oder verändern können.<sup>306</sup> Aus der systematischen Verknüpfung von Informationen resultierende Wissen ermöglicht uns, Sachverhalte zu verstehen und zu erklären, Entwicklungen zu prognostizieren, Entscheidungen zu treffen oder in der optimalen Art

---

<sup>301</sup> Rehäuser/Krcmar (1996), S. 5.

<sup>302</sup> Vgl. Rehäuser/Krcmar (1996), S. 5.

<sup>303</sup> Vgl. Davenport/Prusack (2000), S. 4 und Ackerschott (2001), S. 12.

<sup>304</sup> Ackerschott (2001), S. 13.

<sup>305</sup> Steinmüller (1993), S. 236.

<sup>306</sup> Vgl. Ackerschott (2001), S. 14.

und Weise zu handeln.<sup>307</sup>

### 4.2.3 Formen von Wissen

Wissen kann in unterschiedlichen Formen vorliegen. Vor dem Hintergrund einer strategischen Perspektive scheint es sinnvoll zu sein, Wissen nach Explikationsgrad (explizit-implizit) und Lokalisation (individuell-kollektiv) zu kategorisieren.<sup>308</sup>

Die japanischen Forscher Nonaka und Takeuchi<sup>309</sup> unterscheiden in ihrem Werk „The Knowledge-Creating Company“ zwischen implizitem und explizitem Wissen. Diese Unterscheidung geht auf den Wissenschaftler (Medizin/Physik) Michael Polanyi zurück.

Polanyi stellte schon 1966 fest, dass Menschen mehr wissen, als sie anderen mitteilen können.<sup>310</sup> Als Beispiel gibt er das Erkennen von Gesichtern an. Menschen erkennen eine ihnen bekannte Person unter tausenden Menschen. Sie können aber nicht genau beschreiben, an welchen Merkmalen sie diese Personen erkennen. Diese Feststellung von Polanyi und seine darauf aufbauende Begriffsprägung „tacit knowing“ führten zur zentralen Unterscheidung von implizitem und explizitem Wissen, die wichtige Schlussfolgerungen für den Umgang mit Wissen erlaubt.<sup>311</sup>

„Implizites Wissen ist an seinen Wissensträger gebundenes Wissen, das schwierig zu strukturieren und zu formulieren ist.“<sup>312</sup> Es ist in den Köpfen einzelner Personen gespeichert (embodied knowledge)<sup>313</sup> und besitzt somit einen subjektiven und intuitiven Charakter<sup>314</sup>. Implizites Wissen, z. B. subjektive Einsichten und Intuition, ist tief in den Handlungen von Individuen verankert, ebenso wie in ihren Idealen, Werten und Gefühlen.<sup>315</sup> Deshalb lässt sich implizites Wissen von einem Individuum zu einem anderen nur schwer transferieren. Viele Wissens Elemente kann der Wissensträger

---

<sup>307</sup> Vgl. Ackerschott (2001), S. 13.

<sup>308</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 249.

<sup>309</sup> Sie sind die ersten, die dieses Konzept im Rahmen der Organisationstheorie benutzen. Sie haben die kulturellen Unterschiede der japanischen und der westlichen Mentalität untersucht und beschreiben, dass implizites Wissen eher der japanischen und explizites Wissen eher der westlichen Mentalität entspricht. Nach ihnen wird Wissen in der westlichen Welt als explizit, also formal und systematisch gesehen. Explizites Wissen lässt sich in Worten und Zahlen ausdrücken und leicht mit Hilfe von Daten, wissenschaftlichen Formeln, festgelegten Verfahrensweisen oder universellen Prinzipien mitteilen. Demgegenüber herrscht in Japan ein völlig anderes Verständnis von Wissen. Für japanische Unternehmen stellt das in Worten und Zahlen fassbare Wissen nur die Spitze des Eisbergs dar. Sie sehen das Wissen hauptsächlich als etwas Implizites. Implizites Wissen ist sehr persönlich und schwer formulierbar, es lässt sich schwer mitteilen. Sie vertreten die Auffassung, dass ein Unternehmen nur Wissen generieren und Innovationen hervorbringen kann, wenn es implizites Wissen externalisiert, also kollektiv zugänglich macht.

<sup>310</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 249.

<sup>311</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 249.

<sup>312</sup> Ackerschott (2001), S. 19.

<sup>313</sup> Vgl. Rehäuser/Krcmar (1996), S. 7.

<sup>314</sup> Vgl. Hannig (2002), S. 15.

<sup>315</sup> Vgl. Rehäuser/Krcmar (1996), S. 7.

selbst beim besten Willen nicht weitergeben, da sie ihm nicht bewusst sind oder sie sich sprachlich nur schwer ausdrücken lassen (Ein Beispiel für implizites Wissen ist das Fahrradfahren. Obwohl die meisten Menschen Fahrrad fahren können, wird es vielen schwerfallen, ihre erworbene Fertigkeit anderen beizubringen). Daher nennt man implizites Wissen auch verborgenes oder stillschweigendes Wissen.<sup>316</sup> Diese Art von Wissen wird durch Praxis und jahrelange Erfahrung erworben.

Implizites Wissen lässt sich in die Dimensionen technisch und kognitiv unterteilen: Die technische Dimension umfasst informale und schwer dokumentierbare Fähigkeiten und Fertigkeiten (Know-how). Die kognitive Dimension setzt sich aus Schemata, mentalen Modellen, Überzeugungen und Wahrnehmungen, die der Mensch für selbstverständlich hält, zusammen und spiegelt seine Vorstellung von der Wirklichkeit (Was ist?) und seine Vision der Zukunft (Was sollte sein?) wider.<sup>317</sup>

Als explizit wird Wissen dann bezeichnet, wenn es dokumentiert und übertragbar ist. Diese Art des Wissens wird auch artikulierbares Wissen genannt, da es in formale Sprache gefasst werden kann und nicht an eine Person gebunden ist.<sup>318</sup> Es ist außerhalb der Köpfe einzelner Personen in Medien gespeichert (disembodied Knowledge) und kann so einfach mit Hilfe von Informations- und Kommunikationstechnologien verarbeitet, übertragen und gespeichert werden.<sup>319</sup> Die Beschaffenheit des impliziten Wissens erschwert dessen logische und systematische Verarbeitung, Übertragung und Speicherung in einer Organisation. Aus diesem Grund muss es in dokumentiertes, explizites Wissen überführt werden.<sup>320</sup>

Zur besseren Übersicht werden die wesentlichen Unterschiede zwischen implizitem und explizitem Wissen im Folgenden tabellarisch zusammengestellt (siehe Tab. 3).

Außer der vorangängigen Unterscheidung zwischen implizitem und explizitem Wissen ist eine weitere Kategorisierung von Wissen möglich. Ist Wissen an ein bestimmtes Individuum gebunden, spricht man von individuellem/privatem Wissen. Wenn es aber von mehreren Individuen geteilt wird, bezeichnet man es als kollektives/organisationales Wissen. Wenn implizites Wissen gleich in verschiedenen Köpfen vorhanden ist, so ist es organisational.<sup>321</sup> Ist explizites Wissen für andere nicht, sondern durch Verschluss oder Verschlüsselung nur für ein Individuum

---

<sup>316</sup> Vgl. Ackerschott (2001), S. 19 und Zahn/Foschiani/Tilebein (2000), S. 249.

<sup>317</sup> Vgl. Rehäuser/Krcmar (1996), S. 7.

<sup>318</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 249.

<sup>319</sup> Vgl. Rehäuser/Krcmar (1996), S. 7.

<sup>320</sup> Vgl. ebd., S. 7.

<sup>321</sup> Vgl. ebd., S. 8.

erreichbar, ist es personal.<sup>322</sup> Darüber hinaus macht es aus Unternehmenssicht Sinn, kollektives Wissen weiter danach zu differenzieren, ob es innerhalb oder außerhalb der Unternehmensgrenze angesiedelt ist.<sup>323</sup> So kann es z. B. Gruppen von Individuen innerhalb eines Unternehmens innewohnen (intraorganisationales Wissen) oder über Unternehmensgrenzen hinweg zwischen Organisationen existieren (interorganisationales Wissen).<sup>324</sup>

Implizites Wissen	Explizites Wissen
Stillschweigendes Wissen	Kodierbares Wissen
<ul style="list-style-type: none"> <li>• Einstellungen</li> <li>• Erfahrungen</li> <li>• Fertigkeiten</li> </ul>	Informationen, die in Büchern, Anweisungen, Formeln, Zeichnungen, Diagrammen, Plänen, Filmen, Auf Tonbändern usw. dokumentiert sind.
<ul style="list-style-type: none"> <li>• Können</li> <li>• Kompetenz</li> </ul>	Kenntnisse
Vermittelt durch Anwendungsbezug	Vermittelt durch Unterricht
Erworben durch Nachahmen, Üben und kontinuierliches Verbessern	Erworben durch Studieren

**Tabelle 3: Unterschiede zwischen implizitem und explizitem Wissen**  
 Quelle: Ackerschott (2001), S. 20.

Ausgehend von dem Konzept des impliziten und expliziten Wissens stellten Nonaka und Takeuchi ein dynamisches Modell vor, mit dem sie den Zusammenhang der Wissensgenerierung auf der individuellen und organisationalen Ebene erklären. Ihr Modell beruht auf der Annahme, dass menschliches Wissen durch eine soziale Interaktion von implizitem und explizitem Wissen geschaffen und erweitert wird.<sup>325</sup> Sie nennen diese Interaktion als ein sozialer Prozess zwischen Individuen „Knowledge Conversion“.<sup>326</sup> Ohne Individuen können Organisationen kein Wissen erzeugen<sup>327</sup>, deshalb stellt das implizite Wissen von Individuen die Basis für die Erzeugung des organisatorischen Wissens dar.<sup>328</sup> Organisatorische Wissenserzeugung ist als ein Prozess zu sehen, der das von Individuen erzeugte Wissen organisatorisch verstärkt

<sup>322</sup> Vgl. Rehäuser/Krcmar (1996), S. 8.

<sup>323</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 250.

<sup>324</sup> Vgl. Zahn/Foschiani/Tilebein (2000), S. 250.

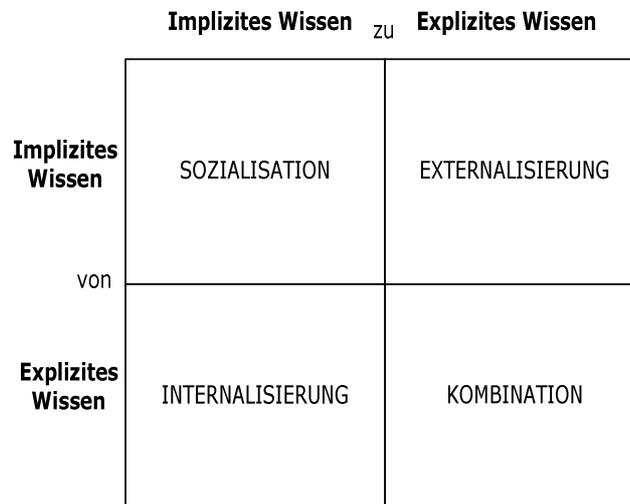
<sup>325</sup> Vgl. Nonaka/Takeuchi (1995), S. 61.

<sup>326</sup> Vgl. Nonaka/Takeuchi (1995), S. 61.

<sup>327</sup> Vgl. Rehäuser/Krcmar (1996), S. 35.

<sup>328</sup> Vgl. Rehäuser/Krcmar (1996), S. 37.

und in das Wissensnetzwerk der Organisation überführt.<sup>329</sup> Grundsätzlich wird in jeder Organisation auf individueller Ebene nach vier Formen Wissen erzeugt, die in dynamischer Beziehung zueinander stehen<sup>330</sup>:



**Abbildung 15: Formen der Wissensumwandlung**

Quelle: in Anlehnung an Nonaka/Takeuchi (1995), S. 72.

1. Von implizit zu implizit (*Sozialisation*):

Sozialisation bezeichnet den direkten Austausch von persönlichem Wissen zwischen zwei Personen. Als Schlüssel zum Erwerb impliziten Wissens dient hier Erfahrung, denn ohne gemeinsame Erfahrungen ist es sehr schwierig, sich in die Sichtweise anderer Personen zu versetzen und sich implizites Wissen anzueignen. Ein wichtiges Beispiel für Sozialisation ist für Nonaka und Takeuchi die Lehre.<sup>331</sup> Durch Beobachtung, Nachahmung und Praxis erwerben Lehrlinge implizites Wissen von ihren Meistern, wobei das Gelernte Teil des eigenen impliziten Wissens wird. Dieses erworbene Wissen wird als „sympathized knowledge“ bezeichnet. Das Wissen wird bei diesem Grundmuster nicht expliziert und steht deshalb nicht dem gesamten Unternehmen zur Verfügung.<sup>332</sup>

2. Von implizit zu explizit (*Externalisation*):

Erst in dieser Form wird neues, für das gesamte Unternehmen verwertbares Wissen erzeugt, da implizites Wissen durch Externalisation für alle zugänglich und dokumentierbar gemacht<sup>333</sup>, d. h. in explizites Wissen umgeformt wird. Nonaka

<sup>329</sup> Vgl. Rehäuser/Krcmar (1996), S. 35.

<sup>330</sup> Vgl. Rehäuser/Krcmar (1996), S. 35.

<sup>331</sup> Vgl. Nonaka/Takeuchi (1997), S. 75.

<sup>332</sup> Vgl. Rehäuser/Krcmar (1996), S. 36.

<sup>333</sup> Vgl. Rehäuser/Krcmar (1996), S. 36.

und Takeuchi nennen dadurch entstandenes, neues Wissen „conceptual knowledge“. Die Externalisation wird typischerweise in der Erstellung eines Konzeptes gesehen und wird durch Dialog oder kollektives Nachdenken ausgelöst.<sup>334</sup>

Innerhalb der vier Formen von Knowledge Conversion stellt die Externalisierung den Schlüssel zur Wissenserzeugung dar, weil sie aus implizitem Wissen explizite Konzepte generiert.<sup>335</sup> Dies geschieht innerhalb eines Prozesses mit Hilfe von Metaphern, Analogien und Modellen. Sie helfen schwer formulierbares implizites Wissen zu artikulieren. Der Prozess läuft kurzgefasst so ab<sup>336</sup>: Zuerst werden widersprüchliche Dinge und Aspekte oder Ideen durch Metaphern verknüpft. Daraufhin löst man diese Widersprüche durch Analogien auf. Abschließend werden die gewonnenen Konzepte herauskristallisiert und in einem Modell abgebildet, mit dem das erzeugte Wissen dem Rest des Unternehmens zugänglich wird.

### 3. Von explizit zu explizit (*Kombination*):

Kombination ist der Prozess der Erfassung von Konzepten innerhalb eines Wissenskomplexes und dient dazu, verschiedene Bereiche expliziten Wissens miteinander zu verbinden.<sup>337</sup> Somit kann neues, explizites Wissen durch die Zusammensetzung bereits bekannten expliziten Wissens erzeugt werden. Wissensaustausch und -kombination erfolgen über Medien wie Dokumente, Besprechungen, Telefon oder Computernetzwerke. Indem bestehende explizite Wissenssegmente neu sortiert, anders zusammengefügt und miteinander kombiniert werden, kann neues Wissen entstehen (z. B. Sehr große Datenmengen aus verschiedenen Bereichen des Unternehmens werden zur Analysezwecken in Data Marts abgelegt. Durch Anwendung von Data Mining-Methoden auf diese aufbereiteten Datenbestände können dann neue Beziehungen und Zusammenhänge zwischen einzelnen Daten identifiziert werden, die dem Menschen vorher nicht bekannt waren). Auf diese Art entstandenes Wissen wird als „systemic knowledge“ bezeichnet.

### 4. Von explizit zu implizit (*Internalisierung*):

Internalisierung stellt den Prozess dar, in dem das explizierte bzw. dokumentierte Wissen durch dessen Aufnahme, Ergänzung und Neuordnung seitens Mitarbeiter in

---

<sup>334</sup> Vgl. Rehäuser/Krcmar (1996), S. 36.

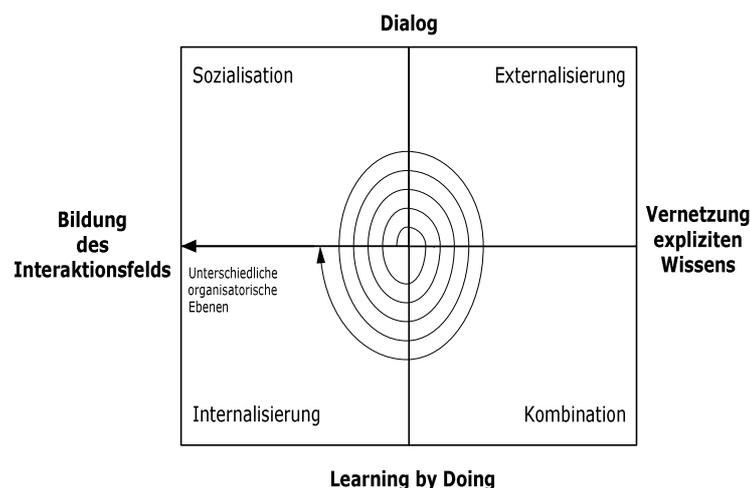
<sup>335</sup> Vgl. Nonaka/Takeuchi (1997), S. 79.

<sup>336</sup> Vgl. Rehäuser/Krcmar (1996), S. 37.

<sup>337</sup> Vgl. Nonaka/Takeuchi (1997), S. 81.

implizites Wissen eingegliedert wird.<sup>338</sup> Dies geschieht immer dann, wenn dokumentiertes Wissen in die täglichen Handlungen übernommen bzw. von den Mitarbeitern gelebt wird.<sup>339</sup> Nonaka und Takeuchi vergleicht diesen Prozess mit dem „Learning by Doing“.<sup>340</sup> Zur Förderung des Übergangs von explizitem zu implizitem Wissen kann Wissen in Dokumenten, Handbüchern oder mündlichen Geschichten festgehalten werden.<sup>341</sup> Die festgehaltenen und somit explizit zur Verfügung gestellten Erfahrungen führen dann durch wiederholtes Anwenden dieses Wissens in Form von Know-how und gemeinsamen mentalen Modellen zu einer Vertiefung der impliziten Wissensbasis von Mitarbeitern.<sup>342</sup> Durch Internalisierung gewonnenes, neues Wissen wird „operational knowledge“ genannt.

Diese vier Formen der Wissensumwandlung sind wechselseitig voneinander abhängig und führen in der Interaktion zu einem Spiralprozess, den Nonaka und Takeuchi die „Spirale des Wissens“ nennen.<sup>343</sup> Implizites und explizites Wissen wirken darin zusammen.<sup>344</sup>



**Abbildung 16: Die Spirale des Wissens**

Quelle: in Anlehnung an Nonaka/Takeuchi (1996), S. 71,  
in: Rehäuser/Krcmar (1996), S. 38.

Diese Interaktion entlang der Spirale wird durch verschiedene Trigger ausgelöst:<sup>345</sup> Der Prozess beginnt mit der Sozialisation, und diese wird durch die Bildung eines Interaktionsfeldes, welches das Teilen der Erfahrungen und mentalen Modelle der

<sup>338</sup> Vgl. Nonaka/Takeuchi (1997), S. 82 und Rehäuser/Krcmar (1996), S. 36.

<sup>339</sup> Vgl. Schreyögg/Geiger (2003), S. 8 f.

<sup>340</sup> Vgl. Nonaka/Takeuchi (1997), S. 82.

<sup>341</sup> Vgl. ebd., S. 83.

<sup>342</sup> Vgl. ebd., S. 83.

<sup>343</sup> Vgl. ebd., S. 84.

<sup>344</sup> Vgl. ebd., S. 85 f.

<sup>345</sup> Vgl. Rehäuser/Krcmar (1996), S. 37.

Unternehmensmitglieder ermöglicht, ausgelöst. Danach kann das durch Sozialisation geschaffene implizite Wissen externalisiert werden. Die Externalisation wird durch Dialog und kollektives Nachdenken angefeuert. Im Durchlauf des Prozesses durch die Spirale folgt dann die dritte Phase, nämlich die Kombination. Diese wird getriggert, wenn neu geschaffenes explizites mit bereits bestehendem explizitem Wissen aus verschiedenen Bereichen des Unternehmens vernetzt werden soll. Als Ergebnis dieser Form der Wissensumwandlung erhält man neue Produkte, Services oder Managementsysteme. Zum Schluss wird die Internalisierung durch learning by doing eröffnet. Die Internalisierung setzt die Spirale wiederum erneut in Gang, allerdings diesmal auf einer höheren ontologischen Ebene.<sup>346</sup>

Während des Durchlaufens dieser Spirale entstehen unterschiedliche Wissensinhalte als Resultat der jeweiligen Wissensumwandlung:<sup>347</sup>

- In der Phase der Sozialisation wird „erlebtes“, „sympathetisches“ Wissen wie gemeinsame, mentale Modelle und technische Fähigkeiten erzeugt.
- In der Phase der Externalisierung generiert „konzeptuelles“ Wissen (z. B. ein Produktkonzept)
- In der Phase Kombination wird „systemisches“ Wissen wie Prototypen, neue Methoden oder Komponenten hervorgebracht.
- In der Phase der Internalisierung wird „operatives“ Wissen wie über das Projektmanagement, den Produktionsprozess oder die Umsetzung eines Unternehmensprogramms generiert.

Die durch die jeweilige Art der Umwandlung geschaffenen Wissensinhalte in Form von sympathetischem, konzeptuellem, systemischem und operativem Wissen wirken in der Wissensspirale zusammen:

„Beispielsweise kann das sympathetische Wissen über Verbraucherbedürfnisse durch Sozialisation und Externalisierung zu konzeptuellem Wissen in Form eines Neuproduktkonzepts führen. Dieses konzeptuelle Wissen wird zur Leitlinie für die Schaffung systemischen Wissens durch Kombination. So kann etwa ein Neuproduktkonzept die Kombinationsphase lenken, in dem neuentwickelte und vorhandene Kenntnisse für den Bau eines Prototyps zusammengefügt werden. Systemisches Wissen wiederum – ein simulierter Produktionsprozess für das

---

<sup>346</sup> Vgl. Güldenbergl (2003), S. 237.

<sup>347</sup> Vgl. Rehäuser/Krcmar (1996), S. 37 und Nonaka/Takeuchi (1997), S. 85 f.

Neuprodukt – verwandelt sich durch Internalisierung in operatives Wissen über die Massenproduktion des betreffenden Produkts. Darüber hinaus löst das operative Erfahrungswissen oft einen neuen Kreislauf der Wissensschaffung aus. So kann das implizite operative Wissen eines Anwenders über ein Produkt durch Sozialisation eine Verbesserung oder Erneuerung des Produkts initiieren.“<sup>348</sup>

Durch die Interaktion von implizitem und explizitem Wissen während des fortdauernden Durchlaufens der Spirale des Wissens wird Wissen in seinem Ausmaß immer größer.<sup>349</sup> „Die Wissensschaffung im Unternehmen ist somit ein Spiral-Prozess, der ausgehend von der individuellen Ebene immer mehr Interaktionsgemeinschaften erfasst und die Grenzen von Sektionen, Abteilungen, Divisionen und sogar das Unternehmen überschreitet.“<sup>350</sup>

Zusammenfassend kann man sagen, dass das Wissen erst vom Individuum erzeugt wird. Dieses Wissen wird dann durch dessen Umwandlung im Unternehmen verstärkt und unternehmensweit zur Verfügung gestellt. In diesem Prozess dringt es in immer höhere ontologische Schichten vor<sup>351</sup>, also von Individuumsebene auf Gruppenebene, von Gruppenebene zur Unternehmensebene. Vor diesem Hintergrund müssen Unternehmen individuelles implizites Wissen mobilisieren und für seine Verteilung innerhalb des Unternehmens Möglichkeiten bieten.<sup>352</sup>

## **4.3 WAS IST WISSENSMANAGEMENT?**

In den Kapiteln 4.1 und 4. 2 wurden die Begriffe Zeichen, Daten, Informationen und Wissen voneinander abgegrenzt. Danach ist zwischen sowie zwischen implizitem und explizitem sowie zwischen individuellem und organisationalem Wissen unterschieden worden. Abschließend wurde der Wissensgenerierungsprozess anhand des von Nonaka und Takeuchi entwickelten Spiralmodells erläutert.

All diese Elemente bilden die theoretischen Grundlagen für das Wissensmanagement.

### **4.3.1 Definitionen**

Die im Kapitel 4.2.1 erwähnten Abgrenzungs- und Definitionsschwierigkeiten des Wissensbegriffs ziehen sich weiter zum Wissensmanagement-Begriff, denn auch hier ist keine einheitliche Definition vorhanden.

---

<sup>348</sup> Nonaka/Takeuchi (1997), S. 86.

<sup>349</sup> Vgl. Rehäuser/Krcmar (1996), S. 37.

<sup>350</sup> Nonaka/Takeuchi (1997), S. 86 und (1995), S. 72.

<sup>351</sup> Vgl. Nonaka/Takeuchi (1997), S. 86.

<sup>352</sup> Vgl. Nonaka/Takeuchi (1997), S. 86.

In der Literatur kann man zwar verschiedene Begriffsformulierungen für das Wissensmanagement finden, aber da Unternehmen in ihren organisatorischen Strukturen, Zielen, Philosophien und Kulturen ganz unterschiedlich und damit einzigartig sind, müssen sich Organisationen, die Wissensmanagement betreiben möchten, selbst auf die Suche nach der für sie gültigen bzw. passenden Begriffsbestimmung für Wissensmanagement machen.

Als Orientierungshilfe werden hier einige Definitionen unterschiedlicher Autoren genannt:

Probst, Raub und Romhardt gehen bei Ihrer Definition des Begriffs Wissensmanagement von der organisationalen Wissensbasis aus. Für sie bildet Wissensmanagement ein Interventionskonzept, das sich mit den Möglichkeiten zur Gestaltung der organisationalen Wissensbasis befasst.<sup>353</sup> Die Interventionen liegen sowohl auf der individuellen (Z. B. Maßnahmen des Personalmanagements) als auch auf der kollektiven bzw. organisationalen (Z. B. Unternehmensentwicklung, strategische Planung oder EDV-Organisation) Ebene.

Nach Wilke ist Wissensmanagement als Element eines Zusammenhanges gesellschaftlicher, organisationaler, technologischer und individueller Faktoren zu betrachten.<sup>354</sup> Er definiert Wissensmanagement als die Gesamtheit organisationaler Strategien zur Schaffung einer intelligenten Organisation.<sup>355</sup> Hier geht es mit Blick auf die Personen um das organisationsweite Niveau der Kompetenzen, Ausbildung und Lernfähigkeit der Mitglieder. Bezüglich der Organisation geht es um die Schaffung, Nutzung und Entwicklung einer kollektiven Intelligenz oder Expertise; hinsichtlich der technologischen Infrastruktur um die Schaffung und effiziente Nutzung der zur Organisation passenden Kommunikations- und Informationsinfrastruktur.<sup>356</sup> Informationstechnologie stellt immer die Basis dar, jedoch soll das Management nicht in die Falle geraten, sich ausschließlich darauf zu konzentrieren.<sup>357</sup>

Davenport und Prusak schlagen folgende Definition für Wissensmanagement vor:

„Wissensmanagement ist eine formale, strukturierte Initiative, um die Erzeugung, Verteilung und Nutzung von Wissen in einem Unternehmen zu verbessern. Es ist ein formaler Prozess, in dem das Wissen des Unternehmens den Wert des Unternehmens

---

<sup>353</sup> Vgl. Probst/Raub/Romhardt (1999), S. 57.

<sup>354</sup> Vgl. Wilke (2001), S. 6.

<sup>355</sup> Vgl. ebd., S. 39.

<sup>356</sup> Vgl. ebd., S. 5 f und 39.

<sup>357</sup> Vgl. Eschenbach/Geyer (2004), S. 148.

vergrößert.“<sup>358</sup>

Vergleichend zu den oben eingeführten Definitionen kann man grundsätzlich sagen, dass unter dem Begriff Wissensmanagement ein systematischer Ansatz verstanden wird, der die Gesamtheit der Managementaufgaben zur Gestaltung eines effizienten und effektiven Wissensflusses umfasst. Dies beinhaltet die Planung, Steuerung und Kontrolle der Prozesse zur Generierung der Ressource Wissen als Ergebnis der Transformation von Daten über Informationen aus unternehmensinternen und – externen Quellen, zielgerechten Nutzung der Ressource Wissen im Rahmen der Geschäftsabläufe und Verteilung der Ressource Wissen inner- und außerhalb des Unternehmens.<sup>359</sup> Dabei kann je nach Anliegen des Unternehmens eine unterschiedliche Orientierung des Ansatzes vorgezogen werden (d. h. technik- oder humanzentriertes Wissensmanagement).<sup>360</sup> Im folgenden Kapitel werden die möglichen Ausrichtungen des Wissensmanagements beschrieben.

### **4.3.2 Gestaltungsdimensionen des Wissensmanagements**

Eine erste Orientierung im breiten Feld der Wissensmanagementkonzepte bietet die Unterscheidung in eine human- und eine technikorienteerte Perspektive.<sup>361</sup>

„Die humanorientierte Sichtweise versteht Wissen als einen Prozess, der dynamisch, kontextgebunden und personenabhängig ist.“<sup>362</sup> Diese Sichtweise des Wissensmanagements stellt das Individuum - als Wissensträger - sowie die Kommunikation zwischen Organisationsmitgliedern und die Gestaltung und Nutzung der Interaktionsprozesse in einer Organisation in den Mittelpunkt,<sup>363</sup> vernachlässigt jedoch die Bedeutung und Notwendigkeit einer geeigneten technologischen Infrastruktur für ein erfolgreiches Wissensmanagement.<sup>364</sup>

Im Gegensatz dazu sieht die technikorienteerte Sichtweise Wissen als teilbares Objekt.<sup>365</sup> So kann Wissen auf einer instrumentell-technischen Ebene als Abbildung der Realität personen- und kontextunabhängig verarbeitet werden.<sup>366</sup> Damit stehen die Informations- und Kommunikationssysteme zur Identifikation, Bearbeitung,

---

<sup>358</sup> Davenport/Prusak (1998), S. 16.

<sup>359</sup> Vgl. Dittmar/Gluchowski (2002), S. 28 f.

<sup>360</sup> Vgl. Mühlethaler (2005), S. 22.

<sup>361</sup> Vgl. Dittmar/Gluchowski (2002), S. 29, Bullinger/Wagner/Ohlhausen (2000), S. 79 und Schüppel (1996), S. 187.

<sup>362</sup> Gaßen (1999), S. 12.

<sup>363</sup> Vgl. Dittmar/Gluchowski (2002), S. 29.

<sup>364</sup> Vgl. Gaßen (1999), S.12.

<sup>365</sup> Vgl. Dittmar/Gluchowski (2002), S. 29 und Gaßen (1999), S. 11.

<sup>366</sup> Vgl. Dittmar/Gluchowski (2002), S. 29.

Erweiterung, Speicherung, Bewertung und Verteilung von Wissen im Mittelpunkt des Wissensmanagements.<sup>367</sup> Der Nachteil dieser technikorientierten Auslegung des Wissensmanagements liegt darin, dass der Einfluss der organisatorischen Regelungen, der Bereitschaft und der Fähigkeiten von Mitarbeitern sowie der Normen und Werte im Unternehmen nicht berücksichtigt wird.<sup>368</sup>

Ausgehend von beiden Sichtweisen und deren Ausrichtungen lässt sich i. d. R. feststellen, dass sich der humanorientierte Ansatz eher mit dem Management des impliziten Wissens beschäftigt, d. h. der Internalisierung von explizitem und der Sozialisation und Externalisierung von implizitem Wissen, während der technologische Ansatz sich primär auf das explizite Wissen bezieht. Hier wird explizites Wissen kombiniert (z. B. durch dessen Integration mit bereits bestehendem explizitem Wissen).

Die ausschließliche Festlegung auf nur einen der Ansätze vernachlässigt jedoch die Wechselwirkungen, die sich im Gesamtbild eines Unternehmens im Zusammenhang mit Wissensmanagement ergeben.<sup>369</sup> Daher müssen beide Ansätze gleichmaßen und unter Berücksichtigung organisatorischer Aspekte betrachtet werden.

So ergeben sich drei Gestaltungsdimensionen für Maßnahmen zur erfolgreichen Einführung von Wissensmanagement in Unternehmen:<sup>370</sup>

- *Human Resource Management (Mensch)*
- *Organisation (Aufbau- und Ablauforganisation)*
- *IuK-Technologien (Technik)*

Diese drei Dimensionen müssen in einer ganzheitlichen Sichtweise integriert werden (siehe dafür Abb. 17).

Denn das Potential des Wissensmanagements kann nur dann optimal ausgeschöpft werden, wenn ein ganzheitlicher Ansatz verfolgt wird.<sup>371</sup> Damit wird nicht nur die informationstechnische Dimension in Form der Bereitstellung geeigneter technischer Infrastrukturen betrachtet, sondern es werden auch organisatorische Aspekte (Unternehmensstrukturen, Prozesse usw.) sowie die diversen Humanressourcen

---

<sup>367</sup> Vgl. Dittmar/Gluchowski (2002), S. 29.

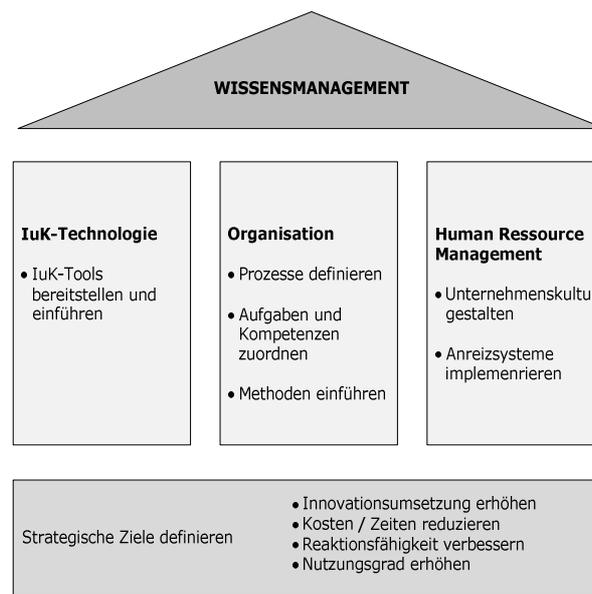
<sup>368</sup> Vgl. Gaßen (1999), S. 12.

<sup>369</sup> Vgl. Bullinger/Wagner/Ohlhausen (2000), S. 79.

<sup>370</sup> Vgl. Dittmar/Gluchowski (2002), S. 29.

<sup>371</sup> Vgl. Bullinger/Wagner/Ohlhausen (2000), S. 79.

(Unternehmenskultur, Motivation usw.) in den Gestaltungsrahmen mit einbezogen.



**Abbildung 17: Gestaltungsdimensionen eines ganzheitlichen Wissensmanagements**

Quelle: in Anlehnung an Bullinger/Wagner/Ohlhausen (2000), S. 80.

- *Human Ressource Management:*

Ein zielorientiertes Human Resource Management zur Gestaltung einer adäquaten Unternehmenskultur, die einen kontinuierlichen Wissenstransfer unterstützt, ist ein wesentlicher Erfolgsfaktor.<sup>372</sup> Der gesamte Prozess des Wissensmanagements ist in hohem Maße von der frühzeitigen Einbindung und der Bereitschaft der Mitarbeiter abhängig.<sup>373</sup> Im Mittelpunkt steht also der Mensch als Wissensträger. Er erwirbt sein Wissen in Lernprozessen und durch Interaktionen mit anderen Organisationsmitgliedern. Dabei stellt die Verwendung von Standards ein wichtiges Instrument dar, das die Kodifizierung von individuellem (Spezialisten-) Wissen erleichtert und den persönlichen Lernprozess positiv beeinflusst.<sup>374</sup> Die Berücksichtigung gewisser Standards im Umgang mit Informationen hilft dem Menschen bei der Wahrnehmung und Strukturierung der auf ihn einwirkenden Informationsflut.<sup>375</sup>

- *Organisation:*

Das Wissen der Mitarbeiter in der Gesamtheit stellt die organisationale Wissensbasis

<sup>372</sup> Vgl. Bullinger/Wagner/Ohlhausen (2000), S. 79.

<sup>373</sup> Vgl. ebd., S. 79.

<sup>374</sup> Vgl. Dittmar/Gluchowski (2002), S. 30.

<sup>375</sup> Im Zusammenhang dazu stehen Visualisierungstechniken (z. B. Mind Mapping), die helfen, individuelles Wissen zu strukturieren, vgl. Dittmar/Gluchowski (2002), S. 30.

eines Unternehmens dar.<sup>376</sup> Im Rahmen der Gestaltungsdimension Organisation sollen die geeigneten Maßnahmen eingeleitet werden, um dem Wissensmanagementkonzept gerecht zu werden. So realisieren Unternehmen meist Wissenslandkarten (Knowledge Maps) bzw. Wissensbranchenbüchern zur Umsetzung von Wissensmanagement. Derartige Abbildungen des Wissens bieten Hilfestellung bei der Identifikation von unternehmensinternem Wissen und verbessern somit die betriebliche Wissenstransparenz.<sup>377</sup> Zudem stehen sogenannte Lessons Learned- bzw. Best Practices-Berichten in Form von papiergebundenen oder elektronischen Dokumenten zur Verfügung, um die Gefahr der Wiederholung von Fehlern zu reduzieren.

Durch verbesserte Kommunikationsmöglichkeiten und -maßnahmen (z. B. Einrichtung von Kommunikationsforen, Förderung von sog. Communities of Practice) werden Mitarbeiter dazu gefördert, ihr Wissen mitzuteilen und unternehmensweit zu verbreiten. In diesem Zusammenhang kommt der Entwicklung sowohl materieller als auch immaterieller Anreizsysteme besondere Bedeutung zu. Der Einsatz solcher Anreizsysteme bringt jedoch alleine nichts, wenn die Unternehmenskultur keine entsprechende Unterstützung bietet, d. h., wenn sie nicht durch Offenheit und Ehrlichkeit sowie durch Vertrauen – Vertrauen in die Mitarbeiter durch Vorgesetzte und unter den Kollegen – geprägt ist.<sup>378</sup> Nur wenn die Mitarbeiter Vertrauen in ihr Unternehmen haben, sind sie bereit, ihr Wissen weiterzutragen. Außerdem werden den Mitarbeitern konkrete Wissensrollen bzw. -funktionen zugeordnet, um die erfolgreiche Umsetzung von Wissensmanagement im Unternehmen gewährleisten zu können. Der sogenannte Chief Knowledge Officer übernimmt somit die Aufgabe, die organisationale Wissensbasis zu gestalten, lenken und entwickeln.<sup>379</sup>

- *Informations- und Kommunikationstechnologien:*

Die dritte Gestaltungsdimension des Wissensmanagements ist die Informations- und Kommunikationstechnologien. IuK-Technologien als *technologischer Enabler* unterstützen die organisationale Dimension und ermöglichen die Sammlung, Organisation, Nutzung und Verteilung von Wissen zwischen Mitarbeitern. Sie dienen als Organisator von explizierbarem Wissen, so dass den Mitarbeitern Wissen zur richtigen Zeit am richtigen Ort in der erforderlichen Qualität zur Verfügung steht.<sup>380</sup>

Internet- und Intranet-Technologien, Computer Supported Cooperative Work (CSCW)-

---

<sup>376</sup> Vgl. Bullinger/Wagner/Ohlhausen (2000), S. 80.

<sup>377</sup> Vgl. Dittmar/Gluchowski (2002), S. 30.

<sup>378</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 8.

<sup>379</sup> Vgl. Dittmar/Gluchowski (2002), S. 31.

<sup>380</sup> Vgl. ebd., S. 31.

Lösungen sowie Dokumenten-Managementsysteme gehören zu den Basissystemen von Wissensmanagement.<sup>381</sup> Diese bestehenden Systemelemente werden dann durch die Integration von innovativen Ansätzen erweitert und ausgebaut. Dazu zählen z. B. Information Retrieval-Verfahren, Push-/Pull-Technologien, Intelligente Agenten, Data Mining-Systeme sowie Ansätze aus dem Bereich der künstlichen Intelligenz.<sup>382</sup> Damit kommt bei dem Wissensmanagement eine Vielzahl von Werkzeugen zum Einsatz.<sup>383</sup>

In der Phase der Wissensgenerierung werden z. B. Meeting- bzw. Group Decision Support-Werkzeuge, On-Line Analytical Processing (OLAP)-Systeme und Data/Text Mining-Systeme verwendet. Zur Wissenskodifizierung werden Speichersysteme, wie klassische Datenbank und Dokumenten-Managementsystem aber auch Data Warehouse-Lösungen eingesetzt. Indexing Engines, Information Retrieval-Systeme oder IT-gestützte Knowledge Maps unterstützen den Anwender bei der Wissenssuche. Die Wissensverteilung innerhalb des Unternehmens wird insbesondere durch CSCW-Systeme und Internet-/Intranet-Technologien ermöglicht.

### **4.3.3 Bausteine des Wissensmanagements**

Mit dem Konzept der „Bausteine des Wissensmanagements“ liefern Probst, Raub und Romhardt einen pragmatischen Wissensmanagement-Ansatz, welcher Führungskräften Methoden aufweist, mit denen sie organisationale Wissensbestände lenken und in ihrer Entwicklung beeinflussen können.<sup>384</sup>

Dieser Ansatz gliedert wissensbezogene Management-Aktivitäten in acht sogenannte Bausteine, die in einem Kreislauf (besteht aus einem inneren und äußeren Teil) nacheinander folgen. Die Bausteine Wissensidentifikation, Wissenserwerb, Wissensentwicklung, Wissensverteilung, Wissensnutzung und Wissensbewahrung werden als Kernprozesse bezeichnet<sup>385</sup>, welche den inneren Kreislauf der folgenden Abbildung 18 ausmachen. Sie bilden die operativen Probleme ab, die im Umgang mit Wissen auftreten.<sup>386</sup> Diese Kernprozesse mit Ausrichtung auf operative Aspekte des Wissensmanagements benötigen noch einen „orientierenden und koordinierenden Rahmen“<sup>387</sup>, welcher die Wichtigkeit strategischer Aspekte und zielgerichteter

---

<sup>381</sup> Vgl. Dittmar/Gluchowski (2002), S. 31.

<sup>382</sup> Vgl. ebd., S. 31.

<sup>383</sup> Vgl. ebd., S. 31. Diese Werkzeuge werden meist nicht primär für das Wissensmanagement angeboten, sind aber im Hinblick auf die Ansprüche des Wissensmanagements geeignet, den bewussten Umgang mit der Ressource „Wissen“ zu ermöglichen, vgl. ebd., S. 31.

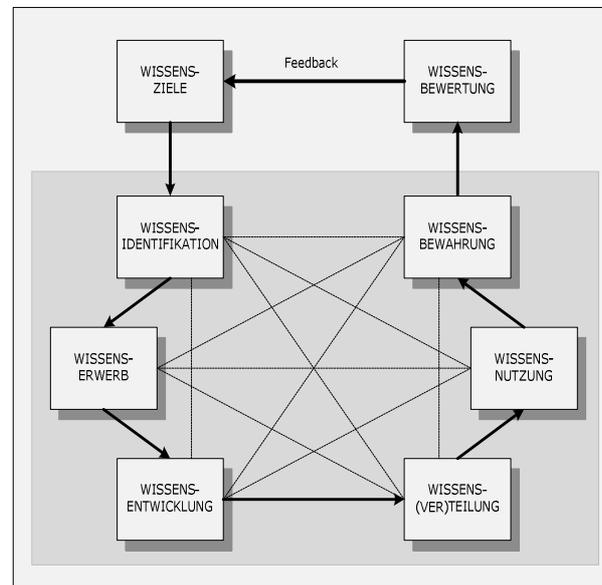
<sup>384</sup> Vgl. Probst/Raub/Romhardt (1999), S. 51.

<sup>385</sup> Vgl. ebd., S. 53.

<sup>386</sup> Vgl. ebd., S. 56.

<sup>387</sup> Vgl. ebd., S. 56.

Steuerung im Wissensmanagement betont<sup>388</sup>. Um diesen strategischen Bezug in das Konzept zu integrieren, werden die sechs Kernprozesse durch die Bausteine Wissensziele und Wissensbewertung ergänzt.<sup>389</sup>



**Abbildung 18: Bausteine des Wissensmanagements**

Quelle: in Anlehnung Probst/Raub/Raumhardt (1999), S. 58.

Dadurch wird das Konzept zu einem Managementregelkreis, der mit seinen insgesamt acht Bausteinen strategische und operative Elemente integriert.<sup>390</sup>

Den Ausgangspunkt dieses Regelkreises bildet die Definition von Wissenszielen, um die strategische Ausrichtung des Wissensmanagements sicherzustellen und die Zielsetzungen für Interventionen abzuleiten. Der Kreislauf wird durch die Wissensbewertung geschlossen. Dieser kommt die Controlling-Aufgabe zu, um den Wissensmanagement-Prozess zielgerichtet steuern zu können. Alle Wissensmanagement-Bausteine stehen in enger Verbindung zueinander, so dass Eingriffe auf einen Bereich bzw. Baustein zwingende Konsequenzen für die anderen Bereiche/Bausteine hervorbringen.<sup>391</sup> Im Folgenden werden die einzelnen Bausteine kurz vorgestellt:

Am Beginn des Wissensmanagement-Prozesses steht der Baustein *Wissensziele*. Probst, Raub und Romhardt sehen die Definition von Wissenszielen als grundlegende Voraussetzung für die Umsetzung des Wissensmanagements.<sup>392</sup> Damit das

<sup>388</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 9.

<sup>389</sup> Vgl. Probst/Raub/Romhardt (1999), S. 56.

<sup>390</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 9 und Probst/Raub/Romhardt (1999), S. 56 f., 59.

<sup>391</sup> Vgl. Probst/Raub/Romhardt (1999), S. 53 und 60.

<sup>392</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 11.

Wissensmanagement-Projekt nicht aus der Bahn gerät, müssen Organisationen überlegen, welche Fähigkeiten und Kompetenzen im Unternehmen gebraucht werden. „Wissensziele geben den Aktivitäten des Wissensmanagements eine Richtung vor.“<sup>393</sup> Organisationen legen damit fest, auf welchen Ebenen welche Fähigkeiten und Kompetenzen für die Zukunft aufgebaut werden sollen und bestimmen somit die Entwicklungsrichtung des Unternehmens.

Nach den Autoren gibt es drei Zielebenen zu unterscheiden:<sup>394</sup>

Normative Wissensziele schaffen das kulturelle Umfeld für Wissensmanagement, erzeugen ein entsprechendes Bewusstsein und schaffen so die Voraussetzungen für ein effektives Wissensmanagement.<sup>395</sup> Die strategischen Wissensziele definieren organisationales Kernwissen, das in der Zukunft erworben oder bewahrt werden soll und beschreiben somit den zukünftigen Kompetenzbedarf eines Unternehmens.<sup>396</sup> Die dritte Ebene der operativen Wissensziele sichert die notwendige Konkretisierung der normativen und strategischen Zielvorgaben und sorgt für die konkrete Umsetzung des Wissensmanagements.<sup>397</sup> Als klar formulierte, konkrete Handlungsanweisungen richten sie sich direkt an Abteilungen oder einzelne Beschäftigte und machen die normativen und strategischen Ziele in der Praxis umsetzbar.<sup>398</sup>

Nachdem es festgestellt worden ist, welche Wissensziele das Unternehmen hat, und bevor weitere Schritte unternommen werden, müssen bereits vorhandene Wissensbestände (intern oder extern) ermittelt und Wissenslücken bzw. -defizite im Unternehmen festgestellt werden. Das erfolgt durch die *Wissensidentifikation*. „Mangelnde Transparenz bei der Identifikation führt zu Ineffizienzen, uninformierten Entscheidungen und Redundanzen.“<sup>399</sup> Ein effektives Wissensmanagement muss daher ein hinreichendes Maß an interner und externer Transparenz schaffen und die Mitarbeiter bei ihren Suchaktivitäten unterstützen.<sup>400</sup>

Nach der Wissensidentifikation, in der es festgestellt worden ist, über welches Wissen das Unternehmen verfügt und wo neues Wissen aufgebaut werden soll, muss nun versucht werden, dieses erforderliche neue Wissen schnell für das Unternehmen verfügbar zu machen. In dieser Hinsicht können Unternehmen kritische Kompetenzen

---

<sup>393</sup> Bullinger/Warschat/Prieto/Wörner (1998), S. 11.

<sup>394</sup> Vgl. Probst/Raub/Romhardt (1999), S. 71-88.

<sup>395</sup> Vgl. Probst/Raub/Romhardt (1999), S. 72.

<sup>396</sup> Vgl. Probst/Raub/Romhardt (1999), S. 78.

<sup>397</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 11.

<sup>398</sup> Vgl. Probst/Raub/Romhardt (1999), S. 87.

<sup>399</sup> Bullinger/Warschat/Prieto/Wörner (1998), S. 12.

<sup>400</sup> Vgl. Bullinger/Warschat/Prieto/Wörner (1998), S. 12.

extern auf verschiedenen Wissensmärkten erwerben<sup>401</sup>, wenn deren interne Beschaffung ihnen aus irgendeinem Grund nicht möglich zu sein scheint oder zuviel Zeit kostet.

Es gibt eine Reihe von Möglichkeiten zum externen Erwerb von Wissen:<sup>402</sup>

Erwerb von Kompetenzen externer Wissensträger durch langfristige oder temporäre Anstellung, Anmietung von Personal oder Consulting durch Experten, Erwerb von Stakeholder-Wissen z. B. über das Feedback von Kunden und Lieferanten, die wertvolle Beiträge zur Verbesserung der Kompetenzen liefern (So können z. B. Schlüsselkunden in Entwicklungsprozesse involviert werden, oder es können mit ihnen Kooperationen in Pilotprojekten eingegangen werden, wodurch das frühzeitige Lernen und Berücksichtigen der Kundenbedürfnisse ermöglicht wird), Erwerb von Wissen anderer Unternehmen durch Übernahme innovativer Firmen oder Eingehen von „product links“ zur gezielten Schließung von Wissenslücken, Erwerb von Wissensprodukten, z. B. Ankauf von Datenbanken, Software, Patenten, CD-Roms und anderen „Wissenskonserven“.

Der Baustein der *Wissensentwicklung* ergänzt das Aufgabenfeld des *Wissenserwerbs*.<sup>403</sup> Hier steht die Erzeugung neuer Fähigkeiten, neuer Produkte, innovativer Ideen und besserer Prozesse im Mittelpunkt. Wissensentwicklung ist also der bewusste Aufbau neuen Wissens im Unternehmen,<sup>404</sup> wodurch nachhaltige Wettbewerbsvorteile erzielt werden können<sup>405</sup>. Daher soll das Management die Organisationsmitglieder dazu bewegen, bisher intern noch nicht bestehende oder intern und extern noch nicht existierende Fähigkeiten zu entwickeln. Die Voraussetzung dafür ist, dass die Unternehmenskultur ein geeignetes, positives Umfeld schafft, das die Mitarbeiter im Prozess der Wissensentwicklung unterstützt.<sup>406</sup>

Probst, Raub und Romhardt unterscheiden drei Arten der Wissensentwicklung:

Individuelle Wissensentwicklung einzelner Beschäftigter, kollektive Wissensentwicklung durch gezielte Bildung von kreativen Teams und Wissensentwicklung durch Forschungsk Kooperationen, in denen Wissen von einem

---

<sup>401</sup> Vgl. Probst/Raub/Romhardt (1999), S. 174.

<sup>402</sup> Vgl. Probst/Raub/Romhardt (1999), S. 163-173.

<sup>403</sup> Vgl. Probst/Raub/Romhardt (1999), S. 54. Der Unterschied zwischen Wissenserwerb und -entwicklung liegt in der Herkunft des erzeugten neuen Wissens: Während beim Wissenserwerb das für das Unternehmen notwendige Wissen (Know-how) von außen beschafft wird, lässt die Wissensentwicklung Unternehmen dieses Wissen selbst, also unternehmensintern erzeugen.

<sup>404</sup> Vgl. Probst/Raub/Romhardt (1999), S. 218.

<sup>405</sup> Vgl. ebd., S. 179.

<sup>406</sup> Vgl. ebd., S. 187 f.

externen Partner speziell für ein Unternehmen entwickelt wird und diesem dann exklusiv zur Verfügung steht.

Sobald das nötige Wissen vorhanden ist, wird es im nächsten Schritt innerhalb des Unternehmens verbreitet, d. h. auf die richtigen Beschäftigten verteilt bzw. an die richtigen Stellen gebracht. Somit wird sichergestellt, dass das richtige Wissen zur richtigen Zeit am richtigen Ort zur Verfügung steht. In der *Wissens(ver)teilung* wird isoliert vorhandenes Wissen (auch in Form von Kenntnissen und Erfahrungen) für das gesamte Unternehmen nutzbar gemacht.<sup>407</sup> Die Steigerung der Effizienz der Beschäftigten ist das wesentliche Ziel der Wissens(ver)teilung.<sup>408</sup> Hier wird allerdings vorausgesetzt, dass die Beschäftigten bereit sind, ihr Wissen zu (ver)teilen.<sup>409</sup>

Zudem sind das Teilen und das Verteilen von Wissen nicht identisch.<sup>410</sup> So bedeutet Wissensverteilung die Bereitstellung von Informationen, Daten und Wissen über Informations- und Kommunikationssysteme. Wissensteilung beschreibt hingegen die Weitergabe von Wissen von einem Menschen zum anderen.<sup>411</sup>

Die Identifikation und (Ver)Teilung des Wissens schaffen zwar die Grundlage für das Wissensmanagement, reichen jedoch alleine nicht für ein effektives Wissensmanagement. Denn erst die produktive Anwendung des vorhandenen, erworbenen oder selbst entwickelten Wissens im Unternehmen macht die Folgen eines effektiven Wissensmanagements sichtbar.<sup>412</sup> Die *Wissensnutzung*, also der produktive Einsatz organisationalen Wissens zum Nutzen des Unternehmens ist das Ziel des Wissensmanagements,<sup>413</sup> demzufolge auch in anderen Bausteinen des Wissensmanagements die Nutzung des Wissens als Ziel berücksichtigt wird.<sup>414</sup>

Mitarbeiter, die das Unternehmen verlassen, Outsourcing, das Auflösen von erfolgreichen Teams sowie ein Mangel an Kommunikation führen zum Verlust des wertvollen Wissens im Unternehmen. Die *Wissensbewahrung* schützt das Unternehmen vor diesem Wissensverlust. Sie sorgt durch gezielte Speicherung von Wissen dafür, dass dieses auch in Zukunft zur betrieblichen Nutzung zur Verfügung

---

<sup>407</sup> Vgl. Probst/Raub/Romhardt (1999), S. 224.

<sup>408</sup> Vgl. ebd., S. 229.

<sup>409</sup> Vgl. ebd., S. 258,

<sup>410</sup> Vgl. ebd., S. 224 f.

<sup>411</sup> Wissensteilung ist eine kulturelle Aufgabe, vgl. ebd., S. 257 und wird durch moderne IuK-Systeme unterstützt, vgl. ebd., S. 243. Der befürchtete Machtverlust des Wissen Abgebenden stellt allerdings ein soziales Problem der Bereitschaft zur Wissensweitergabe dar.

<sup>412</sup> Vgl. Probst/Raub/Romhardt (1999), S. 271 f.

<sup>413</sup> Vgl. ebd., S. 55.

<sup>414</sup> Vgl. ebd., S. 272 und 287.

steht.<sup>415</sup>

Informationen, Fähigkeiten und Erfahrungen werden z. B. in digitaler Form auf Speichermedien abgelegt und zur Wiederverwendung bereitgestellt. Durch den Prozess der Wissensbewahrung kann sich das Unternehmen ein „organisationales Gedächtnis“ schaffen, welches auch die Grundlage für organisationales Lernen bildet.<sup>416</sup> Die drei Hauptprozesse der Wissensbewahrung sind die Selektion, Speicherung und Aktualisierung vorhandener Wissensbestände.<sup>417</sup>

Durch die *Wissensbewertung* wird schließlich der Regelkreis des Wissensmanagements geschlossen. In diesem Baustein werden die Ergebnisse des Wissensmanagements den normativen, strategischen und operativen Wissenszielen gegenübergestellt.<sup>418</sup> So wird geprüft, ob Wissensziele zweckmäßig formuliert und Wissensmanagementmaßnahmen erfolgreich durchgeführt worden sind.<sup>419</sup> Die Ergebnisse der Wissensbewertung fließen als Feedback in den erneuten Zielbildungsprozess ein und veranlassen entsprechende Interventionen in den Bausteinen des Wissensmanagements, wodurch eine systematische und fortlaufende Verbesserung des Wissensmanagementprozesses ermöglicht wird.<sup>420</sup>

## **4.4 WISSENSMANAGEMENT IM CRM**

Viele Unternehmen haben schon erkannt, dass das intellektuelle Vermögen der Kunden wertvoll ist und gezielt ausgeschöpft werden muss.

Eng verknüpft mit dem Konzept des Informationsmanagements<sup>421</sup> und definiert als „[...] a discipline that seeks to improve the performance of individuals and organizations by maintaining and leveraging the present and future value of knowledge assets“<sup>422</sup> wird beim Wissensmanagement häufig diese systematische Ausschöpfung des Kundenwissens in den Vordergrund gestellt.<sup>423</sup> Denn eine Steigerung des Wissens über die Kunden und somit ein erhöhtes Verständnis für deren Bedürfnisse bietet viele Möglichkeiten zur Verbesserung der Kundenbeziehungen, was

---

<sup>415</sup> Vgl. Probst/Raub/Romhardt (1999), S. 289 f.

<sup>416</sup> Vgl. ebd., S. 292.

<sup>417</sup> Vgl. ebd., S. 295.

<sup>418</sup> Vgl. ebd., S. 57.

<sup>419</sup> Vgl. ebd., S. 350.

<sup>420</sup> Vgl. ebd., S. 58.

<sup>421</sup> Vgl. Sexauer (2004), S. 22.

<sup>422</sup> Newman, Conrad (1999), S. 1.

<sup>423</sup> Vgl. Sexauer (2004), S. 22.

letztlich zur Steigerung des Unternehmenswerts führt.<sup>424</sup>

Beide Konzepte - CRM und Wissensmanagement - zielen darauf ab, die knappen Ressourcen von Unternehmen gezielter einzusetzen, um damit gegenüber der Konkurrenz Vorteile zu erzielen.<sup>425</sup>

Die Zielsetzung des CRM ist die Kundenbeziehung so auszugestalten und zu entwickeln, dass man dadurch sowohl für das Unternehmen als auch für die Kunden Vorteile schafft. Als ein kundenorientierter Managementansatz, der durch den Einsatz integrierter kundenorientierter Informationssysteme unterstützt wird, ermöglicht es dem Unternehmen die Sammlung, Bereitstellung und Nutzung von Wissen über den Kunden.<sup>426</sup>

Dagegen versucht das Wissensmanagement, vorhandenes Wissen im Unternehmen effektiver zu nutzen und in Zukunft benötigtes Wissen aufzubauen.<sup>427</sup> Dafür muss relevantes Wissen identifiziert und dieses dann im Unternehmen so gesteuert werden, dass es zielgerichtet in den Geschäftsprozessen zur Verfügung gestellt, genutzt und weitergegeben wird. Durch die bessere Wissensversorgung des Unternehmens können die geplanten Ziele schneller, günstiger und vor allem mit höherer Qualität erreicht werden.

Aus der integrierten Anwendung von CRM und Wissensmanagement entsteht das sogenannte Customer Knowledge Management (CKM). „CKM verbindet die Konzepte des Wissensmanagements und Customer Relationship Management, indem es die erfolgversprechenden Bestandteile integriert.“<sup>428</sup> Dadurch sollen die Risiken und Mängel beider Ansätze vermindert und Synergiepotentiale zwischen CRM und Wissensmanagement ausgeschöpft werden.<sup>429</sup> Dieser kombinierte Ansatz von CRM und Wissensmanagement führt schließlich zu einer Verbesserung der kundenorientierten Geschäftsprozesse in Marketing, Verkauf und Service. Aus diesem Grund ist das gemeinsame Zusammenwirken von CRM und Wissensmanagement entscheidend für eine erfolgreiche kundenorientierte Systementwicklung, in der Wissensmanagement einerseits den ergänzenden Teil und andererseits die notwendige Grundlage für das Management der Kundenbeziehung darstellt.<sup>430</sup>

---

<sup>424</sup> Vgl. Sexauer (2004), S. 22.

<sup>425</sup> Vgl. Kolbe/Österle/Brenner/Geib (2003), S. 4.

<sup>426</sup> Vgl. Oberweis/Paulzen/Sexauer (2002), S. 117.

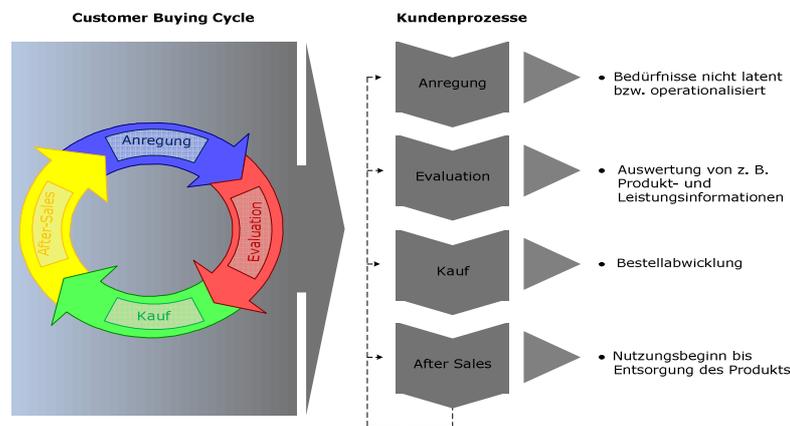
<sup>427</sup> Vgl. Kolbe/Österle/Brenner/Geib (2003), S. 4.

<sup>428</sup> Kolbe/Österle/Brenner/Geib (2003), S. 8.

<sup>429</sup> Vgl. Kolbe/Österle/Brenner/Geib (2003), S. 8.

<sup>430</sup> Vgl. Sexauer (2004), S. 22.

Für eine erfolversprechende Zusammenführung beider Konzepte CRM und Wissensmanagement zu einem integrierten System wurden von Oberweis, Paulzen und Sexauer ein Vorgehensmodell entworfen, welches die Modellierung von Kunden-, CRM- und Wissensprozessen explizit berücksichtigt und somit eine Brücke zwischen ihnen bildet, sodass alle relevanten Prozesse in die Systemgestaltung einfließen.<sup>431</sup>



**Abbildung 19: Kundenprozesse**

Quelle: Oberweis/Paulzen/Sexauer (2002), S. 120.

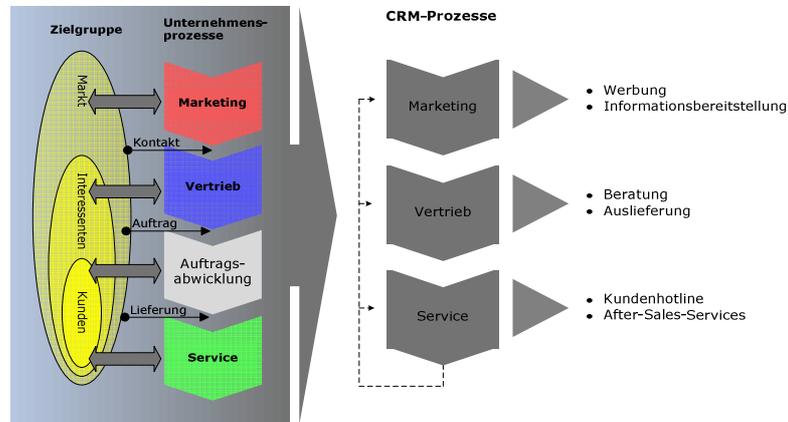
Da im Rahmen von CRM alle Aktivitäten auf die Kundenorientierung ausgerichtet sind, stellen in ihrem Modell die Kundenprozesse den Ausgangspunkt der Systemgestaltung dar (siehe Abb. 19). Der sogenannte Customer Buying Cycle dient hier als Grundlage für die Betrachtung der Kundenprozesse. Er umfasst die Anregungs-, Evaluations-, Kauf- und After-Sales-Phase und spiegelt alle Berührungspunkte zwischen Anbieter und Kunde wider, die im Zusammenhang mit dem Erwerb, Besitz und der Entsorgung einer Marktleistung stehen. Diese systematische Betrachtung der Prozesse ermöglicht dem Unternehmen das Erkennen von Kundenbedürfnissen und vermittelt ein besseres Verständnis über den eigentlichen Kundenprozess.<sup>432</sup>

Im Gegensatz zu den Kundenprozessen, die den Interaktionsprozess der Kunden mit dem Unternehmen beschreiben, stellen die CRM-Prozesse alle unternehmensinternen Prozesse dar, die im direkten Kundenkontakt ablaufen (siehe Abb. 20). Zu diesen gehören die Marketing-, Vertriebs- und Serviceprozesse, die auch als Front Office-Prozesse bezeichnet werden. Die Auftragsabwicklung, im Grunde genommen als eine rein interne Leistungserstellung ohne direkten Kundenkontakt, zählt jedoch nicht dazu. Während der Marketingprozess die Interessentengewinnung und -selektion umfasst und sich damit an die Interessentengruppen richtet, fokussiert der Vertriebsprozess auf die potentiellen zukünftigen Kunden. Demgegenüber richtet sich der

<sup>431</sup> Vgl. Oberweis/Paulzen/Sexauer (2002), S. 118.

<sup>432</sup> Vgl. Oberweis/Paulzen/Sexauer (2002), S. 119.

Serviceprozess ausschließlich an die Kunden, die dem Unternehmen bekannt sind. Diesen drei CRM-Prozessen lassen sich nun die vier Phasen des Customer Buying Cycle zuordnen. So deckt der Marketingprozess die Anregungsphase, der Vertriebsprozess die Evaluations- und Kaufphase und der Serviceprozess die After-Sales-Phase ab.

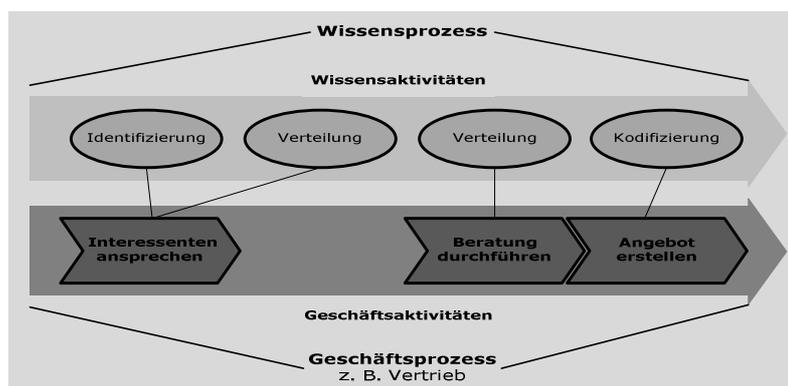


**Abbildung 20: CRM-Prozesse**

Quelle: Oberweis/Paulzen/Sexauer (2002), S. 120.

Da die Ausführung von Kunden- bzw. CRM-Prozessen implizit die Durchführung von Wissensaktivitäten beinhaltet, können diese auch als Wissensprozesse interpretiert werden.<sup>433</sup>

Demnach werden die Wissensaktivitäten zu Wissensprozessen zusammengefasst und dann zur Unterstützung des geschäftsprozessbezogenen Wissensflusses den Kunden- bzw. CRM-Prozessen zugeordnet. So lassen sich die Wissensaktivitäten „Wissen identifizieren“ und „Wissen verteilen“ z. B. der Geschäftsaktivität „Interessenten ansprechen“ zuordnen (siehe Abb. 21).



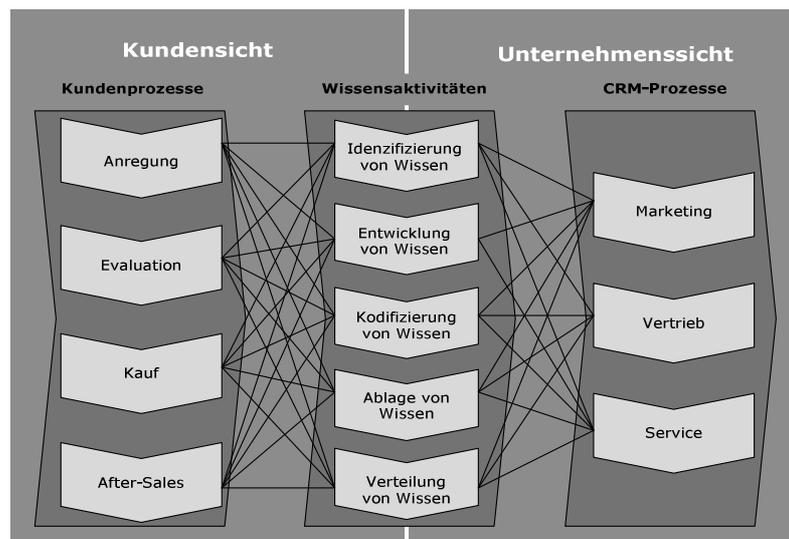
**Abbildung 21: Sicht auf Geschäftsprozesse als Wissensprozesse**

Quelle: Oberweis/Paulzen/Sexauer (2002), S. 121.

Der Vorteil dieser Vorgehensweise liegt darin, dass jede Geschäftsaktivität

<sup>433</sup> Vgl. Oberweis/Paulzen/Sexauer (2002), S. 121.

systematisch im Hinblick auf den Umgang mit Wissen analysiert werden kann.<sup>434</sup>



**Abbildung 22: Interdependenzen zwischen Prozessen**

Quelle: Oberweis/Paulzen/Sexauer (2002), S. 123.

Zudem stellen die Wissensaktivitäten ein Bindeglied zwischen den Kunden- und CRM-Prozessen dar (siehe Abb. 22).<sup>435</sup> „Dies erlaubt die Identifizierung von Schwachstellen sowohl im Prozessablauf als auch im Umgang mit Wissen und somit die Ableitung von Maßnahmen für die Gestaltung von CRM-Systemen.“<sup>436</sup>

Die Unternehmensbereiche Marketing, Vertrieb und Service stellen die primären Kontaktpunkte zwischen Unternehmen und Kunden dar. Zur Verbesserung der Kundenbeziehung muss über alle CRM-Prozesse an diesen Kundenkontakten (Point of Contact PoC) das zur Erfüllung der Aufgaben notwendige Wissen zur Verfügung gestellt werden.<sup>437</sup> Dieses Wissen bezieht sich auf die Leistungen des Unternehmens, aber auch auf das Wissen über seine Kunden. Je mehr Wissen über ihre individuellen Vorlieben, Bedürfnisse und Wünsche verfügbar ist, desto besser kann das Unternehmen seine Kundenbeziehungen gestalten. Aus diesem Grund kommt dem Kundenwissen eine große Bedeutung zu. Auf der einen Seite wird dem Unternehmen mit diesem Wissen ermöglicht, auf die Bedürfnisse des Kunden besser einzugehen und somit die Beziehung des Kunden zum Unternehmen zu vertiefen. Das bildet die Grundlage für ein Umsatzwachstum durch sogenannte Folgekäufe wie in Kapitel 3.7 beschrieben wurde. Durch eine effektive Nutzung des Kundenwissens können zudem die Kosten für die Neukundengewinnung und die bei jeder Transaktion mit dem

<sup>434</sup> Vgl. Oberweis/Paulzen/Sexauer (2002), S. 123.

<sup>435</sup> Vgl. ebd., S. 123.

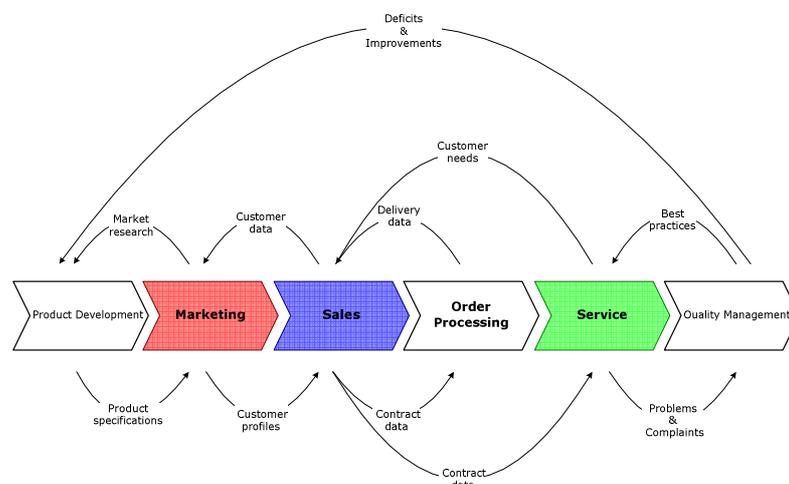
<sup>436</sup> Oberweis/Paulzen/Sexauer (2002), S. 123.

<sup>437</sup> Vgl. Schulze (2002), S. 17.

Kunden zusätzlich anfallenden Kosten eingespart werden. Auf der anderen Seite ist es möglich, auf Basis des über den Kunden gesammelten Wissens, jedem einzelnen Kunden maßgeschneiderte Leistungen anzubieten. Dies nennt man Cross/Up Selling, wie in den vorigen Kapiteln mehrmals erläutert wurde. Weiters dient dieses Wissen bei der Identifikation langfristig profitabler Kundensegmente auf Basis einer gut strukturierten Datenbasis, beispielsweise mit Hilfe eines Data Warehouse-Systems (siehe dafür auch Kapitel 3.6.2).

Um diese Potentiale von Wissen für CRM nutzbar machen zu können, ist eine Koordination des Wissens notwendig. Die Realisierung des Wissensmanagements beim CRM setzt daher eine Regelung der Wissensflüsse zwischen CRM-Prozessen voraus.<sup>438</sup> Das Wissensmanagement wird von den CRM-Systemen praktisch umgesetzt, indem sie das notwendige Wissen für alle CRM-Prozesse integriert bereitstellen.<sup>439</sup> Somit wird das in einem Prozess entstandene Wissen nach der Erfassung im CRM-System für alle CRM-Prozesse verfügbar.

Die Abbildung 23 zeigt die Wissensflüsse zwischen den CRM-Prozessen sowie dem vorgelagerten Produktentwicklungsprozess und dem nachgelagerten Qualitätsmanagementprozess.



**Abbildung 23: Wissensflüsse im CRM-Prozess**

Quelle: Schulze (2002), S. 19.

Zur Deckung des Wissensbedarfs übernimmt das CRM-System die Koordination der Wissensflüsse zwischen den Prozessen, indem es ermöglicht, dass die Mitarbeiter verschiedener Prozessbereiche gemäß ihren Berechtigungen im CRM-System auf die zentrale Datenbasis zugreifen und somit das erforderliche Wissen zum richtigen

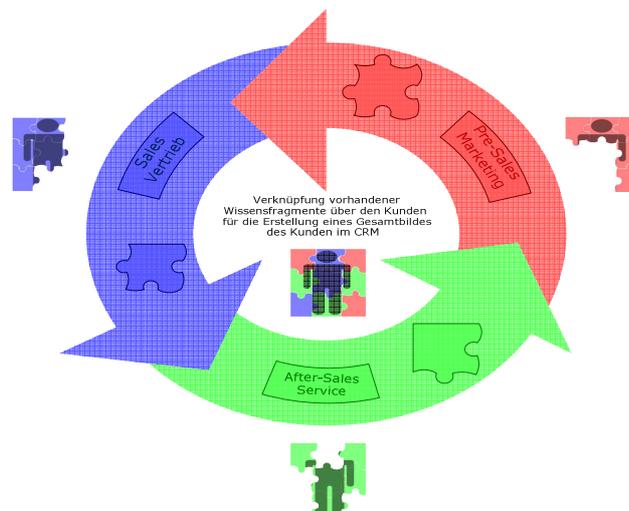
<sup>438</sup> Vgl. Schulze (2002), S. 17.

<sup>439</sup> Vgl. ebd., S. 18-19.

Zeitpunkt erhalten.<sup>440</sup> Die Voraussetzung dafür ist, dass die Mitarbeiter motiviert sind, ihr Wissen auch mit anderen Berechtigten im System zu teilen. Denn ohne den gemeinsamen Austausch von Wissen wird weder die Generierung, die Repräsentation noch die Nutzung von Wissen zum Tragen kommen.

Die folgende Aussage betont die Notwendigkeit der Wissensteilung unter Mitarbeitern: „In post-capitalism, power comes from transmitting information to make it productive, not from hiding it.“<sup>441</sup> Trotzdem ist die Wissensteilung einerseits aufgrund von befürchteten Machtverlusten der Mitarbeiter und andererseits aufgrund der organisatorischen Trennung der Prozessbereiche Marketing, Vertrieb und Service im Unternehmen nicht so effektiv umgesetzt.<sup>442</sup> Jeder Bereich verwaltet getrennt von den anderen Bereichen sein eigenes Kundenwissen und macht sich ein ganz verschiedenes Bild vom selben Kunden.

Aufgrund des getrennten Sammelns von Wissen jeweils in den drei Bereichen entstehen sogenannte Wissenslücken.<sup>443</sup>



**Abbildung 24: Erstellung eines Gesamtbildes des Kunden im CKM**

Quelle: Eigene Darstellung in Anlehnung an Peter Scheir (2002).

Mit Hilfe eines bereichsübergreifenden Wissensmanagements müssen diese Wissenslücken geschlossen und die Wissensfragmente aus unterschiedlichen Kundenkontaktpunkten zusammengefügt werden, um ein möglichst solides Bild des Kunden zu schaffen (siehe Abb. 24). Dieses ganzheitliche Kundenwissen wird dann jedem Bereich zur Verfügung gestellt. Ferner soll an jedem Kundenkontaktpunkt neu gewonnenes Wissen schnell in das CRM-System eingebracht werden, um eine

<sup>440</sup> Vgl. Schulze (2002), S. 19.

<sup>441</sup> Drucker (1993), S. 120, zit. in: Schulze (2002), S. 20.

<sup>442</sup> Vgl. Schulze (2002), S. 120.

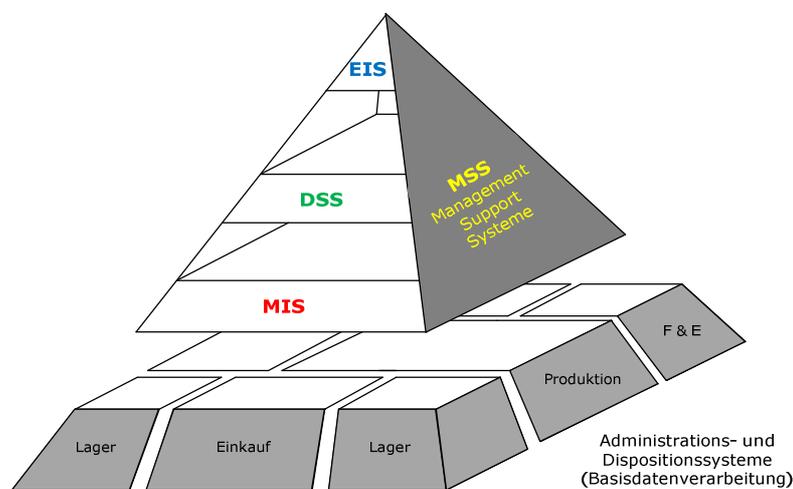
<sup>443</sup> Vgl. Scheir (2002), S. 47-48.

kontinuierliche Erweiterung der Wissensbasis zu gewährleisten. So kann das Unternehmen in nachfolgenden (Folgekäufe) oder parallel ablaufenden (Cross/Up Selling) Sales Cycles besser auf die Wünsche und Bedürfnisse seiner Kunden eingehen und sich nach außen mit „One Face to the Customer“ präsentieren.

## 4.5 COMPUTERUNTERSTÜTZUNG IM MANAGEMENT

Der Wunsch, den Produktionsfaktor Information als Unterstützung von Unternehmensentscheidungen bestmöglich auszunutzen, ist nicht neu. Bereits seit den sechziger Jahren sind elektronische DV- und Informationssysteme zur Entscheidungsunterstützung im Einsatz. Im Laufe der Zeit und durch die rasanten technologischen Entwicklungen bei den Datenbanktechnologien, den Personal Computers und den Vernetzungsmöglichkeiten wurden diese deutlich verbessert, so dass sie schließlich interaktiver geworden sind.<sup>444</sup>

Die Struktur der Informationssysteme zur betrieblichen Informationsversorgung lässt sich anhand einer Systempyramide darstellen (siehe Abb. 25). So bilden die wert- und mengenorientierten Administrations- und Dispositionssysteme auf der unteren Stufe der Pyramide die Basisabläufe des operativen Unternehmensgeschehens ab.<sup>445</sup>



**Abbildung 25: Systempyramide**

Quelle: In Anlehnung an Gluchowski/Gabriel/Dietmar (2008), S. 87.

Die Administrationssysteme, die teilweise auch als Transaktionssysteme bezeichnet werden und auf umfangreiche, in Datenbanken strukturiert abgelegte Datenbestände zugreifen, dienen hauptsächlich der Rationalisierung der Massendatenverarbeitung.<sup>446</sup> Im Vordergrund stehen dabei einfache Tätigkeiten wie z. B. die Verwaltung von

<sup>444</sup> Vgl. Schildhauer/Grothe/Braun/Schultze (2004), S. 25.

<sup>445</sup> Vgl. Muksch/Behme (1998), S. 13.

<sup>446</sup> Vgl. Muksch/Behme (1998), S. 13.

Kunden-, Lieferanten- und Produktstammdaten, die Erfassung, Bearbeitung und Kontrolle von Kundenaufträgen, Lagerbeständen, Produktionsvorgaben und Bestellungen oder Adressenschreiben und Tabellendrucke. Dagegen unterstützen die Dispositionssysteme die Lösung von einfachen (gut strukturierten) Entscheidungsproblemen bei Routine- bzw. Standardaufgaben.<sup>447</sup> Beispiele hierfür sind Bestelldispositionssysteme, Produktionssteuerungssysteme oder auch das Mahnwesen einer Debitorenbuchhaltung. Somit bilden die Administrations- und Dispositionssysteme die Grundlage der als management-unterstützende Systeme bezeichneten zweiten Stufe der Systempyramide. Sie benutzen die auf der ersten Ebene erfassten Daten und verdichten sie zum Zwecke der Entscheidungsunterstützung.

Resultierend aus den unterschiedlichen Anforderungen ergeben sich jeweils typische Merkmale für die operativen und für die management-unterstützenden Informationssysteme, wie in der folgenden Tabelle 4 dargestellt ist.

MERKMAL	Operative DV-Systeme	Managementunterstützende Systeme
Typische Datenstruktur	flache, nicht hierarchische Tabellen	multidimensionale Strukturen
Identifikationskriterium	eindimensional	mehrdimensional
Datenmanipulation	zeilenbezogen/aktualisierend	sichtspezifisch/analysierend
Datenmenge/Transaktion	klein	sehr umfangreich
Betrachtungsebene	detailliert	aggregiert
Zeithorizont	gegenwärtig	historisch, gegenwärtig, zukünftig

**Tabelle 4: Gegenüberstellung der operativen u. managementunterstützenden Datenverarbeitungs-Systeme**

Quelle: Eckerson (1994), S. 6 bei Muksch/Behme (1998), S. 14.

Historisch gesehen stellen die *Management Information Systems (MIS)* die ersten Bemühungen der EDV-Unterstützung des Managements dar. Sie entstanden Ende der 60er Jahre, um den Managern verschiedener Hierarchieebenen zu erlauben, detaillierte und verdichtete Informationen aus der operativen Datenbasis ohne (meist aufwendige) Modellbildung und logisch-algorithmische Bearbeitung zu extrahieren.<sup>448</sup> Basierend auf den operativen Basissystemen zur Administration und Disposition

<sup>447</sup> Vgl. Muksch/Behme (1998), S. 14 und Gluchowski/Gabriel/Dittmar (2008), S. 6.

<sup>448</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 56.

greifen sie verdichtend auf deren Daten zu und bieten somit ein DV-gestütztes Standardberichtswesen mit einfachen algorithmischen Auswertungen.<sup>449</sup>

Die damaligen Realisierungsversuche, die auf den verschiedenen Unternehmensebenen und -bereichen anfallenden Daten für ein Real-Time Management zu verdichten und in einem MIS bereitzustellen, sind jedoch aus zwei Gründen gescheitert:<sup>450</sup> Erstens war die gegebene Hard- und Software nicht in der Lage, Datenbestände in gewünschtem Volumen und notwendiger Schnelligkeit vorzuhalten. Zweitens war der starke Wunsch danach, mit einem Informationssystem alle Informationsbedürfnisse der Entscheidungsträger (Manager) abzudecken, ein konzeptioneller Denkfehler. Außerdem wurden diese Systeme damals häufig kritisiert, dass sie das vorhandene Informationsdefizit durch eine Informationsflut ersetzen, weil die Methoden zur Filterung, Säuberung und Verdichtung der Daten damals nicht hinreichend ausgereift waren.

Anfang der 70er Jahre folgten dem MIS als dessen Weiterentwicklung die *Decision Support Systems (DSS)* - oder *Entscheidungsunterstützungssysteme (EUS)*.

Im Gegensatz zu den Management Information-Systemen, bei denen die reine Datenversorgung des Managements mit zeit- und sachgerechter Information im Vordergrund steht, fokussierten sie auf die effektive Unterstützung im Planungs- und Entscheidungsprozess mit dem Ziel, das Urteilsvermögen des Managers und dadurch seine Entscheidungsqualität zu verbessern.<sup>451</sup>

Eine allgemeingültige Definition der DSS lautet wie folgt: DSS sind interaktive EDV-gestützte Systeme, die Manager (Entscheidungsträger) mit Modellen, Methoden und problembezogenen Daten in ihrem Entscheidungsprozeß bei der Lösung in strukturierten und semi-strukturierten Entscheidungssituationen unterstützen.<sup>452</sup> Charakteristisch für diese Systeme ist die Modell- und Methodenbildung, die eine situationspezifische Unterstützung des Managers im Sinne einer Assistenz ermöglicht.<sup>453</sup> Dabei werden oft wissensbasierte Techniken (Techniken der künstlichen Intelligenz) eingesetzt, die sowohl zur intelligenten Lösungsfindung als auch zur Repräsentation und Modellierung von Wissen bzw. Information dienen.<sup>454</sup>

---

<sup>449</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 57.

<sup>450</sup> Vgl. Muksch/Behme (1998), S. 16f und Gluchowski/Gabriel/Dittmar (2008), S. 61f.

<sup>451</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 63.

<sup>452</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 57. Wenn aber eine Notwendigkeit vorliegt, können die DSS dann auch bei unstrukturierten Problemen oder für die strategische Planung (strategisches Management) eingesetzt werden.

<sup>453</sup> Vgl. ebd., S. 63.

<sup>454</sup> Vgl. ebd., S. 64.

Der Nachteil der früher erschienenen DSS liegt jedoch darin, dass sie sich auf Teilprobleme spezialisierten und daher keine unternehmensübergreifende Modelle zur Simultanplanung anbieten konnten.<sup>455</sup> Zudem waren sie für die intuitive Problemerkennung und Wahrnehmung von Signalen durch den Manager ungeeignet.<sup>456</sup> Diese Systeme leisten aber bei eingegrenztem Problemlösungsumfang und strukturierten Problemen gute Dienste, daher werden sie zurzeit häufig in die ERP-Systeme von Unternehmen integriert und sie sind flächendeckend im Einsatz.

Mit der fortschreitenden Vernetzung der DV-Systeme, der Dezentralisierung von EDV-Leistungen und der Verbreitung von leistungsstarken und benutzerfreundlichen PCs kam Mitte der 80er Jahre eine neue Welle der Verbesserung des MIS-Ansatzes aus den USA – die *Executive Information Systems (EIS)*.

Diese Systeme zeichnen sich durch eine starke Präsentations-Orientierung aus. Als dialog- und datenorientiertes Abfrage- und Berichtssystem mit ausgeprägten Kommunikationselementen bieten sie den Managern, auf Knopfdruck, aktuelle, entscheidungsrelevante, interne und externe Informationen über intuitiv benutzbare und individuell anpassbare Benutzeroberflächen an.<sup>457</sup> Sie können sowohl in den frühen Phasen des Planungs- und Entscheidungsprozesses, in denen der Entscheidungsträger einen explorativen „Data Support“<sup>458</sup> benötigt, um frühzeitig die Entwicklungstendenzen zu erkennen und die notwendigen Analysen dafür durchzuführen, als auch in der Kontrollphase zur Überprüfung der Auswirkungen angeordneter Maßnahmen eingesetzt werden.<sup>459</sup> Im Vergleich zu den Decision Support- Systemen sind sie aber eher modell- und methodenarm<sup>460</sup>, wodurch sie sich im täglichen Gebrauch als zu starr und unflexibel erwiesen, um mit den wechselnden Anforderungen betrieblicher Entscheidungsträger umgehen zu können.<sup>461</sup>

Durch die Zusammenführung von DSS und EIS in einem integrierten bzw. kombinierten Ansatz entstanden die sogenannten *Executive Support Systems (ESS)*. Diese Systeme streben eine ganzheitliche, phasen- und problemübergreifende Unterstützung des Managements an, indem sie einerseits die Visualisierungs- und

---

<sup>455</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 74.

<sup>456</sup> Vgl. ebd., S. 63. Ein entscheidungsunterstützendes Informationssystem soll generell genug Raum für die subjektive Urteilskraft des Managers lassen, damit er auch seine eigenen Erfahrungen für die Problemlösung nutzen kann.

<sup>457</sup> Vgl. ebd., S. 75.

<sup>458</sup> Der Begriff Data Support alleine bezeichnet die reine Bereitstellung von Informationen ohne gezielte Ausrichtung auf den Benutzer.

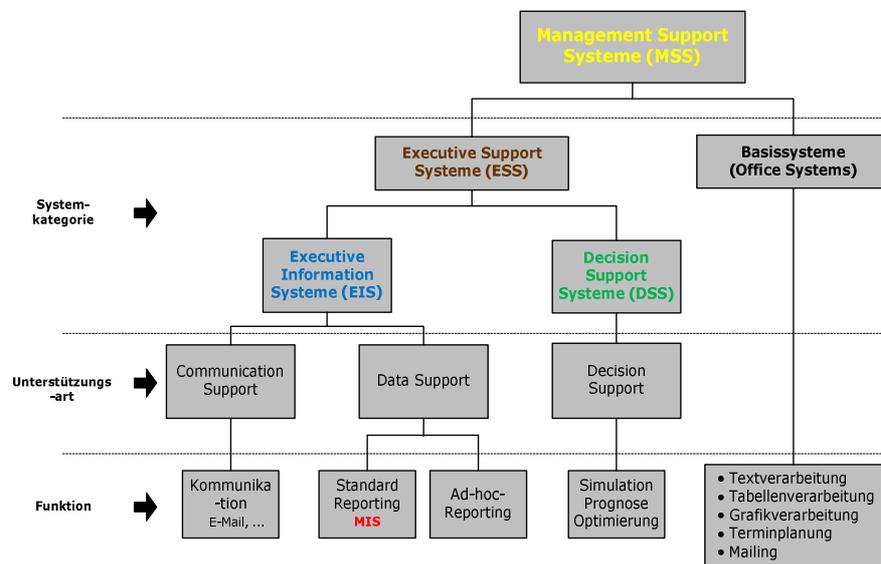
<sup>459</sup> Vgl. ebd., S. 76.

<sup>460</sup> Vgl. ebd., S. 76.

<sup>461</sup> Vgl. Chamoni/Gluchowski (2006), S. 8.

Präsentationsformen von EIS zur raschen Aufdeckung von Zusammenhängen nutzen und andererseits die betriebswirtschaftlichen Kausalmodelle und Methoden zur Analyse, Prognose, Simulation und Optimierung im Sinne einer DSS-Unterstützung anbieten.<sup>462</sup> Sie sind leistungsfähiger als die der separat betrachteten und genutzten EIS und DSS, da durch die Verbindung der Benutzerfreundlichkeit, der graphischen Informationsaufbereitung und der Reduktion komplexer Informationszusammenhänge des EIS mit der entscheidungsunterstützenden Modell- und Analysefunktion des DSS Synergien ausgeschöpft werden können.<sup>463</sup>

Wie in der Abbildung 26 zu sehen ist, lassen sich alle diese Entscheidungsunterstützungssysteme letztlich unter dem Begriff Management Support-Systeme (MSS) zusammenfassen.



**Abbildung 26: Bestandteile von Management Support-Systemen**

Quelle: in Anlehnung an Gluchowski/Gabriel/Dittmar (2008), S. 85.

Dieser Begriff umfasst dann „[...] alle Einsatzformen von Datenverarbeitungs-, Informations- und Kommunikationstechnologien zur Unterstützung unternehmerischer Aufgaben“<sup>464</sup>, wobei sie eine „[...] individuelle, konzeptionelle Lösung zur Steuerung und Kontrolle von Unternehmungen auf informationstechnisch modernstem Niveau“<sup>465</sup> versprechen und das gesamte Unterstützungsspektrum von Managern durch den Einsatz von Computern und Informations- bzw. Kommunikationstechnologien abdecken.<sup>466</sup>

Die grundlegenden Probleme früherer entscheidungsorientierter Informationssysteme

<sup>462</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 82f.

<sup>463</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 83.

<sup>464</sup> Scott Morton (1983), zit. in: Gluchowski/Gabriel/Dietmar (2008), S. 86.

<sup>465</sup> Krallmann/Rieger (1987), S. 29, zit. in: Gluchowski/Gabriel/Dittmar (2008), S. 86.

<sup>466</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 86.

waren, wie oben schon erwähnt, die fehlende Interaktivität und übermäßige Starrheit, die durch die Überflutung der Informationen verursachte mangelhafte Datenversorgung sowie die unzureichende Integrierbarkeit in bestehende DV-technische und organisatorische Infrastrukturen. Zur effizienten Bewältigung dieser Probleme wurde der Business Intelligence-Ansatz entwickelt, welcher vorwiegend auf die Weiterentwicklungen von Data Warehousing, On-Line Analytical Processing (OLAP) und Data Mining zurückgreift. Basierend auf diesen technologischen Konzepten lassen sich dann betriebswirtschaftliche Anwendungslösungen für analytisches Customer Relationship Management entwerfen und implementieren.

## **4.6 BUSINESS INTELLIGENCE**

Business Intelligence ist heute von vielen Unternehmen als eines der wichtigsten Konzepte zur Generierung von Wissen über das Unternehmen anerkannt und bildet somit den Ausgangspunkt und die Grundlage der Unternehmensentwicklung.

In der Literatur sind verschiedene Beschreibungen des Begriffs Business Intelligence (BI) zu finden, doch existiert keine allgemein anerkannte Definition dafür. Fachleute sind sich aber heute darüber einig, dass es sich bei Business Intelligence eher um einen Oberbegriff handeln sollte, unter welchem viele unterschiedliche Ansätze und Werkzeuge zur Analyse von Geschäftsprozessen und erfolgskritischen Wirkungszusammenhängen zusammengefasst werden können.<sup>467</sup> Dazu zählen alle Werkzeuge und Anwendungen mit entscheidungsunterstützendem Charakter, die zur besseren Einsicht in das Unternehmensgeschäft und damit zum besseren Verständnis in die Ursache/Wirkung- Beziehungen führen.<sup>468</sup>

Die Abbildung 27 veranschaulicht die Einordnung unterschiedlicher Konzepte und Technologien im Business Intelligence. Die Darstellung der Abbildung erfolgt durch zwei verschiedene Dimensionen.

Die vertikale Dimension deckt dabei die Wissensgenerierung (von der Datenbereitstellung bis zur Datenauswertung) ab, während die horizontale Dimension die Ausrichtung der eingesetzten Systemkomponenten bestimmt (technikgetrieben oder anwendungszentriert). Auf der Grundlage dieser Darstellung der Abbildung unterscheidet Gluchowski drei Arten des BI-Verständnisses:<sup>469</sup>

- *Weites BI-Verständnis*

---

<sup>467</sup> Vgl. Dittmar/Gluchowski (2002), S. 32.

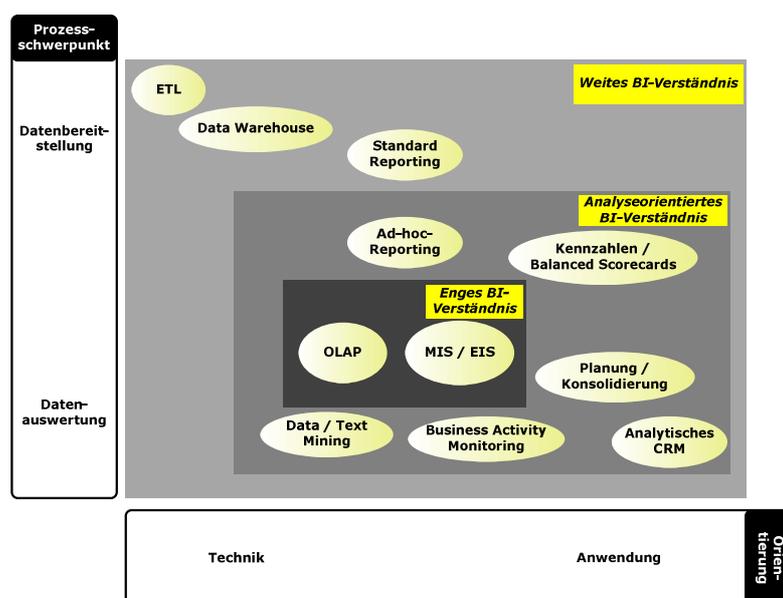
<sup>468</sup> Vgl. Dittmar/Gluchowski (2002), S. 32.

<sup>469</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 90.

- *Analyseorientiertes BI-Verständnis*
- und *Enges BI-Verständnis*

Nach dem weiten BI-Verständnis gehören alle Systemkomponenten zu Business Intelligence, die operatives Datenmaterial zur Informations- und Wissensgenerierung aufbereiten und speichern sowie Auswertungs- und Präsentationsfunktionalität anbieten. Hierzu zählen die ETL-Werkzeuge, Data Warehouses und sämtliche analytische Applikationen für Standard Reporting.

Dagegen fokussiert sich das analyseorientierte BI-Verständnis auf die Tools und Anwendungen, die modell- und methodenbasiert eine zielgerichtete Analyse von vorhandenem Datenmaterial ermöglichen. In diesem Kontext umfassen die BI-Tools insbesondere die On-Line Analytical Processing (OLAP)-Tools und Data Mining-Produkte. Die Werkzeuge zur Verarbeitung unstrukturierter Daten sowie zur Analyse bzw. Überwachung der Geschäftsprozesse (Business Activity Monitoring- und Process Performance Management-Systeme) und die analytischen CRM-Lösungen, bei denen statistische Methoden häufig angewendet werden, gehören aber auch zum analyseorientierten Business Intelligence. Letztlich können die Balanced Scorecard- und Kennzahlensysteme sowie die Systeme zur Planung und Budgetierung, Konzernkonsolidierung bzw. zum Risiko-Management auch dem analyseorientierten BI zugeordnet werden, sofern sie den analytisch arbeitenden Anwender bei seinen Aufgaben unterstützen.



**Abbildung 27: Einordnung unterschiedlicher Technologien und Abgrenzungen von Business Intelligence**

Quelle: in Anlehnung an Gluchowski/Gabriel/Dittmar (2008), S. 85.

Die engere Auffassung des BI-Verständnisses bezieht sich auf den Bereich analyseorientierter Produkte und Anwendungen, die eine Aufbereitung und Präsentation von multidimensional organisiertem Datenmaterial mit den fortgeschrittenen Techniken wie „Slicing und Dicing“<sup>470</sup> oder „Color Coding“<sup>471</sup> ermöglichen. Als BI-Tools werden dann herstellereigene Client-Lösungen, Briefing-Books, Excel-Add-Ins oder Browser-Erweiterungen bezeichnet, die den Umfang der verfügbaren Standardfunktionalität von OLAP-Tools und MIS/EIS ergänzen.

Betrachtet man sich diese Darstellung entlang der Dimensionen weiter, erkennt man unterschiedliche Bereiche, an denen verschiedene Systeme ansetzen.

So sind im oberen Teil der Darstellung die Ansätze positioniert, die vorwiegend zur Sammlung, Bereitstellung und Speicherung analyserelevanter Basisdaten dienen. Der untere eher analyseorientierte Bereich hebt die Aspekte hervor, zu deren Realisierung die methodischen Systemkomponenten verwendet werden, die das vorhandene Datenmaterial für die weiterführenden Analysen nutzen. Der mittlere Bereich als reine Auswertungsebene ist durch die Werkzeuge geprägt, die durch verschiedene Techniken besonders auffällige oder interessante Daten identifizieren und deren Gegebenheit entsprechend das Datenmaterial als Geschäftsgrafik präsentieren können. Die Konzepte im linken Teil sind durch ihre technische Ausrichtung charakterisiert und werden daher öfters in IT-Abteilungen und Stabstellen des Unternehmens genutzt. Die Systeme im rechten Teil, mit denen die betriebswirtschaftlichen themenspezifischen Aufgabenbereiche des Unternehmens abgedeckt werden, sind dagegen anwendungsorientiert,.

Neben den oben diskutierten werkzeugs- und anwendungsorientierten Sichtweisen auf BI existiert noch die Möglichkeit, Business Intelligence aus einer prozessfokussierten Perspektive zu betrachten.<sup>472</sup> In diesem Fall wird Business Intelligence folgendermaßen definiert: „Business Intelligence (BI) bezeichnet den analytischen

---

<sup>470</sup> Slicing und Dicing sind dynamische Anzeigetechniken zur interaktiven Navigation in einer multidimensionalen Datentabelle oder -matrix, die auch als Datenwürfel (Data Cube) bezeichnet wird. Die Zahl der möglichen Betrachtungsschichten innerhalb einer multidimensionalen Datenstruktur steigt exponentiell mit der Anzahl der Dimensionen/Ebenen. So hat eine zweidimensionale Datenmatrix zwei Betrachtungsschichten, eine dreidimensionale hat sechs, eine vierdimensionale 24 und eine fünf dimensionale 120 Betrachtungsschichten. Ranging oder auch Dicing (d. h. in Würfel schneiden) ermöglicht dem Benutzer, beliebige Elemente von Dimensionen eines Datenwürfels ein- oder auszublenden. Dadurch kann er eine Anzeige auf diejenigen Elemente (auf eine Teilmenge) beschränken, die ihn besonders interessieren. Dadurch wird ein kleinerer Würfel erzeugt, der einen Teilbereich des Gesamtwürfels enthält. Bei Slicing (d. h. Ausschneiden von Scheiben aus dem Würfel) geht es darum, dass der Benutzer jedesmal unterschiedliche Dimensionen auswählt und sich dann nur diese anzeigen lässt.

<sup>471</sup> Eine spezielle Technik zur Repräsentation von Systemausgaben durch Farben. In einem EIS werden z. B. die auftretenden Abweichungen mit farblichen Markierungen versehen, wodurch sich der Manager schnell einen umfassenden Eindruck über die aktuelle Geschäftssituation in einer geographischen Region verschaffen kann.

<sup>472</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 91.

Prozess, der – fragmentierte – Unternehmens- und Wettbewerbsdaten in handlungsgerichtetes Wissen über die Fähigkeiten, Positionen, Handlungen und Ziele der betrachteten internen oder externen Handlungsfelder (Akteure und Prozesse) transformiert.“<sup>473</sup>

Grothe und Gentsch machen es mit dieser Definition deutlich, dass Business Intelligence als kontinuierlicher Prozess zu verstehen ist. Das verfügbare Datenmaterial als Ausgangsbasis für jeden Analysefall ändert sich ständig, so dass immer wieder neue Analysepotentiale entdeckt werden können. Die sich ändernden Umwelt- und Umfeldbedingungen und ihre Einwirkungen auf die Ausgangssituation machen eine ständige Anpassung notwendig, die im Idealfall durch einen kontinuierlichen Prozess erreicht werden kann.<sup>474</sup> Jede Anpassung generiert neues Wissen, und dieses muss dann gezielt für die weiteren Aktionen verwendet werden, die bereits vorhandene Strukturen und Abläufe verändern können.<sup>475</sup> Somit wird es dem Unternehmen ermöglicht, die Chancen und Risiken frühzeitig zu erkennen und diese für die eigenen Ziele nutzbar zu machen. Zur Datenanalyse und –abfrage müssen auch bei der prozessorientierten BI-Sichtweise viele Techniken und Werkzeuge verwendet werden. Diese Verwendung von Werkzeugen und Techniken trägt zu einer Veredelung der Datenbasis bei.<sup>476</sup> Die Basisdaten werden dabei schrittweise so aufbereitet, dass sie eine verbesserte Identifikation der ihnen zugrunde liegenden Zusammenhänge ermöglichen.

So lässt sich leicht feststellen, dass das prozessorientierte Business Intelligence in drei Prozessphasen strukturiert werden kann:<sup>477</sup>

- *Data Delivery*: Die Bestimmung und Bereitstellung von Basisdaten als Grundlage für künftige Analysen (z. B. in einem Data Warehouse). Die Daten können strukturiert, wie auch unstrukturiert, oder quantitativ bzw. qualitativ sein.
- *Discovery of Relations, Patterns, and Principles*: Die Entdeckung relevanter Zusammenhänge, Muster und Musterbrüche oder Diskontinuitäten innerhalb der vorgehaltenen Datenbasis. Diese Entdeckung kann entweder hypothesenfrei oder anhand einer vorbestimmten Hypothese erfolgen.

---

<sup>473</sup> Grothe/Gentsch (2000), S. 19.

<sup>474</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 91.

<sup>475</sup> Diese Aktionen reichen von einer kurzfristigen Marketingaktivität bis zu einer organisatorischen Modifikation, z. B. in der Aufbau- und Ablauforganisation.

<sup>476</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 91.

<sup>477</sup> Vgl. Grothe/Gentsch (2000), S. 19-21.

- *Knowledge Sharing*: Die Kommunikation für die Verteilung des aus Daten generierten Wissens im Unternehmen und dessen Integration in die operativen Unternehmensabläufe. Dafür ist ein erfolgreiches Wissensmanagement relevant (siehe dafür auch Kapitel 4.3 und 4.4).

Diese oben erwähnten unterschiedlichen Betrachtungsweisen des BI (die technische-, anwendungsorientierte- und prozessorientierte BI-Sichtweise) könnte allerdings dafür sorgen, dass man beim Business Intelligence Verständnis in Verwirrung gerät. Um dies zu verhindern, kann hier die folgende Definition für Business Intelligence vorgeschlagen werden: Business Intelligence versucht eine Vielzahl von Konzepten und Technologien im Umfeld der entscheidungsunterstützenden Systeme zu bündeln (technische Begriffsdefinition), um somit die entscheidungsorientierte kontinuierliche Sammlung und Aufbereitung von Daten über das Unternehmen und dessen Umwelt sowie deren Darstellung in Form von geschäftsrelevanten Informationen für die Analyse-, Planungs- und Steuerungszwecke zu ermöglichen (anwendungs- bzw. prozessorientierte Begriffsdefinition).<sup>478</sup>

#### **4.6.1 Das Data Warehouse-Konzept**

Die Informationsbereitstellung bildet die Grundlage von Managementunterstützungs- bzw. Business Intelligence-Systemen.<sup>479</sup> Die Sammlung und Verdichtung entscheidungsrelevanter Informationen ist nur auf der Basis einer konsistenten unternehmensweiten Datenhaltung möglich.

Der Begriff Data Warehouse bezeichnet generell eine von den operationalen DV-Systemen isolierte relationale Datenbank, die als eine unternehmensweite Datenbasis für alle Formen der managementunterstützenden Systeme dient und durch eine strikte Trennung der operationalen Daten zur Administration und Steuerung von den entscheidungsunterstützenden Daten und Informationen zur Analyse, Planung und Kontrolle gekennzeichnet ist.<sup>480</sup>

Von vielen Autoren wird dieser Begriff oft in die deutsche Sprache als Daten-Warenhaus übersetzt (im Wörterbuch steht jedoch das Wort „Warehouse“ für Lagerhaus), da man den Datenfluss im Unternehmen mit dem Warenfluss im Handel vergleichen kann.<sup>481</sup> Wie in einem Warenhaus holt sich der Kunde nach seinem Bedarf und in Selbstbedienung die Ware „Information“ aus den Regalen in seinen Warenkorb.

---

<sup>478</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 93.

<sup>479</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 117.

<sup>480</sup> Vgl. Muksch/Behme (1998), S. 36 und Meyer/Weingärtner/Döring (2001), S. 98.

<sup>481</sup> Vgl. Chamoni/Gluchowski (2006), S. 130 und Gluchowski/Gabriel/Dittmar (2008), S. 119.

Die Regale sind nach Themen geordnet und das Warenangebot ist kundenorientiert.

William H. Inmon gilt als geistiger Vater des Data Warehouse-Konzeptes und definierte den Begriff folgendermaßen: „A data warehouse is a subject-oriented, integrated, time-variant, nonvolatile collection of data in support of management’s decision-making process.“<sup>482</sup>

Nach dieser Definition ist ein Data Warehouse durch die Merkmale Themenorientierung, Integration, Zeitraumbezug und Beständigkeit gekennzeichnet. Diese Merkmale, die im Folgenden erläutert werden, beziehen sich auf die Datenhaltung und haben somit Auswirkungen auf die Wahl des geeigneten Datenmodells.

#### 1. Themenorientierung:

„Im Gegensatz zur Funktions- und Anwendungsorientierung der operativen Systeme (Z. B. Einkauf, Lagerhaltung, Verkauf u. a.) steht bei der Konzeption des Data Warehouse die Orientierung an Gegenstandsbereichen (Subjekten) des Unternehmens im Vordergrund.“<sup>483</sup> Die innerbetrieblichen Funktionen und Prozesse sind für die Entwicklung des Data Warehouses von untergeordnetem Interesse und haben somit einen geringen Einfluss auf dessen Struktur.<sup>484</sup>

In einem Data Warehouse liegt der Fokus auf den inhaltlichen Themenschwerpunkten wie z. B. Produkte und Kunden, die das Handeln und den Erfolg des Unternehmens bestimmen.<sup>485</sup> Operative Daten, die nur für die Durchführung der täglichen Geschäftsprozesse wichtig sind und somit der Entscheidungsunterstützung kaum dienen, werden daher nicht im Data Warehouse-Umfeld berücksichtigt. Häufig betrachtete Gegenstandsbereiche sind die Unternehmensstruktur (Z. B. Geschäftsbereiche, Organisationsstruktur, rechtliche Einheiten), die Produktstruktur (Z. B. Produktfamilie, Produktgruppe, Artikel), die Regionalstruktur (Z. B. Land, Gebiet, Bezirk), die Kundenstruktur (Z. B. Einzelkunden, Kundengruppen), die Zeitstruktur (Z. B. Monat, Quartal, Jahr, Geschäftsjahr), betriebswirtschaftliche Kenngrößen (Z. B. Umsatz, Deckungsbeiträge, Gewinn) sowie deren Ausprägung (Z. B. Plan, Soll, Ist, Abweichungen).<sup>486</sup>

---

<sup>482</sup> Inmon (1996), S. 33.

<sup>483</sup> Meyer/Weingärtner/Döring (2001), S. 98.

<sup>484</sup> Vgl. Muksch/Behme (1998), S. 40.

<sup>485</sup> Vgl. Muksch/Behme (1998), S. 40.

<sup>486</sup> Vgl. Muksch/Behme (1998), S. 40.

## 2. *Integration:*

„Das DW-Konzept strebt eine unternehmensweite Integration von Daten in einem einheitlich gestalteten System an.“<sup>487</sup> Das Ziel dieser Integration besteht darin, einen konsistenten Datenbestand aufzubauen, der sich trotz der großen Heterogenität von Datenquellen als logisch und akzeptabel darstellt.<sup>488</sup> Diese Integration der Daten erfolgt über eine Struktur- und Formatvereinheitlichung bei der Datenübernahme.<sup>489</sup> Zur Herstellung der semantischen Integrität können verschiedene Umsetzungstabellen angelegt werden, die beispielsweise das Angleichen international unterschiedlicher Kontenrahmen oder die Währungsumrechnung unterstützen.<sup>490</sup>

## 3. *Zeitraumbezug:*

Die zeitpunktbezogene Betrachtung von Daten, wie sie in operationalen Systemen vorgenommen wird, ist für die Managementunterstützung sinnlos, da es dafür Daten benötigt werden, welche die Entwicklung des Unternehmens über einen bestimmten Zeitraum darstellen können.<sup>491</sup> Nur solche Daten werden zur Erkennung und Analyse von Trends herangezogen. Der in einem Data Warehouse abgebildete Zeithorizont von fünf bis zu zehn Jahren macht es möglich, die Zeitreihenanalysen über historische Daten durchzuführen.<sup>492</sup> Jeder Datenimport in das Data Warehouse bietet einen Schnappschuss des Unternehmensgeschehens.<sup>493</sup> Die aus einem DW abgerufenen Daten entsprechen folglich der Aktualität dieses Schnappschusses. Selbst der neueste Schnappschuss von Daten kann schon zum Zeitpunkt der Nutzung dieser Daten als Informationen veraltet sein. Für die analytischen Auswertungen reichen dennoch diese Daten vollkommen aus, wenn sie ein mäßiges Maß an Aktualität aufweisen.<sup>494</sup> Zur Herstellung des Zeitraumbezugs wird der betrachtete Zeitraum (Tag, Woche oder Monat) als Schlüssel in die Daten eingebunden.<sup>495</sup>

---

<sup>487</sup> Meyer/Weingärtner/Döring (2001), S. 100.

<sup>488</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 119.

<sup>489</sup> Bei verschiedenen Anwendungen werden bestimmte Attribute mit unterschiedlichen Abkürzungen belegt werden. So können z. B. bei dem Attribut Geschlecht in einer Anwendung die Abkürzungen „m“ für männlich und „w“ für weiblich stehen, während in einer anderen für die Bezeichnung des Geschlechts „0“ und „1“ verwendet werden. Dasselbe gilt für verschiedene Datumsformaten oder Wertseinheiten zur Größen- und Mengenbestimmung. Z. B. die englische Darstellungsform für Datum „Monat/Tag/Jahr“ unterscheidet von der deutschen „Tag/Monat/Jahr“, vgl. Muksch/Behme (1998), S. 42.

<sup>490</sup> Vgl. Muksch/Behme (1998), S. 43.

<sup>491</sup> Vgl. Muksch/Behme (1998), S. 41.

<sup>492</sup> Vgl. Muksch/Behme (1998), S. 41.

<sup>493</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 120.

<sup>494</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 120.

<sup>495</sup> Vgl. Muksch/Behme (1998), S. 41.

#### 4. Nicht-Volatilität:

Den Begriff „Volatilität“ benutzt man, um damit den Grad der Veränderung von Daten im Laufe ihrer Nutzung zu beschreiben.<sup>496</sup> Um die Nicht-Volatilität der im Data Warehouse gespeicherten Daten zu gewährleisten, müssen nahezu alle Datenzugriffe lesend erfolgen.<sup>497</sup>

Im Gegensatz zu den Daten der operativen Systeme, die nur für einen begrenzten Zeitraum gehalten, innerhalb dessen sie durch verschiedene Operationen geändert und anschließend ausgelagert oder gelöscht werden<sup>498</sup>, verfügen die in einem DW gespeicherten Daten über einen langen Benutzungszeitraum, und sie werden nach der fehlerfreien Übernahme nur in Ausnahmefällen aktualisiert oder verändert. In einem Data Warehouse benutzt man eigentlich nur zwei Operationen, die auf den Daten ausgeführt werden<sup>499</sup>:

- Einmaliges Laden der Daten aus den operativen Systemen
- Zugriff auf die Daten (mit read-only Option)

Ein idealtypisches Data Warehouse-Konzept umfasst vier grundsätzliche Bereiche, wie in der folgenden Abbildung 28 dargestellt ist. Im Folgenden werden diese Bereiche und ihre Elemente vorgestellt und kurz erläutert:

Im Rahmen der Datengewinnung kann es prinzipiell zwischen unternehmensinternen und unternehmensexternen Datenquellen unterschieden werden.

Die unternehmensinternen Daten werden größtenteils aus den operationalen DV-Systemen gewonnen, die im Unternehmen zumeist als Teilsysteme - unabhängig voneinander - eingesetzt sind. Dementsprechend besitzen die aus diesen diversen Teilinformationssystemen gewonnenen Daten ganz unterschiedliche Datenformaten und -strukturen. Durch geeignete Extraktions- und Transformationsprozessen<sup>500</sup> bzw. -programmen lassen sich aber diese Daten ohne Medienbruch in das Data Warehouse übernehmen.<sup>501</sup> Die Unternehmen erstellen ihr Data Warehouse zunächst auf der Basis von unternehmensinternen Daten, da diese bereits vorhanden und strukturiert sind.<sup>502</sup>

---

<sup>496</sup> Vgl. Muksch/Behme (1998), S. 43.

<sup>497</sup> Vgl. Muksch/Behme (1998), S. 43.

<sup>498</sup> um die Performance dieser Systeme (insbesondere hinsichtlich der Response-Time) nicht zu belasten.

<sup>499</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 101.

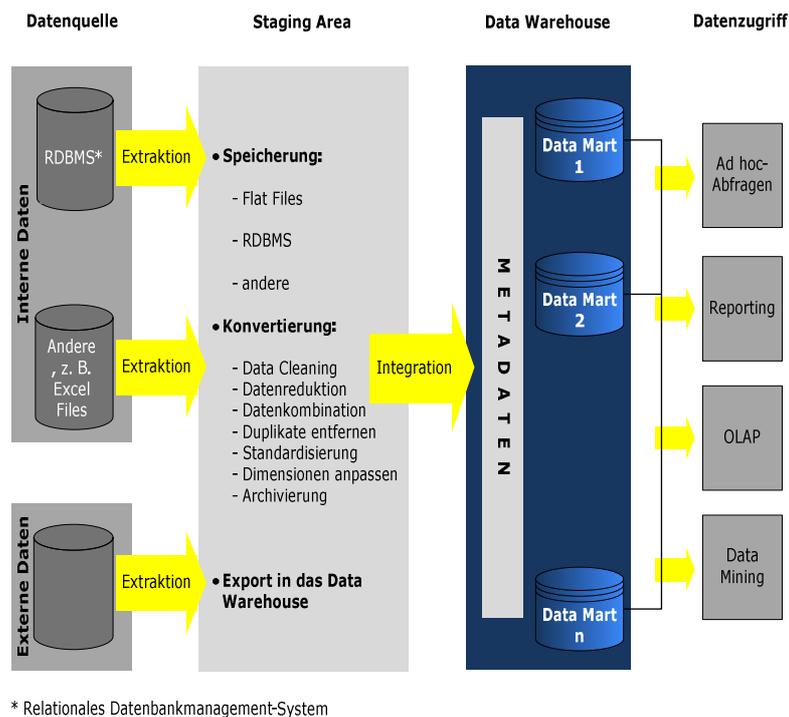
<sup>500</sup> Eine ausführliche Erklärung des Extraktions-, Transformations- und Ladeprozesses findet sich bei Gluchowski/Gabriel/Dittmar (2008), S. 135-139.

<sup>501</sup> Vgl. Muksch/Behme (1998), S. 58.

<sup>502</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 102.

Danach kann auch eine Integration von unternehmensexternen Daten erfolgen.<sup>503</sup>

Viele Auswertungen basierend auf den unternehmensinternen Daten, erlangen erst durch den Vergleich mit den unternehmensexternen Daten eine für den Entscheidungsträger signifikante Bedeutung.<sup>504</sup> Aus diesem Grund kommt externen Daten heutzutage eine wichtige Bedeutung zu. Sie sind im Vergleich zu internen Daten schwer zu beschaffen und liegen meist in unstrukturierter Form vor.



**Abbildung 28: Die Elemente eines Data Warehouse**

Quelle: in Anlehnung an Meyer/Weingärtner/Döring (2001), S. 102.

Vor der Übernahme in das Data Warehouse sollen die internen und externen Daten einer Qualitätskontrolle unterzogen werden. Das erfolgt in der sogenannten Staging Area. Sie umfasst eine Vielzahl von Prozessen, die Daten aus den unterschiedlichen Quellen extrahieren, transformieren und für die eigentliche Nutzung im Data Warehouse vorbereiten.<sup>505</sup> Die Daten werden bis zum Laden in das Data Warehouse sozusagen in der Staging Area zwischengespeichert.<sup>506</sup> Man kann die Staging Area in drei wesentliche Komponenten Monitor, Konvertor und Integrator aufteilen:

- **Monitor:** Über die Monitore werden die Datenquellen auf die Änderungsoperationen überwacht, um verschiedene Aktualisierungs- und

<sup>503</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 102.

<sup>504</sup> Vgl. Muksch/Behme (1998), S. 59.

<sup>505</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 103.

<sup>506</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 131.

Änderungsprozesse anzustoßen.<sup>507</sup> Für die Sicherstellung des erfolgreichen Datenübernahmeprozesses werden vier verschiedene Verfahren angewendet. Das sind Triggering, Modifikation der Anwendungsprogramme, Logfile-Auswertung und Schnappschuss-Mechanismus. Die ersten beiden Verfahren dienen zur sofortigen Aktualisierung der DW-Datenbasis, wogegen die letzten beiden Alternativen dazu benutzt werden, eine Veränderung zu einem bestimmten Zeitpunkt auszulösen.<sup>508</sup>

- *Konverter*: Der Konverter hat die Aufgabe, die über den Monitor erfassten Quelldaten in ein Format zu transferieren, um diese Daten dann in das Data Warehouse integrieren zu können.<sup>509</sup> Er bietet eine einheitliche Schnittstelle zu den verschiedenen Datenquellen und kapselt die Eigenheiten der lokalen Anfrageverarbeitung.<sup>510</sup> Für eine relationale Datenquelle werden bspw. die Anfragen des Data Warehouse in SQL(Structured Query Language)-Code übersetzt oder für ein File System in Funktionsaufrufe des Application Programming Interface (API) umgewandelt.
- *Integrator*: Der Integrator übernimmt die vom Konverter vereinheitlichten Daten, stellt fest, welche Änderungen im Data Warehouse vorgenommen werden müssen und fügt sie in die DW-Datenbasis ein.<sup>511</sup> Hierfür stehen verschiedene Methoden und Techniken zur Auswahl:
  - Bei einem *Data Refresh*<sup>512</sup> werden in bestimmten Abständen alle Daten neu in das Data Warehouse kopiert, und das gesamte Data Warehouse wird neu berechnet, unabhängig davon, ob sich die Daten geändert haben oder nicht.
  - Bei einem *Data Update*<sup>513</sup> werden die Daten in vorher definierten Zeitintervallen automatisiert in das Data Warehouse transportiert. Es werden nur die Daten übertragen, die sich seit dem letzten Update in der Quellumgebung geändert haben.

---

<sup>507</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 103.

<sup>508</sup> Vgl. ebd., S. 103.

<sup>509</sup> Vgl. ebd., S. 103.

<sup>510</sup> Vgl. ebd., S. 103.

<sup>511</sup> Vgl. ebd., S. 103.

<sup>512</sup> ist die einfachste Methode eines DW-Migrationsprozesses, aber für große Datenmengen ungeeignet. Diese Methode wird bestenfalls für das erste Laden des DW und bei nicht zeitkritischen Applikationen verwendet, vgl. ebd., S. 104.

<sup>513</sup> ist bei großen Datenmengen vorteilhaft, da in der Regel nur ein geringer Teil der Quelldaten übertragen werden muss und damit der Transport- und Ladeprozess weniger Zeit beansprucht, vgl. ebd., S. 104.

- Darüber hinaus kann der Integrator eine Verdichtung der Daten vornehmen. Dieser Vorgang wird als *Data Reduction* bezeichnet.

Ein weiterer Bestandteil der Staging Area ist ein Archivierungssystem, das die Bereiche Datensicherung und -archivierung abdeckt. Die Datensicherung dient der Wiederherstellung des Data Warehouse im Falle eines Programm- oder Systemfehlers.<sup>514</sup>

Ein Archivierungssystem im Data Warehouse-Konzept unterstützt das Ziel, durch eine effiziente Speicherung und Verarbeitung großer Datenmengen auch für komplexe ad hoc-Analysen kurze Antwortzeiten zu gewährleisten.<sup>515</sup> Die Notwendigkeit der Datenarchivierung ergibt sich aus dem Verdichtungsprozess der Daten im Data Warehouse. In Abhängigkeit der festgelegten Verdichtungsstufen und -frequenzen werden die Daten der untersten Detaillierungsstufen aus dem Data Warehouse ausgelagert und auf den Offline-Datenträgern archiviert.<sup>516</sup> Diese Reduzierung des Datenvolumens im Data Warehouse trägt zu einer Steigerung der Performance bei.<sup>517</sup>

Zur Staging Area ist noch abschließend anzumerken, dass die Gestaltung ihrer Prozesse die schwierigste und zeitintensivste Aufgabe in einem Data Warehouse-Projekt darstellt.

Die Datenablage in einem zentralen, relationalen Data Warehouse erweist sich als wenig effizient, wenn es darum geht, sehr umfangreiche Datenbestände rasch und flexibel zu analysieren. Häufig werden deshalb bestimmte Daten aus dem Datenbestand selektiert und zur weiteren Verarbeitung aufbereitet. Dadurch bildet man personen-, anwendungs-, funktionsbereichs- oder problemspezifische Teilmengen eines Data Warehouse-Datenbestandes, die auch als Data Marts bezeichnet werden.

Bei der Implementierung von Data Marts innerhalb einer Data Warehouse-Lösung können verschiedene Ansätze unterschieden werden:

- *Ansatz 1: Top down Implementation*

Die Einrichtung eines zentralen Data Warehouse kommt nur dann in Betracht, wenn die operationalen DV-Systeme eines Unternehmens durch einen zentralen DV-Bereich betrieben werden.<sup>518</sup> Durch den Aufbau eines zentralen Data

---

<sup>514</sup> Vgl. Muksch/Behme (1998), S. 61.

<sup>515</sup> Vgl. ebd., S. 61.

<sup>516</sup> Vgl. ebd., S. 61.

<sup>517</sup> Vgl. ebd., S. 61.

<sup>518</sup> Vgl. Muksch/Behme (1998), S. 69.

Warehouse können konsistente Datendefinitionen und unternehmensweit geltende Geschäftsregeln gewährleistet werden.<sup>519</sup> Die Data Marts in verschiedenen Fachabteilungen werden direkt aus dem Data Warehouse gespeist. Um dies verwirklichen zu können, wird zunächst das Data Warehouse implementiert und mit Daten gefüllt. Erst danach können die Data Marts aufgebaut werden. Aus Gründen der Performance und Übersichtlichkeit werden sie dann in die Fachabteilungen ausgelagert.<sup>520</sup>

Die Vorteile eines zentralen Data Warehouse sind der unkomplizierte Zugriff auf die Daten der im DW abgebildeten Unternehmensbereiche und die zentrale Ausführung von Auswertungen und Analysen.<sup>521</sup>

Dieser Ansatz ist jedoch nicht so erfolgreich, da eine zufriedenstellende unternehmensweite Datenmodellierung damit schwer zu realisieren ist.<sup>522</sup> Zudem ist es oft kaum möglich, die Veränderungen (Z. B. aufgrund geänderter Geschäftsprozesse) am Datenmodell vorzunehmen, ohne dabei das Gesamtmodell zu beeinträchtigen. Das Datenmodell muss dann in vielen Fällen komplett neu modelliert werden.

- *Ansatz 2: Bottom up Implementation*

Der Bottom Up-Ansatz ermöglicht die Planung und Inbetriebnahme von einzelnen Data Marts.<sup>523</sup> Dafür ist keine unternehmensweite DV-Infrastruktur vorausgesetzt, d. h. Data Marts können vorher oder zeitgleich zur Entwicklung eines unternehmensweiten Data Warehouse erstellt werden.

Im Data Warehouse-Konzept wird der Bottom Up-Ansatz oft bevorzugt, da er schnell und mit geringen Kosten realisierbar ist.<sup>524</sup> Es ist jedoch zu beachten, dass dieser Ansatz die sorgfältige Planung und Konzipierung der einzelnen Data Marts erfordert. Wenn sie nicht aufeinander abgestimmt sind, entstehen dann Insellösungen.<sup>525</sup>

Außerdem kann die unabhängige Entwicklung der Data Marts zu Problemen wie Inkonsistenzen und Datenredundanzen führen, die eine Zusammenführung der

---

<sup>519</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 105.

<sup>520</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 105.

<sup>521</sup> Vgl. Muksch/Behme (1998), S. 69.

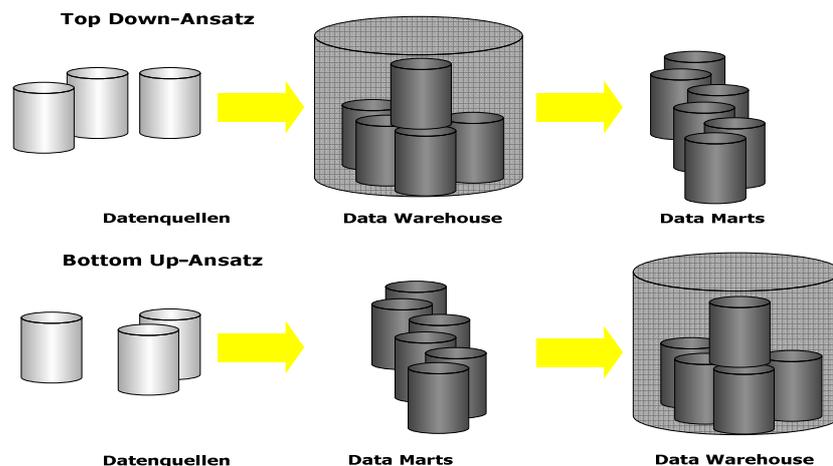
<sup>522</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 105.

<sup>523</sup> Vgl. ebd., S. 106.

<sup>524</sup> Vgl. ebd., S. 106.

<sup>525</sup> Vgl. ebd., S. 106.

Data Marts in ein unternehmensweites Data Warehouse verhindern.



**Abbildung 29: Top Down und Bottom Up Implementation im Vergleich**

Quelle: in Anlehnung an Meyer/Weingärtner/Döring (2001), S. 105-106

Beide Ansätze, sowohl der Top Down- als auch der Bottom Up-Ansatz alleine sind nicht so zufriedenstellend. Eine Mischung der beiden Ansätze als dritter Ansatz wäre vielleicht besser geeignet, um mit den sich ständig ändernden Anforderungen an die Datenhaltung innerhalb eines Unternehmens umzugehen.

- *Ansatz 3: Skalierbarer Data Mart*

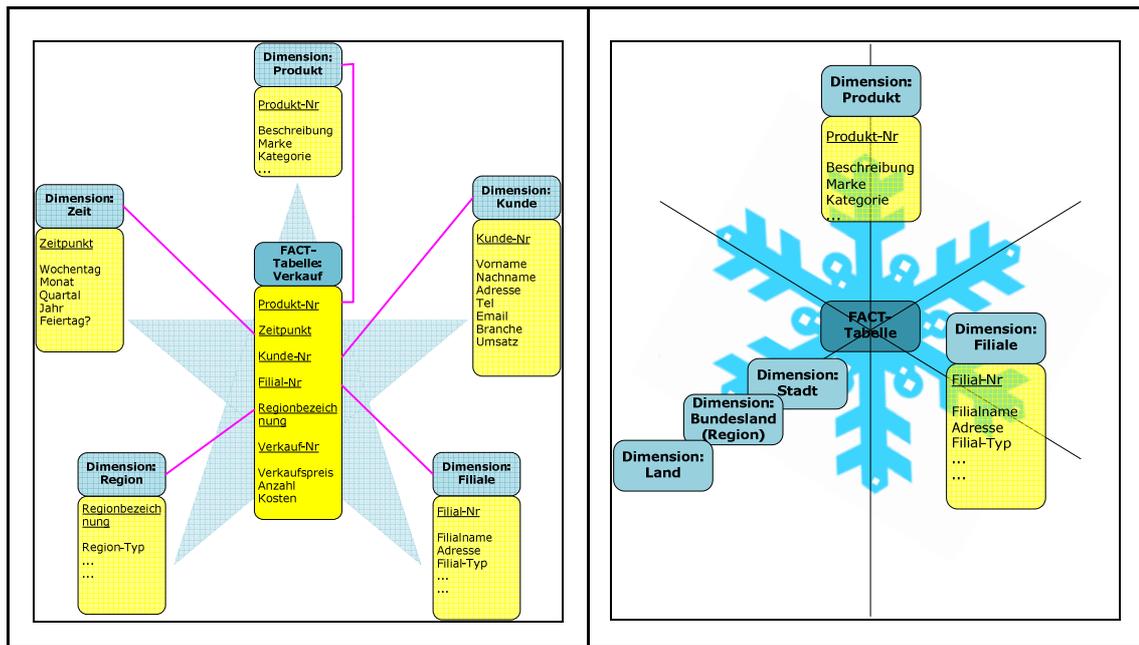
Ein Data Warehouse muss so gestaltet werden, dass es sich den veränderten Geschäftsprozessen und damit auch den veränderten Datenstrukturen anpassen kann. Somit ist es in der Lage, mit dem Unternehmen zu wachsen und sich zu verändern. Hier spielt die Skalierbarkeit des Data Warehouse eine wesentliche Rolle. Unter Skalierbarkeit versteht man die Fähigkeit eines Systems, inkrementell zu wachsen, wobei die einzelnen skalierbaren Komponenten sowohl Hard- und Softwaretechnologien als auch Entwurfs- und Implementierungstechniken umfassen.<sup>526</sup> Die skalierbaren Data Marts zielen auf eine Gesamtlösung im Sinne eines unternehmensweiten Data Warehouse ab, lassen jedoch eine schrittweise Implementierung einzelner Themenbereiche zu.<sup>527</sup>

Um in der Zukunft weitere Themenbereiche (Data Marts) zu einem ersten Themenbereich hinzufügen bzw. den ersten Themenbereich erweitern zu können, muss ein logisches Datenbank-Design erstellt werden, welches skizzenhaft alle möglichen Themenbereiche des Unternehmens enthält. Zur Abbildung und Modellierung einzelner Themenbereiche kann das Star- oder Snowflake-Schema

<sup>526</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 107.

<sup>527</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 107.

genutzt werden.



**Tabelle 5: Star- und Snowflake-Schemen**

Quelle: Hannig (2002), S. 12-13

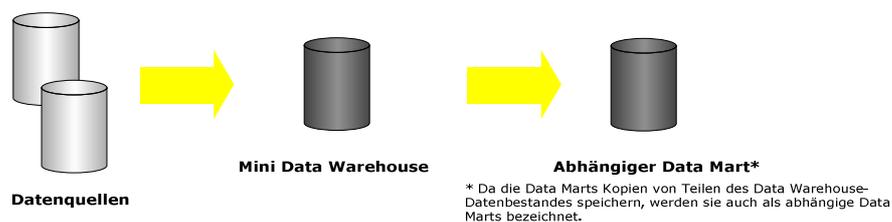
Der Name Star-Schema kommt von der sternförmigen Anordnung der Datentabellen her. Im Mittelpunkt eines Star-Schemas steht eine Faktentabelle, die mit mehreren Dimensionstabellen verbunden ist. Die Faktentabelle beschreibt ein standardisiertes Ereignis, wie z. B. den Verkauf von Waren in einer Filiale. Die Dimensionstabellen übernehmen dann die Aufgabe, die zu diesem Ereignis gehörenden Grunddaten näher zu beschreiben. Eine Dimensionstabelle beschreibt z. B. einen Kunden näher mit all seinen Kontaktinformationen, dem getätigten Umsatz usw. Eine andere Tabelle beschreibt das Produkt oder die Dienstleistung und eine weitere die Unternehmensfiliale. Ein solches Star-Schema ist nicht auf das anfängliche Einsatzgebiet beschränkt. Wenn eine gemeinsame Nutzung der Tabellen notwendig und sinnvoll erscheint, können auch andere bereits vorhandene oder neu hinzugefügte Sterne die Dimensionstabellen z. B. für Kunde und Produkt eines Sterns verwenden. So entsteht im Laufe der Zeit ein unternehmensweites Netz aus Data Mart-Sternen, denen jeweils eine Faktentabelle zugrunde liegt und die sich die Nutzung der Dimensionstabellen teilen.<sup>528</sup>

<sup>528</sup> Vgl. Schmidt/Grimm (2002), S. 249. Dieses Datenmodell wird dann auch als Galaxy-Schema bezeichnet. In einigen Fällen mag die Modellierung eines Star- oder Snowflake-Schemas mit einer Faktentabelle ausreichen. Die Geschäftssituation ist jedoch meistens komplexer, so dass mehrere Faktentabellen zur adäquaten Datenmodellierung notwendig sind. Im Gegensatz zu Star- und Snowflake-Schemen, bei denen eine einzige Faktentabelle die Primärschlüssel der Dimensionen als Fremdschlüssel enthält, können beim Galaxy-Schema mehrere Faktentabellen über gemeinsame Dimensionstabellen miteinander verknüpft werden. Diese Möglichkeit besteht immer dann, wenn die verschiedenen Faktendaten durch teils unterschiedliche, teils gleiche Dimensionen beschrieben sind, vgl. Holthuis (1998), S. 178f und Hannig (2002), S. 13.

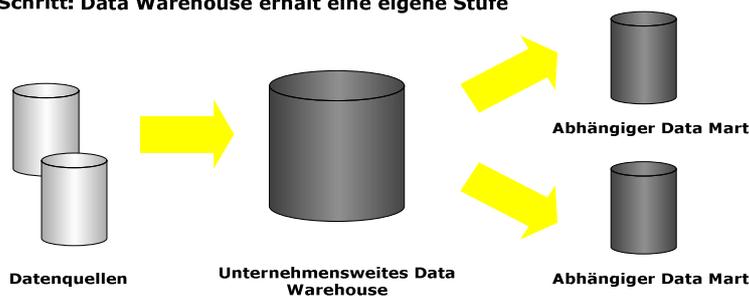
Beim Snowflake-Schema erinnert die Anordnung der Datentabellen an Schneeflocken. Die Faktentabellen bleiben im Zentrum unverändert, und es werden nur die Dimensionen um die Attribute erweitert. Die Tabelle 5 auf der vorigen Seite stellt beide Schemen graphisch dar.

Ein Data Warehouse kann mit Hilfe von skalierbaren Data Marts wie folgt implementiert werden:<sup>529</sup> Zuerst werden die Daten detailliert aus den operationalen Systemen für den ersten Themenbereich in das Data Warehouse geladen und von dort aggregiert in den angeschlossenen Data Mart weitergeleitet.

**1. Schritt: Data Mart und Data Warehouse - eine Plattform, eine Datenbank**



**2. Schritt: Data Warehouse erhält eine eigene Stufe**



**Abbildung 30: Von der zweistufigen zur dreistufigen Architektur**

Quelle: in Anlehnung an Meyer/Weingärtner/Döring (2001), S. 10

Das Data Warehouse enthält zu diesem Zeitpunkt nur Informationen zu einem einzigen Themenbereich. Das Data Warehouse und das Data Mart befinden sich auf derselben Plattform. Falls aber noch weitere Themenbereiche dem Data Warehouse hinzugefügt werden sollen, wächst das Mini-Data Warehouse und erhält eine eigene Stufe, eine größere und skalierbare Plattform. Die Konstruktion der abhängigen Data Marts verschafft dem Unternehmen eine gemeinsame Informationsbasis.

Die Suche und Zugriffe auf die im DW gespeicherten Daten erfolgt durch ein Meta-Datenbanksystem, in dem alle Informationen über die Data Warehouse-Komponenten gehalten und verwaltet werden. Das ist eine Art Hilfesystem bestehend aus einem Informationskatalog und einer Navigationshilfe.<sup>530</sup> Damit wird ein schnelles und

<sup>529</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 107.

<sup>530</sup> Vgl. Muksch/Behme (1998), S. 64.

sicheres Auffinden der benötigten Daten gewährleistet.<sup>531</sup>

Die Metadaten können als Daten über Daten bezeichnet werden, d. h. sie stellen eine Abstraktion von Daten dar.<sup>532</sup> Sie haben die Aufgabe, Auskunft über die Herkunft, Bedeutung, Transformation, Struktur, Qualität aller im Data Warehouse enthaltenen Informationen zu geben.<sup>533</sup> In diesem Sinne verfügen die Metadaten über alle Informationen bezüglich der Daten für die Generierung des Data Warehouse (Z. B. Datenquellen, Speicherort, Generierungszeitpunkt), der Kontrolldaten (Z. B. letzter Aktualisierungszeitpunkt, Gültigkeitsdauer, Zugriffsrechte) und der Anwendung für die Nutzung des Data Warehouse (Z. B. Datenstruktur im Data Warehouse, Semantik der Daten, logische Beziehungen zwischen den Daten).<sup>534</sup> Die aus den Transaktions- und Extraktionsprozessen der Staging Area resultierenden Regeln, Zuordnungen und Definitionen bilden die Grundlage für die Metadaten.<sup>535</sup>

Zum Abruf der Daten aus dem Data Warehouse existieren auf der Anwenderseite verschiedene Werkzeuge und Verfahren. Diese reichen von einfachen Standard Reporting Tools und Tabellenprogrammen über Werkzeuge für komplexe ad hoc-Abfragen bis OLAP- und Data Mining-Anwendungen.

#### **4.6.2 Database Marketing im CRM**

Mit Hilfe eines Data Warehouse, das im vorigen Kapitel bereits erläutert wurde, wird im Rahmen des Customer Relationship Management die zielgerichtete Nutzung aller kundenbezogenen und sonstigen Daten ermöglicht, die auch als Database Marketing bezeichnet werden kann.

Das Database Marketing ist eine Art dialogorientiertes Marketing, welches sich auf die aktuellen, kundenindividuellen und in einer Kundendatenbank gespeicherten Informationen bezieht und sich durch folgende Merkmale charakterisieren lässt:<sup>536</sup>

- Der Einsatz eines Database Marketing-Systems dient zur Erfassung, Bearbeitung und Bereitstellung von Informationen über Zielpersonen als Basis für das Management der Kundenbeziehung.
  
- Das Database Marketing ermöglicht die Nutzung der Daten und Informationen

---

<sup>531</sup> Vgl. Holthuis (1998), S. 95.

<sup>532</sup> Vgl. Holthuis (1998), S. 95.

<sup>533</sup> Vgl. Wieken (1998), S. 278.

<sup>534</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 108.

<sup>535</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 109.

<sup>536</sup> Vgl. Mentzl/Ludwig (1998), S. 479 und Holland/Huldi/Kuhfuß/Nitsche (2001), S. 62.

für den Einsatz der Instrumente und Kanäle des CRM.

- Mit dem Database Marketing kann ein langfristiger, interaktiver und individueller Dialog mit dem Kunden aufgebaut werden. Dadurch ist eine Kundenbetreuung zu geringen Kosten möglich.
- Das Database Marketing dient zur Ermittlung der individuellen Kundenbedürfnisse und ermöglicht somit die Ansprache des Kunden zur richtigen Zeit mit einem auf seine Bedürfnisse abgestimmten Informations- und Leistungsangebot.
- Das Database Marketing hat einen Systemcharakter, der durch die permanente Sammlung kommunikationsnotwendiger Daten und die langfristige Perspektive beschrieben wird.

Aus diesen Merkmalen lässt sich leicht schließen, dass der Einsatz des Database Marketing auf die Erstellung eines möglichst engen Kundenkontaktes abzielt.<sup>537</sup> Dadurch wird der Aufbau einer dauerhaften profitablen Kundenbeziehung möglich.

Das Database Marketing deckt nahezu alle Aufgaben des Direktmarketing von der Neukundengewinnung bis hin zur Erfolgskontrolle ab.<sup>538</sup> Die Bewertung der Kunden oder Interessenten auf der Basis einer Kundendatenbank bildet die Grundlage für den Einsatz des zielgerichteten Direktmarketing und damit auch des Database Marketing.<sup>539</sup>

Die stetig steigende Kundenzahl und die differenzierte Kundenansprache machten es im Rahmen des Customer Relationship Management notwendig, verschiedene Modelle und Verfahren zu entwickeln, die in der Lage sind, den einzelnen Kunden möglichst objektiv zu bewerten und darauf aufbauend maßgeschneiderte Konzepte für diesen Kunden anzubieten.<sup>540</sup> Zu diesen Kundenbewertungsverfahren bzw. -modellen zählen die ABC-Analyse, die RFMR- und FRAT-Methode, das Kundenscoring, der Customer Lifetime Value und die Kundenportfolios.

Im Folgenden werden diese einzeln angeführt und kurz beschrieben:

- *ABC-Analyse*: Bei der ABC-Analyse werden die Kunden nach ihrem Umsatz oder Deckungsbeitrag in drei oder auch mehrere Gruppen eingeteilt. Aus der

---

<sup>537</sup> Vgl. Holland (2004a), S. 120.

<sup>538</sup> Vgl. ebd., S. 107.

<sup>539</sup> Vgl. ebd., S. 107.

<sup>540</sup> Vgl. ebd., S. 107.

Analyse ergibt sich das Verhältnis zwischen dem Kaufanteil und dem Wert der Kunden. Bei vielen Unternehmen entspricht dieses Verhältnis oft dem Pareto-Prinzip<sup>541</sup>. Dieses Prinzip besagt, dass 20 Prozent der Kunden 80 Prozent des Umsatzes oder Deckungsbeitrags erwirtschaften. Für die unterschiedlichen Kundengruppen können dann differenzierte Marketingmaßnahmen im Rahmen des CRM angeordnet werden.

- *RFMR- und FRAT-Methode*: Das ist ein Scoring-Modell, welches zur Auswertung der in den Kundendatenbanken gespeicherten Daten dient. Dieses Modell basiert auf einem Punktbewertungssystem, in das eine Vielzahl von Kriterien eingeht, die vom Unternehmen jeweils nach ihren Aussagewerten bezüglich der Vorhersage des weiteren Verlaufs einer Kundenbeziehung ausgewählt werden. So werden für jeden einzelnen Kunden, ausgehend von den Kaufdaten Recency (letztes Kaufdatum), Frequency (Kaufhäufigkeit) und Monetary Ratio (Umsatzhöhe), Punkte vergeben und diese dann zusammenaddiert.
  - R = Recency → Je weniger der letzte Kauf des Kunden zurückliegt, desto mehr Punkte bekommt er.
  - F = Frequency → Je öfter der Kunde im Verlauf der Geschäftsbeziehung oder im Laufe des letzten Jahres gekauft hat, desto größer ist die Wahrscheinlichkeit, dass er wieder kauft, und damit der Punktwert, der ihm zugeordnet wird.
  - MR = Monetary Ratio → Der kumulierte Umsatz oder Bestellwert des Kunden wird in einen Punktwert umgerechnet und zu den aus Recency und Frequency errechneten Punkten addiert.

Die FRAT-Methode bezieht neben den oben genannten Kriterien auch die Produktsortimente, aus denen bestellt wurde, in die Analyse ein.

- (F) Frequency bekommt in diesem Modell das höchste Gewicht.
- (R) Recency wird auch hier einbezogen.
- (A) Amount of Purchase entspricht der Höhe des Umsatzes oder Bestellwertes.

---

<sup>541</sup> Entsprechend einer Pareto-Verteilung sind A-Kunden jene, die zahlenmäßig die höchsten Umsätze erbringen (z. B. 20% der Kunden erbringen 80% des Umsatzes). Als B-Kunden werden diejenigen bezeichnet, die geringere Umsätze erbringen, jedoch zahlenmäßig größer sind (z.B. 30% der Kunden erbringen 15% des Umsatzes). Die Kundengruppe C ist die zahlenmäßig größte Gruppe, erbringt jedoch den geringsten Umsatzanteil (z. B. 50% der Kunden erbringen nur 5% des Umsatzes).

- Zusätzlich besitzt diese Methode einen durch das (T) symbolisierten Sortimentsbereich oder die Warenart, aus der gekauft wurde.

Hier bietet sich der Einsatz von Fuzzy-Logik besonders an.<sup>542</sup> Die neuesten Scoring-Modelle können beliebig viele Kriterien für die Bewertung berücksichtigen und diese auf der Basis eines Fuzzy-Logik-Ansatzes intelligent verarbeiten.

- *Kundenscoring*: Beim Kundenscoring werden dynamische Faktoren zur Definition eines Kundenwertes eingesetzt. Aus den Kundeninformationen in der Datenbank werden die Segmentierungskriterien ausgewählt, die für die entsprechende Fragestellung relevant sind. Anders als bei der RFMR-Methode, die fest definierte Bestimmungsgrößen verwendet, kann somit flexibler auf die spezifischen Probleme reagiert werden.<sup>543</sup> Die entsprechend ausgewählten Faktoren werden dann bewertet, und jedem Kunden wird ein Wert zugeordnet. Um diese relevanten Faktoren aus der Datenbank herauszufiltern und zu bewerten, werden Data Mining-Methoden eingesetzt.
- *Customer Lifetime Value*: Der Customer Lifetime Value geht davon aus, dass ein Kunde im Laufe seines Lebens eine Vielzahl von Einkäufen bei einem Unternehmen tätigt, und stellt diesen lebenslangen Wert in den Vordergrund der Betrachtung.<sup>544</sup> Dementsprechend wird der Kunde nicht anhand des Ertrags aus einem einzigen Kauf bewertet, sondern auf der Basis des insgesamt möglichen Umsatzes. Dieser Wert bezieht sowohl die schon getätigten Umsätze als auch die zukünftigen potentiellen Umsätze ein, so dass das Chancenpotential einer dauerhaften Bindung des Kunden abgeschätzt werden kann.<sup>545</sup> Der abgezinste Kapitalwert einer Kundenbeziehung wird mittels der dynamischen Investitionsrechnung berechnet, indem man alle Einnahmen und Ausgaben in die dynamische Berechnung mit einbezieht und diese auf einen einheitlichen Bezugszeitpunkt abzinst.<sup>546</sup>
- *Kundenportfolios*: Mit Hilfe eines Kundenportfolios kann eine Klassifizierung der Kunden vorgenommen werden, indem zwei Segmentierungskriterien zueinander ins Verhältnis gesetzt werden.<sup>547</sup> Durch die Einordnung der

---

<sup>542</sup> Vgl. Holland (2004a), S. 109.

<sup>543</sup> Vgl. ebd., S. 111.

<sup>544</sup> Vgl. ebd., S. 111.

<sup>545</sup> Vgl. ebd., S. 111.

<sup>546</sup> Vgl. ebd., S. 111.

<sup>547</sup> Vgl. ebd., S. 111.

Kundengruppen in dieses Koordinatensystem wird die Position des jeweiligen Kunden visuell dargestellt. So kann man z. B. leicht herausfinden, welche Kommunikationsstrategie dem jeweiligen Kunden passt, wenn die Kriterien Kundenpotential und Kundenloyalität zueinander in Beziehung gesetzt werden.<sup>548</sup> Mit Hilfe dieser Kundenpositionen im Portfolio werden die optimalen Maßnahmen für jeden einzelnen Kunden bestimmt.

Die Datenbank bildet die Informationsgrundlage für die computergestützten Marketingentscheidungen. Nahezu alle Unternehmen verfügen heute über eine Datenbank, um ihre Kunden und deren Kaufverhalten individuell bewerten zu können. Nur mit einer Kundendatenbank ist es möglich, die Kunden zu selektieren und nach bestimmten Kriterien Gruppen zuzuordnen.<sup>549</sup> Anhand der Selektion können dann die inaktiven und unrentablen Kunden identifiziert werden.

Das Database Marketing ermöglicht neben dieser Selektion von Kunden auch deren Segmentierung für bestimmte Marketing-Aktivitäten.<sup>550</sup> So lassen sich aus der Kundendatenbank aufgrund der gespeicherten Informationen beispielsweise jene Personensegmente herausfiltern, von denen man ein gesteigertes Interesse für bestimmte Angebote erwarten kann.

Ein weiterer Vorteil der Kundendatenbank liegt darin, dass die Kundenbefragungen damit effizient durchgeführt werden.<sup>551</sup> Da die Datenbank gestützte Befragung mit bekannten und damit nicht mit anonymen Kunden erfolgt, ergibt sich die Möglichkeit, nicht nur den prozentmäßigen Anteil der zufriedengestellten Kunden festzustellen, sondern auch die Einstellungen und Meinungen der einzelnen Kunden beispielsweise in Bezug auf einen bestimmten Geschäftsablauf zu erfahren. Ferner können aus den gespeicherten Kundendaten zusätzliche Informationen gewonnen werden. Durch statische Auswertungen lassen sich aus diesen Daten wertvolle Informationen über das jeweilige Kaufverhalten der Kunden ableiten.

Ein Database Marketing-System besteht im engeren Sinne aus einer Marketing-Datenbank und den vier Prozessstufen Analyse, Planung, Durchführung und Reaktion, die die Bestandteile des Direktmarketing-Regelkreises verkörpern (siehe Abb. 31).<sup>552</sup> Auf Basis einer solchen Marketing-Datenbank unterstützt das Database Marketing-System eine Analyse der Daten in der Datenbasis, die Planung konkreter Maßnahmen

---

<sup>548</sup> Vgl. Holland (2004a), S. 112.

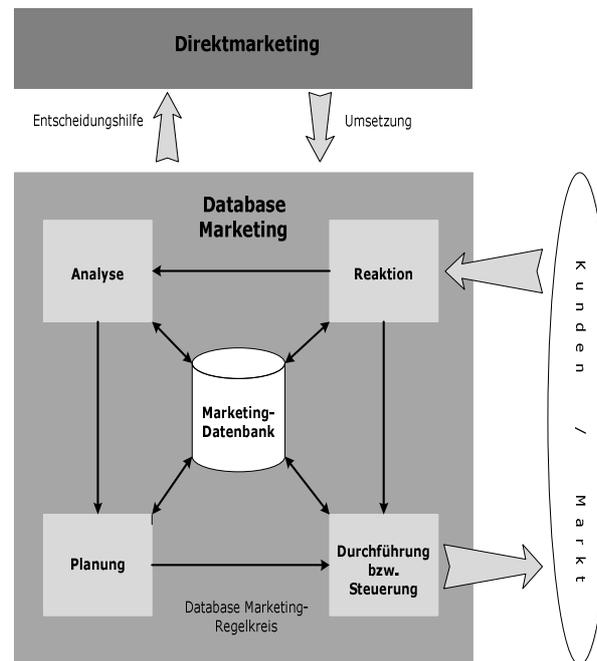
<sup>549</sup> Vgl. ebd., S. 113.

<sup>550</sup> Vgl. ebd., S. 114.

<sup>551</sup> Vgl. ebd., S. 115.

<sup>552</sup> Vgl. Mentzl/Ludwig (1998), S. 481.

und eine individuelle als auch koordinierende Ansprache bestehender oder potentieller Kunden sowie die Erfassung der Kundenreaktionen und deren Auswertung.



**Abbildung 31: Bestandteile des Database Marketing im Regelkreisprinzip**

Quelle: in Anlehnung an Mentzl/Ludwig(1998), S. 481

Die Implementierung des Database Marketing-Systems soll jedoch nicht nur auf die reine Erstellung einer Marketing-Datenbank eingeschränkt werden.<sup>553</sup> Ein Database Marketing-System kann seinen wahren Nutzen für das Unternehmen nur dann erbringen, wenn alle Prozessstufen des Marketing-Regelkreises optimal aufeinander abgestimmt sind. Außerdem soll es unternehmensübergreifend eingesetzt werden.

Die Abbildung 32 stellt ein logisches Modell für die unternehmensübergreifende Implementierung eines Database Marketing-Systems dar.

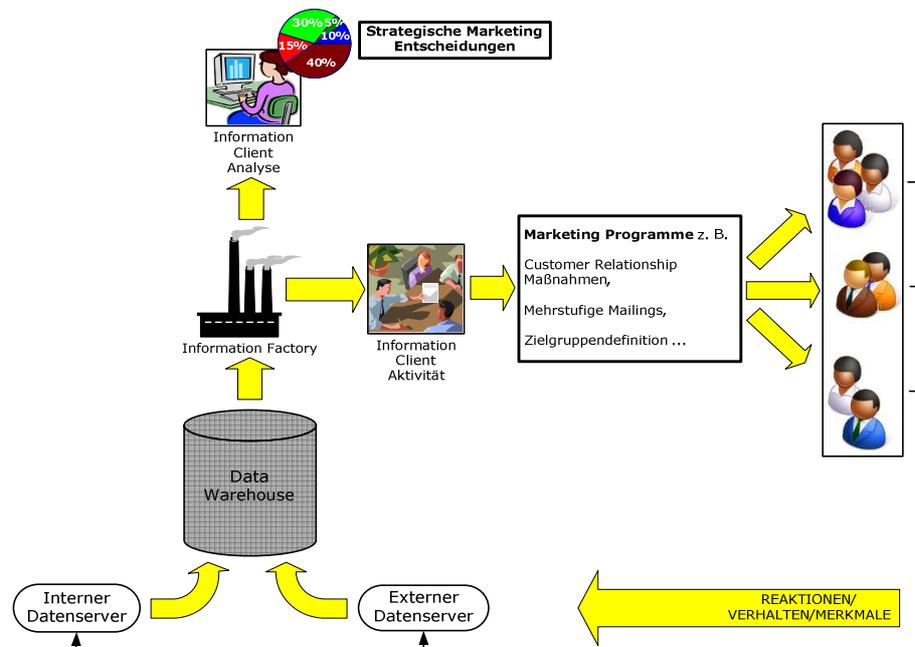
Im Mittelpunkt des Modells befindet sich das Data Warehouse. Das Data Warehouse wird nicht ausschließlich aus den Daten in der Marketing-Datenbank, sondern aus allen relevanten unternehmensinternen und -externen Daten gespeist.<sup>554</sup> Die Zusammenführung dieser Daten im DW ermöglicht die gesamtheitliche Betrachtung der Kunden zu Analyse Zwecken. In der Information Factory werden die im DW gespeicherten Daten für die multidimensionale Betrachtung und marketingspezifischen Auswertungen aufbereitet. In der Information Client Analyse werden mit Hilfe des durch Decision Support Systeme (DSS), OLAP- und Data Mining-Tools erzeugten

<sup>553</sup> Vgl. Mentzl/Ludwig (1998), S. 483.

<sup>554</sup> Vgl. Mentzl/Ludwig (1998), S. 486.

Wissens strategische Marketingentscheidungen getroffen.

Für die Steuerung der in der Information Client-Analyse festgelegten Marketingprogramme ist der Bereich Information Client-Aktivität zuständig. Die Information Client-Aktivitäten richten sich direkt auf den Dialog mit dem bestehenden oder potentiellen Kunden aus.<sup>555</sup>



**Abbildung 32: Das logische Modell eines Database Marketing-Systems**

Quelle: in Anlehnung an Mentzl/Ludwig(1998), S. 487.

Das Verhalten von Kunden in Bezug auf diese Aktivitäten wird dann anhand von Responseelementen erfasst und als Reaktion in das Data Warehouse zurückgesendet.

Die empfangenen Kundenreaktionen werden in der Information Client-Analyse einer Auswertung unterzogen, was folglich zu einer besseren Selektionsgrundlage für die nächsten Marketing-Aktionen und einem Lernprozess führt.<sup>556</sup> Schließlich führt die Information Client-Aktivität den nächsten Schritt in einer vorab festgelegten Marketing-Aktion durch (Z. B. eine Nachfassaktion).<sup>557</sup>

Ein fortlaufender Dialog mit den Kunden macht es möglich, den Ablauf zu wiederholen, bis die Ansprache durch die Auswertung der Maßnahmen in der Information Client-Analyse und durch deren Steuerung in der Information Client-Aktivität an die individuellen Wünsche und Bedürfnisse der Kunden angepasst wird.

<sup>555</sup> Vgl. Holland (2004a), S. 121.

<sup>556</sup> Vgl. ebd., S. 121.

<sup>557</sup> Vgl. ebd., S. 121.

# KAPITEL 5

---

## DER EINFLUSS VON DATA MINING AUF CRM

### 5.1 EINFÜHRUNG

*„Wir ertrinken in Informationen - aber wir hungern nach Wissen.“*

*[Rutherford D. Rogers, Bibliotheksvorstand, Yale*

*The New York Times, 25.02.1988]*

*„The real voyage of discovery consists not in seeking new landscapes but in having new eyes.“*

*[Marcel Proust, französischer Schriftsteller]*

Die ständig wachsende Menge der in einem Unternehmen gespeicherten Daten und die mit der Erzielung und Erhaltung von Wettbewerbsvorteilen für ein Unternehmen verbundene Notwendigkeit einer schnelleren Analyse des gespeicherten Datenbestands haben dazu geführt, dass neuere Ansätze der Datenanalyse entwickelt werden mussten.

Heutzutage stehen den Unternehmen mehr Informationen und Daten über die Kunden zur Verfügung als jemals zuvor: „Alle 20 Monate verdoppelt sich die Menge der Informationen auf der Welt.“<sup>558</sup> Aufgrund der entstandenen riesigen Datenvolumina sind konventionelle bzw. manuelle Methoden der Datenanalyse meist überfordert.<sup>559</sup> Man braucht deswegen eine neue Generation von Methoden und Tools, die „[...] assist humans in extracting useful information (knowledge) from the rapidly growing volumes of digital data.“<sup>560</sup>

---

<sup>558</sup> Mertens/Wieczorrek (2000), S. 211.

<sup>559</sup> Vgl. Mertens/Wieczorrek (2000), S. 211.

<sup>560</sup> Fayyad/Shapiro/Smyth (1996b), S. 37.

Solche Methoden und Tools sind der Gegenstand von „Knowledge Discovery in Databases (KDD)“<sup>561</sup>, einem Ansatz der Datenanalyse, der darauf abzielt, in umfangreichen Datenbeständen implizit vorhandenes Wissen zu entdecken und explizit zu machen.<sup>562</sup> In diesem Kapitel werden zunächst die Grundlagen des Knowledge Discovery in Databases dargestellt und seine Bedeutung für CRM erläutert.

## **5.2 WAS BEDEUTET KNOWLEDGE DISCOVERY IN DATABASES?**

Die Bezeichnung Knowledge Discovery in Databases beschreibt „[...] the nontrivial extraction of implicit, previously unknown, and potentially useful information from data.“<sup>563</sup>

Diese ursprüngliche Bedeutung von KDD hat sich aber im zeitlichen Verlauf verändert und umfasst heute zwei Auslegungen:<sup>564</sup> Die eine ist durch die Weiterentwicklung der Bedeutung im Hinblick auf die Forderung nach einem Automatismus des Knowledge Discovery in Databases bestimmt. Demzufolge setzt die Durchführung des Knowledge Discovery in Databases gar keine oder nur eine geringe Partizipation des Anwenders voraus. Die andere Auslegung betont die Prozesseigenschaft des Knowledge Discovery in Databases. KDD umfasst somit eine komplexe, inhaltlich abgeschlossene, zeitliche und sachlogische Abfolge von Tätigkeiten.

Diese beiden Auslegungen basieren jedoch auf einer inhaltlichen Unterscheidung der Begriffe Knowledge Discovery in Databases und Data Mining, die von den Autoren Fayyad, Shapiro und Smyth thematisiert wurde.

Sie definieren Knowledge Discovery in Databases als, „[...] the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data.“<sup>565</sup> KDD ist demnach darauf ausgerichtet, im Rahmen eines nichttrivialen Prozesses („nontrivial process“) Beziehungsmuster („patterns“), wie z. B. Regelmäßigkeiten oder Auffälligkeiten, in den Daten zu ermitteln.<sup>566</sup>

Diese Beziehungsmuster sollen für einen möglichst großen Anteil des Datenbestandes Gültigkeit haben („valid“) und bislang unbekannte („novel“), potentiell nützliche („potentially useful“) und leicht verständliche (ultimately understandable“)

---

<sup>561</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 37.

<sup>562</sup> Vgl. Düsing (2006), S. 242.

<sup>563</sup> Frawley/Shapiro/Matheus (1992), S. 58.

<sup>564</sup> Vgl. Düsing (2006), S. 242.

<sup>565</sup> Fayyad/Shapiro/Smyth (1996b), S. 40f.

<sup>566</sup> Vgl. Düsing (2006), S. 243.

Zusammenhänge in den Daten artikulieren.<sup>567</sup>

Aus den ermittelten Mustern wird schließlich durch Interpretation explizites Wissen abgeleitet, das zur Unterstützung von Entscheidungen herangezogen werden kann.<sup>568</sup>

Dieses explizit gemachte Wissen wird im Rahmen des Customer Relationship Management zum besseren Kundenverständnis genutzt.

Knowledge Discovery in Databases ist ein iterativ und interaktiv ablaufender Prozess der Wissensentdeckung in Datenbeständen, der die Auswahl, Aufbereitung, Analyse und Interpretation der Daten umfasst.<sup>569</sup> Im Gegensatz dazu bildet Data Mining nur einen kleinen Schritt im KDD-Prozess (siehe auch Kapitel 3.6.2.2).

Fayyad, Shapiro und Smyth definieren Data Mining wie folgt: „Data Mining is a step in the KDD process that consists of applying data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns (or models) over the data.“<sup>570</sup> Diese Definition stellt klar, dass Data Mining als eine Komponente des KDD-Prozesses nur die algorithmische Extraktion und Enumeration von Mustern aus den Daten mit Hilfe von spezifischen Suchalgorithmen beinhaltet. Dagegen bezieht sich KDD auf den gesamten, iterativ und interaktiv ablaufenden Prozess der Wissensentdeckung, der die Auswahl, Aufbereitung und Anwendung von Data Mining-Methoden zur Analyse der Daten und die Interpretation von aus Daten extrahierten Mustern umfasst. Die Interpretation dient schließlich zur Entscheidung, ob die gefundenen Muster als neues Knowledge betrachtet werden können.<sup>571</sup>

Trotz dieser ursprünglichen inhaltlichen Unterscheidung zwischen Knowledge Discovery in Databases und Data Mining werden heute beide Begriffe oft synonym verwendet.<sup>572</sup> Im Hinblick auf KDD wird Data Mining im weiten Sinne als ein vollständiger und eigenständiger Prozess zugrunde gelegt. Im engeren Sinne kann es jedoch als die effiziente Suche nach versteckten, aber potentiell nützlichen Informationen in großen Datenbeständen verstanden werden.<sup>573</sup>

---

<sup>567</sup> Vgl. Düsing (2006), S. 243.

<sup>568</sup> Vgl. Fayyad/Shapiro/Smyth (1996c), S. 31 und Düsing (2006), S. 243.

<sup>569</sup> Vgl. Düsing (2006), S. 246 und Fayyad/Shapiro/Smyth (1996b), S. 42.

<sup>570</sup> Fayyad/Shapiro/Smyth (1996b), S. 41.

<sup>571</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 42.

<sup>572</sup> Vgl. Düsing (2006), S. 242.

<sup>573</sup> Vgl. Mertens/Wieczerrek (2000), S. 211.

### 5.2.1 Evolution des Data Mining

Der zeitliche Ursprung des Knowledge Discovery in Databases als Forschungsgebiet ist das Ende der 80er Jahre.<sup>574</sup> Durch die zunehmende Verfügbarkeit leistungsfähiger DV-Systeme sowie ein vermehrtes Angebot von Analysewerkzeugen hat das Interesse an der Wissensextraktion aus großen Datenbeständen ständig zugenommen. Die Data Mining-Ansätze zur analytischen Auswertung großer Datenmengen sind jedoch erst in jüngster Zeit in den Mittelpunkt des praktischen Interesses gerückt.<sup>575</sup> Deren Einsatz wurde durch folgende Entwicklungen unterstützt:<sup>576</sup>

- Die Menge der elektronisch gespeicherten Informationen wächst exponentiell. Die Etablierung des Data Warehouse in vielen Unternehmen hat jedoch die Durchführung von komplexen Analysevorgängen erleichtert. Die Data Mining-Software ist in zunehmendem Maße in der Lage, sehr große Datenbestände zu verarbeiten.
- Alle 18 Monate verdoppelt sich die Leistung der Rechnersysteme, wodurch es möglich ist, den hohen Anforderungen an rechenintensive Analyseprozesse gerecht zu werden. Außerdem sind die Hardwarepreise kontinuierlich gesunken.
- Die Suche nach Erkenntnissen in den eigenen, vorhandenen Daten ist nach der Ausschöpfung der bisherigen Möglichkeiten zur Kostenreduzierung und Ergebnisverbesserung der letzte Schritt in der Reihe der möglichen Maßnahmen zur Steigerung der Ertragspotentiale des Unternehmens.
- Die Analyseergebnisse können so aufbereitet werden, dass sie auch für Endanwender in den Fachbereichen verständlich sind.

Verfolgt man die Entwicklungsschritte von der einfachen Datensammlung in den Anfängen des Computerzeitalters bis zum heutigen Data Mining, erkennt man gleich, dass im letzten Evolutionsschritt die frühere retrospektivische Datenbeschaffung und –navigation durch eine prospektivische und proaktive Informationsbeschaffung ersetzt wurde (siehe Tab. 6).

<i>Evolutionstufe</i>	<i>Geschäftsfragen</i>	<i>Technologie</i>	<i>Eigenschaften</i>
<b>Data Collection</b> (60er Jahre)	Was war der Durchschnitt aller Einkünfte der letzten fünf Jahre?	Computer, Bänder, Festplatten	retrospektivische, statische Datenbeschaffung

<sup>574</sup> Vgl. Düsing (2006), S. 242.

<sup>575</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 25.

<sup>576</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 25.

<b>Data Access</b> (80er Jahre)	Was waren im letzten März in New England die Verkaufszahlen pro Einheit?	Relationale Datenbanken, Structured Query Language (SQL), ODBC	retrospektivische, dynamische Datenbeschaffung auf Datensatzbasis
<b>Data Navigation</b> (90er Jahre)	Was waren im letzten März in New England die Verkaufszahlen pro Einheit? Verfeinern bis Boston.	Online Analytic Processing (OLAP), multidimensionale Datenbanken, Data Warehouse	retrospektivische, dynamische Datenbeschaffung auf verschiedenen Ebenen
<b>Data Mining</b> (2000er Jahre)	Was für Verkaufszahlen werden wir voraussichtlich nächsten Monat in Boston für diese Einheit haben? Und warum?	Erweiterte Algorithmen, Mehrprozessor-Rechner, parallele Datenbanken	prospektivische, proaktive Informationsbeschaffung

**Tabelle 6: Evolutionsschritte des Data Mining**

Quelle: Krahl/Windheuser/Zick (1998), S. 25f.

## 5.2.2 Data Mining und herkömmliche Statistik - Parellellitäten und Unterschiede

Theoretisch geht es sowohl im Data Mining als auch in der Statistik darum, Daten zu analysieren und Wissen basierend auf dieser Analyse zu kreieren. Gerade von diesem Blickwinkel her kann Data Mining als eine Bereicherung der Statistik<sup>577</sup> angesehen werden. Es verfügt über sämtliche Analysemethoden und –werkzeugen, mit denen die Methoden der Statistik – als der klassische Ansatz zur Wissensgenerierung aus Daten – ergänzt und erweitert (Z. B. Neuronale Netze, Entscheidungsbäume) werden.

Zudem sehen Fayyad, Shapiro und Smyth den Wissenentdeckungsprozess (KDD) als eine ausgedehntere bzw. weitere Sicht der Modellierung in Statistik, welcher Werkzeuge für die Automatisierung des gesamten Prozesses der Datenanalyse zur Verfügung stellt und die statistische Kunst der Hypothesenselektion in sich behält.<sup>578</sup>

Jedoch gibt es wesentliche Unterschiede zwischen den Methoden und Verfahren beider Disziplinen, die hier zum besseren Verständnis erläutert werden sollen:

Während Verfahren des Data Mining typischerweise auf problembezogenen heuristischen Überlegungen basieren, setzen Verfahren der klassischen Statistik auf die traditionelle stochastische Modellbildung, die eine theoretische Analyse der Güte und Robustheit von Verfahren auf der Basis des mathematischen Hintergrunds der Statistik ermöglicht.

<sup>577</sup> Manche sehen Data Mining als eine Teilmenge von Statistik. Diese bislang allgemein akzeptierte Ansicht entspricht aber nicht der Realität, da Data Mining außer Theorien von Statistik auch Ideen, Werkzeuge und Methoden von anderen wissenschaftlichen Disziplinen und Gebieten verwendet.

<sup>578</sup> Vgl. (eigene Übersetzung) Fayyad/Shapiro/Smyth (1996c), S. 29 und (1996b), S. 40.

Statistik hat eine gewisse Bescheidenheit und neigt dazu, ad hoc-Handlungen zu vermeiden und zieht somit die Genauigkeit vor. Vor diesem Hintergrund gehen viele Statistiker mit der Analyse von großen Datenmengen sehr vorsichtig um, da es ihnen gewiss ist, dass nach scheinbaren Strukturen in großen Datenmengen zu suchen - wie es in Data Mining impliziert wurde - sehr aufwendig wird, insbesondere wenn die Suche rein auf Zufall basiert.<sup>579</sup> Der mathematische Background und die Hervorhebung der Genauigkeit von Statistik motivierte die Notwendigkeit, nachzuweisen, dass eine vorgeschlagene Methode funktionieren wird, schon im Vorfeld von deren tatsächlicher Anwendung.<sup>580</sup>

Im Gegensatz dazu hat Data Mining die gewagte Haltung von seinem protagonistischen Vorgänger, nämlich dem Maschinellen Lernen geerbt. Beim Data Mining wird vor der Lösung einer Problemstellung das geeignete, an die Problemstellung angepasste heuristische Verfahren bestimmt, dessen Güte und Robustheit aber vom jeweils zu lösenden Problem abhängt. Heuristische Verfahren bzw. Heuristiken sind als Arbeitsweisen zu verstehen, die geeignet sind, für ein bestimmtes Problem eine zumindest hinreichend gute oder plausible Lösung mit geringem Rechenaufwand und kurzer Laufzeit zu finden, ohne deren Optimalität zu gewährleisten.<sup>581</sup>

Während klassische Verfahren einerseits die optimale Rechenzeit und andererseits die optimale Lösung zu garantieren versuchen, begnügen sich Heuristiken mit dem Auffinden einer als hinreichend gut eingeschätzten Lösung, um bei sehr komplexen Aufgaben einen Kompromiss zwischen dem Rechenaufwand und der Güte der gefundenen Lösung einzugehen. Dies bedeutet aber nicht, dass Data Mining Anwender keinen Wert auf die Genauigkeit legen, sondern impliziert nur, dass sie bereit sind, auf die Genauigkeit zu verzichten, wenn somit auf die Ergebnisse gelangt werden können.<sup>582</sup> Die Anwender von Data Mining sind viel mehr an der Verständlichkeit interessiert als an der Exaktheit oder Vorhersagbarkeit.<sup>583</sup> Weiters sind die Recheneffizienz und Skalierbarkeit sehr wichtig<sup>584</sup>, da es in Data Mining üblicherweise vorkommt, dass Anwendungen eine sehr große Anzahl von unterschiedlichen Variablen

---

<sup>579</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 16.

<sup>580</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 16.

<sup>581</sup> In der Praxis sind viele Probleme in ihrer Struktur sehr kompliziert, so dass die optimale Lösung gar nicht oder nur mit unverhältnismäßig hohem Aufwand ermittelt werden kann. In solchen Fällen werden die exakt optimierenden, mathematisch-formalen Ansätze aber zu aufwendig und durch Näherungsmethoden, Plausibilitätsansätzen oder heuristische Verfahren, ersetzt.

<sup>582</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 16.

<sup>583</sup> Vgl. (eigene Übersetzung) Gilmour/Madigan/Pregibon/Smyth (1996), S. 25.

<sup>584</sup> Vgl. (eigene Übersetzung) Gilmour/Madigan/Pregibon/Smyth (1996), S. 26.

und Messungen einschließen.<sup>585</sup> Die Aspekte der statistischen Konsistenz dürfen hier sekundär (zweitrangig) berücksichtigt werden.<sup>586</sup>

Bei den Methoden der Statistik steht das Modell im Vordergrund.<sup>587</sup> Man bediente sich in der Vergangenheit starre, gering parametrisierte Modelle, die zwar eine einfache Berechenbarkeit und theoretische Behandlung auch komplexer Fragestellungen gewährleisten, sich jedoch oft als nicht robust gegenüber Verletzungen der Modellannahmen erwiesen und daher für die Praxis eher ungeeignet waren.<sup>588</sup> Auch die multivariate Statistik<sup>589</sup>, welche sich durch flexible, komplexe und dennoch robuste Modelle charakterisiert, wirft selbst mit modernen Hochleistungscomputern insbesondere bei größeren Datensätzen beträchtliche Rechenprobleme auf.<sup>590</sup> Vor diesem Hintergrund wird bei Data Mining auf globale Modelle – von denen die Lösungsansätze geleitet sind – verzichtet, was den Algorithmus und dessen rechnerische Umsetzung in den Vordergrund rückt.<sup>591</sup> Der Algorithmus spielt also eine zentrale Rolle im Data Mining.<sup>592</sup> „The key role of programs has lead to an increased emphasis on algorithms in data mining, in contrast to the emphasis on models in statistics. The idea is that one applies the algorithm to data sets, learning how it behaves and what properties it has, regardless of any notion of an underlying model (or pattern) which it might be building.“<sup>593</sup> D. h., statt einheitliche globale Modelle zu bilden, welche alle interessierenden Variablen in sich haben, werden in Data Mining durch Algorithmen Mengen von Aussagen über lokale Abhängigkeiten unter Variablen produziert (in Form von Regeln).<sup>594</sup>

Ein weiterer Unterschied zwischen Data Mining und Statistik ist, dass für Statistik relativ umfangreiche Stichproben vorliegen müssen. Der Grundgedanke der Zuhilfenahme von Stichproben ist das Induktionsprinzip, bei dem von besonderen auf allgemeine Fälle geschlossen wird. Ohne Stichproben können in Statistik kaum relevante Aussagen gemacht werden. Wenn sehr große Datenmengen vorliegen, dann ist es verständlich, warum man aus praktischen Gründen nur mit einer Teilmenge arbeiten will und basierend auf dieser Teilmenge noch Aussagen über die gesamte

---

<sup>585</sup> Vgl. (eigene Übersetzung) Gilmour/Madigan/Pregibon/Smyth (1996), S. 25.

<sup>586</sup> Vgl. (eigene Übersetzung) Gilmour/Madigan/Pregibon/Smyth (1996), S. 26.

<sup>587</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 17.

<sup>588</sup> Vgl. Hudec (o. J.), S. 4.

<sup>589</sup> Die multivariate Statistik befasst sich mit Problemstellungen, in denen mehrere Variablen gemeinsam und gleich bedeutend betrachtet werden sollen.

<sup>590</sup> Vgl. Hudec (o. J.), S. 4.

<sup>591</sup> Vgl. Hudec (o. J.), S. 4.

<sup>592</sup> Vgl. Hand (1999b), S. 17.

<sup>593</sup> Hand (1999a), zit. in: Hudec (o. J.), S. 4.

<sup>594</sup> Vgl. (eigene Übersetzung) Gilmour/Madigan/Pregibon/Smyth (1996), S. 26.

Datenmenge macht. Die neuen Ansätze des Data Mining sind im Gegenteil an keine bestimmte Stichprobe gebunden, denn häufig steht die Gesamtpopulation (-bevölkerung) von Daten für Data Mining-Probleme zur Verfügung (Z. B. alle Kunden bzw. ihre Daten in der Datenbank, alle Transaktionen, die mit denen im letzten Jahr getätigt worden sind u. a.) und man braucht extra von keiner Stichprobe abhängig zu sein, die aus der Gesamtpopulation ausgezogen wird.<sup>595</sup> Darüber hinaus basieren allgemein übliche statistische Techniken auf der Annahme, dass die Datenelemente unabhängig und aus derselben Verteilung gezogen worden sind. Dagegen sind Data Mining-Probleme weit entfernt von der idealisierten i.i.d-Situation (identisch verteilt und stochastisch unabhängig<sup>596</sup>). Sehr große Datenmengen werden wohl kaum in dieser Art und Weise (idd) anfallen<sup>597</sup>, sondern es ist ziemlich wahrscheinlich, dass zu verschiedenen Zeiten manche Bereiche des Variablenraums schwerlich mehr als die anderen gesampelt werden (Z. B. unterschiedliche Zeitzonen bedeuten, dass Geschäftstransaktionen oder Telefonanrufe nicht wahllos bzw. stichprobenartig in ganz USA gemacht werden.).<sup>598</sup> Dies kann Zweifel an die Gültigkeit der maßgebenden Schätzungen aufkommen lassen sowie spezielle Probleme für sequentielle Schätzung und Suchalgorithmen bereiten.<sup>599</sup>

Obwohl es in beiden Disziplinen – Data Mining und Statistik mehr oder weniger um das Gleiche handelt, unterscheiden sie sich in ihren Ansichten zur Datenanalyse. So hat die Statistik mit primärer Datenanalyse zu tun.<sup>600</sup> Das heißt, die Daten werden

<sup>595</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 17.

<sup>596</sup>

Zwei Zufallsvariablen  $X$  und  $Y$  heißen *identisch* verteilt, wenn sie aus demselben Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, P)$  definiert sind und für die Verteilungsfunktionen  $F_X$  von  $X$  und  $F_Y$  von  $Y$   $F_X = F_Y$  gilt.

Eine Menge von Zufallsvariablen  $X_i, i$  aus irgendeiner Indexmenge  $I$ , ist dann *identisch* verteilt, wenn  $X_i = X_j$  für jedes Paar  $i, j \in I$  gilt.

Zwei Ereignisse heißen *stochastisch unabhängig*, wenn  $P(A \cap B) = P(A) \cdot P(B)$  gilt, d. h., wenn die Wahrscheinlichkeit, dass beide Ereignisse gleichzeitig eintreten, gleich dem Produkt ihrer Einzelwahrscheinlichkeiten ist. Das Eintreten des einen Ereignisses beeinflusst das Eintreten des anderen Ereignisses nicht, beide Ereignisse sind also voneinander unabhängig.

Eine Ansammlung von Zufallsvariablen heißt *unabhängig identisch verteilt*, wenn die  $X_i$ 's gegenseitig unabhängig (jede endliche Teilmenge von  $X_i$  ist unabhängig) und identisch verteilt sind.

Dies wird in der Regel mit dem Kürzel "i. i. d" - independent and identically distributed gekennzeichnet.

<sup>597</sup> In der Statistik nimmt man meistens an, dass für Zwecke der statistischen Schlussfolgerung die Beobachtungen in einer Stichprobe mehr oder weniger i. d. d. sind. Die Annahme (Anforderung), dass Beobachtungen i. d. d. sind, sorgt dafür, dass die zugrunde liegende Mathematik von vielen statistischen Methoden vereinfacht wird. In praktischen Anwendungen kann sie allerdings nicht realistisch sein.

<sup>598</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 114.

<sup>599</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 114.

<sup>600</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 112.

besonders im Hinblick auf Fragen oder eine Menge von Fragen absichtlich gesammelt, um die vorgegebenen Fragen zu beantworten und damit das spezielle Problem zu lösen. Die Analyse von Daten im Data Mining ist sekundär<sup>601</sup> oder, besser gesagt retrospektiv.<sup>602</sup> In der Tat darf Data Mining als der Prozess der sekundären Analyse von großen Datenbanken definiert werden, das sich zum Ziel setzt, vorher unvermutete Beziehungen, welche für den Anwender der Datenbank von großem Wert bzw. interessant sind, zu finden.<sup>603</sup> Data Mining ist eher eine induktive Aufgabe als eine hypothetisch-deduktive Vorgehensweise.<sup>604</sup> Data Mining-Systeme mit ihrer prozessorientierten Methode<sup>605</sup> decken einen größeren Problembereich und damit einen breiteren Interessentenkreis ab als klassische Systeme der computergestützten Datenanalyse, deren Primärfunktion vorwiegend in der Methodenanwendung liegt und deren Anwenderkreis eher Personen mit fundierten statistischen Kenntnissen sind.<sup>606</sup> Data Mining-Systeme erlauben nicht nur die Überprüfung vorformulierter Hypothesen, sondern können bei Anwendung entsprechender Methoden diese auch selbständig generieren.<sup>607</sup>

Klassische Statistik geht lediglich mit numerischen Daten um.<sup>608</sup> Heutzutage können aber Datenbanken auch andere mögliche Formen von Daten beinhalten (Z. B. Bild-Daten, Audio-Daten, Text-Daten und geographische Daten). Logische Daten ergeben sich vorwiegend bei der Suche nach Mustern, die sich aus den konjunktiven und disjunktiven Kombinationen von Elementen zusammensetzen.<sup>609</sup> Das heißt, die Elemente der Analyse können oft z. B. Bilder, Texteinheiten, Sprachsignals u. a. sein, was ebensoviel wie einfache numerische Daten auf die Fragen und Probleme von Data Mining – interessante Muster und Strukturen in den Daten zu finden – zutrifft. Dementsprechend ist „Mining the Internet“ ein ausgeprägter Teilbereich von Data Mining geworden.

Zur Präsentation von Mustern hat sich das visuelle Data Mining herausgebildet, welches automatische Mining-Methoden mit Techniken der Informationsvisualisierung verbindet und auf diese Weise numerisch ermittelte Daten graphisch darzustellen und einfach abzulesen ermöglicht. Seine hochentwickelten Visualisierungstechniken

---

<sup>601</sup> Vgl. Hand (1998), S. 112.

<sup>602</sup> Vgl. Gilmour/Madigan/Pregibon/Smyth (1996), S. 25.

<sup>603</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 112.

<sup>604</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 112.

<sup>605</sup> Vgl. Fayyad/Shapiro/Smyth, (1996b), S. 41 f.

<sup>606</sup> Vgl. Bensberg/Grob (1999), S. 4.

<sup>607</sup> Vgl. Bensberg/Grob (1999), S. 4.

<sup>608</sup> Vgl. (eigene Übersetzung) Hand (1998), S. 115 und (1999b), S. 18.

<sup>609</sup> Vgl. (eigene Übersetzung) Hand (1999b), S. 18.

machen das Data Mining einem breiteren Anwenderkreis zugänglich, wodurch auch technisch weniger versierte Business Manager ihre Märkte analysieren und auf diese Weise ihre Geschäftsentscheidungen besser treffen können. Die Beschränkung der klassischen Visualisierung auf drei Koordinatenachsen x, y und z wird hier durch den Einsatz von neueren geometrischen Visualisierungstechniken wie Parallelkoordinatensystemen<sup>610</sup> oder Icons/Glyphen<sup>611</sup> aufgehoben.

Ein weiterer Aspekt, der als ein wesentlicher Schwachpunkt beurteilt werden könnte, ist die Schwierigkeit Expertenwissen (sog. Domain Knowledge) im Rahmen der klassischen statischen Methodologie in den Analyseprozess zu integrieren.<sup>612</sup> Um diese Schwäche der Statistik überwinden zu können, wird beim Data Mining gezielt versucht, Informationen aus unterschiedlichen Quellen möglichst in Form eines interaktiven Dialogs im Mining-Prozess zu berücksichtigen.

### **5.2.3 Verallgemeinerter Data Mining-Prozessablauf**

Im Rahmen des Knowledge Discovery in Databases wird häufig zur Entdeckung und Explikation von in großen Datenbeständen implizit vorhandenem Wissen auf Ansätze und Verfahren von verschiedenen Disziplinen wie Statistik, Datenbanken und Maschinelles Lernen zurückgegriffen.<sup>613</sup> Dennoch sollte man KDD nicht mit diesen Forschungsrichtungen gleichsetzen.<sup>614</sup> KDD unterscheidet sich von Statistik, Datenbanken und Maschinelles Lernen besonders dadurch, dass es nicht auf bestimmte Phasen des Prozesses der Wissensentdeckung ausgerichtet ist, sondern den gesamten Entdeckungsprozess umfasst.<sup>615</sup>

„KDD focuses on the overall process of knowledge discovery from data, including how the data is stored and accessed, how algorithms can be scaled to massive datasets and still run efficiently, how results can be interpreted and visualized, and how the overall human-machine interaction can be modelled and supported.“<sup>616</sup>

Das wesentliche Merkmal des Knowledge Discovery in Databases ist ein interaktiver

---

<sup>610</sup> Beim Prinzip der Parallelen Koordinaten wird jedes Merkmal durch eine parallele Achse repräsentiert; Punkte auf der Achse sind die jeweilige Ausprägung. Ausprägungen eines Datensatzes werden durch eine Linie über alle Achsen verbunden, so dass sich Zusammenhänge zwischen Datensätzen als parallele Linien darstellen, vgl. Hagedorn/Bissantz/Mertens (1997), S. 608.

<sup>611</sup> Glyphen sind grafische Symbole, die jeweils ein Objekt in seinen Merkmalsausprägungen durch unterschiedliche (visuell kodierte) Formen und Farben darstellen, vgl. Hagedorn/Bissantz/Mertens (1997), S. 608.

<sup>612</sup> Vgl. Hudec (o. J.), S. 4.

<sup>613</sup> Vgl. Düsing (2006), S. 245.

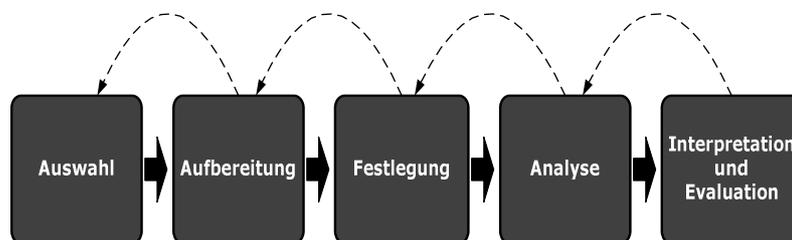
<sup>614</sup> Vgl. Düsing (2006), S. 245.

<sup>615</sup> Vgl. Düsing (2006), S. 245.

<sup>616</sup> Fayyad/Shapiro/Smyth (1996c), S. 29.

und iterativer Prozess.<sup>617</sup> Dieser Prozess durchläuft in Interaktion mit dem Anwender mehrere Phasen iterativ in Rückkopplungsschleifen. Da der Anwender anhand des Ergebnisses einzelner Phasen oder des gesamten Prozesses den weiteren Ablauf des KDD bestimmen wird<sup>618</sup>, muss er von Beginn an eng in den KDD-Prozess eingebunden werden.<sup>619</sup> Wenn der Anwender eine Notwendigkeit dafür sieht, kann er einzelne Phasen oder den gesamten Prozess erneut durchlaufen lassen.

In der Literatur ist eine Vielzahl von Vorgehensmodellen des Knowledge Discovery in Databases zu finden. Diese unterscheiden sich hinsichtlich des Grads der Detaillierung und Ausrichtung an KDD-Werkzeugen.<sup>620</sup> Der Einfachheit halber wird im Folgenden ein vereinfachtes Vorgehensmodell dargestellt, das auf das ursprüngliche KDD-Modell von Fayyad und Shapiro zurückzuführen ist und den KDD-Prozess in fünf Phasen gliedert:



**Abbildung 33: Vorgehensmodell des Knowledge Discovery in Databases**

Quelle: Düsing (2006), S. 246.

1. **Auswahlphase:** In dieser ersten Phase wird die Ausgangslage des Knowledge Discovery in Databases bestimmt. Dafür soll zunächst die Aufgabenstellung des KDD ausgewählt werden. Obwohl die Entdeckung unbekannter Sachverhalte und Zusammenhänge eines der typischen Hauptziele des KDD darstellt, ist es dennoch sinnvoll, eine Aufgabenrichtung vorzugeben und diese angemessen konkret zu definieren. Dabei ist allerdings zu beachten, dass die ausgewählte Aufgabe einerseits hinreichend allgemein formuliert ist, um bei der Analyse die Generierung von Hypothesen über den durch die Daten abgebildeten Gegenstandsbereich zu ermöglichen, und andererseits genügend konkret definiert ist, um eine Erfolgskontrolle des KDD zu gewährleisten.<sup>621</sup>

Die Datenbestände können prinzipiell auch ohne Vorgabe konkreter Ziele analysiert werden. In diesem Fall ist jedoch die Analyse nicht besonders

<sup>617</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 42 sowie (1996c), S. 30

<sup>618</sup> Vgl. Düsing (2006), S. 246.

<sup>619</sup> Vgl. Düsing (2006), S. 246. Die intensive Einbindung des Anwenders in den KDD-Prozess hebt die Eigenschaft hervor, dass KDD ein „user-driven“ Prozess ist, den der Anwender durch seine Entscheidungen unterstützt, vgl. Fayyad/Shapiro/Smyth (1996b), S. 42 und (1996c), S. 30. Diese begründet außerdem die fehlende volle Automatisierbarkeit des Knowledge Discovery in Databases-Prozesses, vgl. Düsing (2006), S. 246 und Wilde (2001), S. 13.

<sup>620</sup> Vgl. Düsing (2006), S. 246.

<sup>621</sup> Vgl. Düsing (2006), S. 247.

effektiv und bringt keinen ökonomischen Nutzen im Rahmen des CRM. Eine freie Suche nach irgendwelchen Auffälligkeiten in den Daten ist vor diesem Hintergrund eher sinnlos.<sup>622</sup> Vor der Durchführung der Analyse ist daher festzulegen, worauf sich die Analyse ausrichtet und entsprechend dieser Ausrichtung welche Art von Ergebnissen bzw. Mustern erwartet wird. Außerdem muss man bei der Auswahl der Aufgabenstellung sicherstellen, dass das im Rahmen der Aufgabenstellung zu entdeckende und explizit zu machende Wissen anwendbar und bezüglich seiner Anwendung wirtschaftlich ist.<sup>623</sup>

Nachdem die Aufgabestellung des KDD bestimmt wurde, muss nun die für diese Aufgabestellung relevante Datenbasis ausgewählt werden. Bei der Auswahl der Datenbasis ist insbesondere zu klären, welche Daten in welchem Umfang benötigt werden.<sup>624</sup> Wenn kein Data Warehouse vorhanden ist oder dieses nicht alle benötigten Daten enthält, wird es dann auf operative Datenbestände oder externe Quellen zurückgegriffen. Falls sich der Analyseprozess über einen längeren Zeitraum erstreckt, muss sichergestellt werden, dass sich der zu verarbeitende Datenbestand in dieser Zeit nicht verändert.<sup>625</sup> Darüber hinaus soll überprüft werden, ob der Datenbestand im Rahmen des KDD eine Analyse zulässt bzw. dafür verfügbar ist.<sup>626</sup> Die Verfügbarkeit eines Datenbestands für eine Analyse könnte aus technischen, organisatorischen oder rechtlichen Gründen eingeschränkt sein.

2. **Aufbereitungsphase:** Die Qualität der Ergebnisse des Analyseprozesses hängt wesentlich von der Qualität der zu verarbeitenden Daten ab. Das Ziel der Aufbereitungsphase ist es, die Qualität der ausgewählten Datenbasis im Hinblick auf die Vollständigkeit und Konsistenz der Daten zu verbessern.<sup>627</sup> Diese Phase umfasst die Integration, Bereinigung, Anreicherung und Reduktion des Datenbestands.

Die ausgewählten Daten befinden sich meistens in diversen, voneinander unabhängigen Datenquellen und weisen spezielle Charakteristiken in Bezug auf Datendarstellung und -organisation auf.<sup>628</sup> Solche Daten müssen für eine

---

<sup>622</sup> Vgl. Neckel/Knobloch (2005), S. 83 und Krahl/Windheuser/Zick (1998), S. 30.

<sup>623</sup> Vgl. Düsing (2006), S. 247.

<sup>624</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 30.

<sup>625</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 31.

<sup>626</sup> Vgl. Düsing (2006), S. 247.

<sup>627</sup> Vgl. Düsing (2006), S. 248.

<sup>628</sup> Vgl. Düsing (2006), S. 248.

gemeinsame Analyse durch eine physische und logische Integration in eine einheitliche Datenbasis überführt werden.<sup>629</sup>

Die Datenbereinigung beinhaltet die Bearbeitung von fehlerhaften oder unvollständigen Daten. Um eine möglichst hohe Ergebnisqualität zu gewährleisten, ist eine Säuberung der Daten erforderlich, indem Duplikate oder irrelevante Daten ausgefiltert werden und festgelegt wird, wie unvollständige oder sich im Zeitverlauf ändernde Daten bearbeitet werden sollen. Für die Bearbeitung fehlender und fehlerhafter Merkmalausprägungen des Datenbestands bieten sich verschiedene Ansätze an:

Die fehlenden oder fehlerhaften Daten werden entweder komplett aus dem Datenbestand entfernt, indem sie gelöscht werden, oder man korrigiert sie vor der Durchführung der Analyse.<sup>630</sup> Die nachträgliche Erhebung der fehlenden und fehlerhaften Daten - falls sie trotz des hohen Risikos, falsche Analyseergebnisse zu erhalten, dennoch in die Analyse mit einbezogen werden - ist zwar möglich, aber mit sehr viel Aufwand verbunden, und sie erweist sich im Falle von Daten aus vergangenen Zeitperioden als nicht durchführbar.<sup>631</sup> Als eine weitere Möglichkeit können die fehlenden und fehlerhaften Daten ersetzt werden, indem man dafür das sogenannte Imputationsverfahren anwendet.<sup>632</sup>

Durch eine Anreicherung der Daten können dem Datenbestand weitere Merkmale, die für die Bearbeitung der Aufgabenstellung des KDD zweckmäßig sind, hinzugefügt werden. Diese Datenanreicherung erfolgt entweder auf der Grundlage von aus den vorhandenen Merkmalen der Datenbasis abgeleiteten Merkmalen oder durch die Aufnahme von aus anderen Datenquellen, wie z. B. Statistischen Ämtern oder Marktforschungsinstituten bezogenen Merkmalen.<sup>633</sup> Die Anzahl dieser Merkmale des Datenbestands kann schließlich bei der Datenreduktion verringert werden, was folglich zu einer Reduzierung der Analysekomplexität führt. Dies ist im Fall umfangreicher Datenbestände eine wichtige Voraussetzung für die Durchführbarkeit der Analyse.<sup>634</sup> Alle diese Schritte von der Datenbereinigung bis zur Datenreduktion verfolgen ein gemeinsames Ziel: den Datenbestand so bereitzustellen, dass er im folgenden Schritt verwendet, d. h. entsprechend der Aufgabenstellung analysiert werden

---

<sup>629</sup> Vgl. Düsing (2006), S. 248.

<sup>630</sup> Vgl. Düsing (2006), S. 249 und Nakhaeizadeh (1998), S. 21.

<sup>631</sup> Vgl. Düsing (2006), S. 248f.

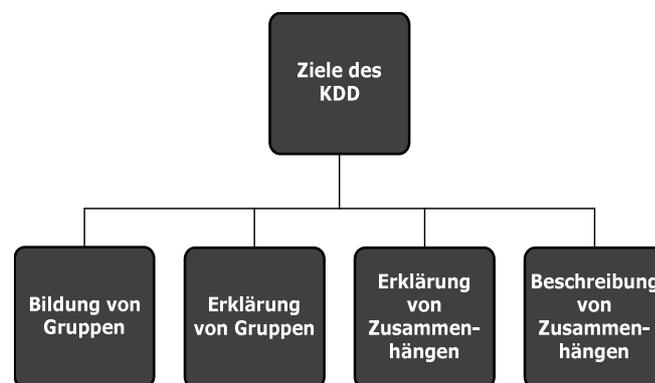
<sup>632</sup> Vgl. ebd., S. 249.

<sup>633</sup> Vgl. ebd., S. 249

<sup>634</sup> Vgl. ebd., S. 249.

kann.

3. **Festlegungsphase:** In der Festlegungsphase werden sämtliche Tätigkeiten zur Vorbereitung der Analyse der Datenbasis ausgeführt. Zu diesen Tätigkeiten gehören die Festlegung der Zielsetzung, des Analyseverfahrens sowie des Umfangs und der Darstellung der zu analysierenden Daten.<sup>635</sup> Je nach Bedarf können verschiedene Zielsetzungen im Rahmen des Knowledge Discovery in Databases festgelegt werden. Eine typische Klassifizierung von Zielen ist in der folgenden Abbildung 34 dargestellt. Mit der Festlegung des Analyseverfahrens werden das oder die Verfahren zur Analyse der Datenbasis sowie die Reihenfolge der Ausführung der Verfahren bestimmt.



**Abbildung 34: Klassifikation der Ziele des Knowledge Discovery in Databases**

Quelle: Düsing (2006), S. 250.

Je nach festgelegter Zielsetzung können verschiedene Analyseverfahren bzw. -werkzeuge zum Einsatz kommen. Diese Verfahren sind z. B. Clusterverfahren, Entscheidungsbaumverfahren, Konnektionistische Systeme (Neuronale Netze) und Assoziationsregelverfahren. Es besteht die Möglichkeit, die Ergebnisse eines mit einem bestimmten Verfahren ausgeführten Analyseschrittes durch den Einsatz anderer Verfahren und Werkzeuge zu überprüfen bzw. zu verbessern.<sup>636</sup> Dies zeigt sogar, dass KDD ein iterativer Prozess ist. Weiters soll die Anzahl der Datensätze der zu analysierenden Datenbasis bestimmt werden. Die Festlegung des Datenumfangs dient zur Entscheidung, ob die Analyse auf der Grundlage des gesamten Datenbestands, oder einer Stichprobe durchgeführt werden soll.<sup>637</sup>

<sup>635</sup> Vgl. Düsing (2006), S. 249.

<sup>636</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 31.

<sup>637</sup> Mit der Verwendung einer Stichprobe ist oftmals eine Verkleinerung des Datenraums und damit eine Reduzierung der Analysekomplexität verbunden. Dennoch besteht die Gefahr, dass man durch die Verwendung einer Stichprobe öfters zu einem im Vergleich zur Verwendung des gesamten Datenbestands schlechteren Analyseergebnis gelangt, vgl. Düsing (2006), S. 252.

4. **Analysephase:** In der Analysephase werden anhand eines oder mehrerer Analyseverfahren die Zusammenhänge und Auffälligkeiten in den Daten ermittelt und als Beziehungsmuster abgebildet. Diese Analysetätigkeiten können auch unter der Bezeichnung Data Mining zusammengefasst werden.<sup>638</sup>

Data Mining als ein Ansatz zur Analyse von Daten ist darauf ausgerichtet, die Beziehungsmuster (patterns) in der zugrunde liegenden Datenbasis zu ermitteln und durch logische und mathematische Beschreibungen abzubilden.<sup>639</sup> Es ist ein Teilprozess des Knowledge Discovery in Databases.<sup>640</sup> Im Rahmen des CRM wird nach entsprechender Selektion und Aufbereitung der Kundendaten das Verhalten von Kunden in einem Modell abgebildet. Auf der Grundlage dieses Modells kann ein Data Mining-System sowohl hypothesenbasierte Aussagen über die Vergangenheit machen (Verifikationsmodelle, konfirmatorische Datenanalyse der traditionellen Statistik) als auch durch automatische Entdeckung neuer Beziehungsmuster in den Daten die Zukunft bezüglich des Kundenverhaltens prognostizieren (vorhersagende Entdeckungsmodelle, explorative Datenanalyse).<sup>641</sup>

Dem Data Mining wird häufig eine Menge bestimmter Datenanalysemethoden zugeordnet. Allerdings soll es hier zwischen top down- und bottom up-Analyseverfahren unterschieden werden.

Möchte man eine Annahme, Vermutung oder Theorie anhand verfügbarer Daten verifizieren oder falsifizieren, so geht man „hypothesengetrieben“<sup>642</sup> vor. Die Datenbestände werden in diesem Fall nach Sachverhalten durchgesucht, die die definierten Hypothesen stützen oder widerlegen.<sup>643</sup> Diese Vorgehensweise wird auch als klassische Datenanalyse bezeichnet. Die Hypothesenverifikation ist das Ziel von top down-Analysen.<sup>644</sup> Typische Vertreter der top down-Analysemethode sind die traditionelle Statistik, die Datenbankabfragen mittels SQL und das On-line Analytical Processing (OLAP). Verfolgt man hingegen ohne konkrete Annahmen das Ziel, in vorliegenden Daten interessante Sachverhalte zu entdecken, um auf diese Weise zu

---

<sup>638</sup> Vgl. Düsing (2006), S. 252.

<sup>639</sup> Vgl. Düsing (2006), S. 252.

<sup>640</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 40 und (1996c), S. 28.

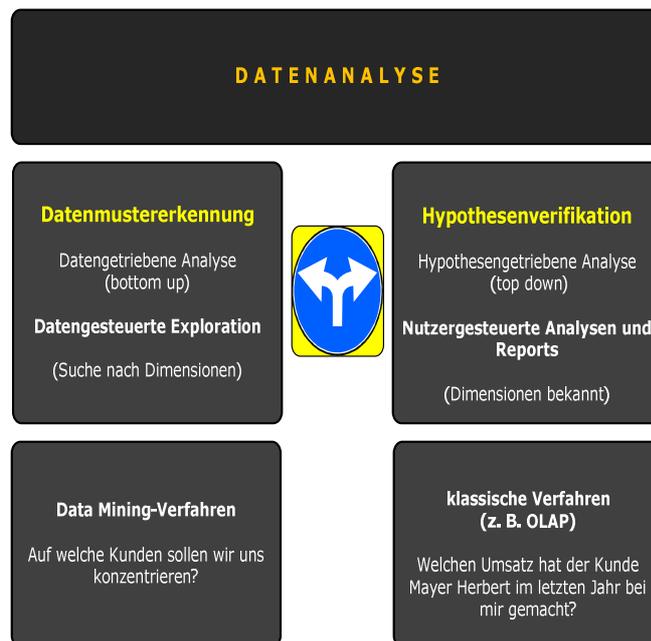
<sup>641</sup> Vgl. Zipser (2001), S. 44, Schildhauer/Grothe/Braun/Schultze (2004), S. 33 und Fayyad/Shapiro/Smyth (1996b), S. 43.

<sup>642</sup> Bei der hypothesengetriebenen Datenanalyse werden von einem einschränkenden Modell der Wirklichkeit Hypothesen abgeleitet, die meistens mit einer Stichprobe überprüft werden, vgl. Lusti (2002), S. 263.

<sup>643</sup> Vgl. Knobloch/Neckel (2005), S. 79.

<sup>644</sup> Vgl. Knobloch/Neckel (2005), S. 79.

Erkenntnissen zu gelangen, wählt man ein bottom up-Analyseverfahren. Man spricht dann von „datengetriebenen“ Analysen.



**Abbildung 35: top down-versus bottom up-Analyseverfahren**

Quelle: in Anl. an Knoblauch/Neckel (2005), S. 83 und Zipser (2001), S. 42.

Im Vergleich zu top down-Analysen ist der Suchraum bei bottom up-Analysen wesentlich komplexer, da hier keine Begrenzung durch konkrete Annahmen vorgegeben ist.<sup>645</sup>

Eine Datenanalysemethode wird dann als ein eigentliches Data Mining-Verfahren angesehen, wenn sie dazu geeignet ist, auf der Grundlage eines gesamten Datenbestands bzw. einer Datengesamtheit datengetriebene Untersuchungen durchzuführen, die das Ziel der Datenmustererkennung verfolgen.<sup>646</sup> Das Data Mining-Verfahren kann demnach auch als „hypothesenfreies Suchen“<sup>647</sup> bezeichnet werden. Während sich klassische Verfahren eher auf die quantitative Dimension des zu analysierenden Datenbestands konzentrieren, erlaubt ein Data Mining-Verfahren auch qualitative Aussagen über diesen Datenbestand und liefert somit fundiertes Wissen für zukünftige wirtschaftliche Entscheidungen.

5. **Interpretations- und Evaluationsphase:** Nach der Durchführung der Analyse werden die Ergebnisse interpretiert bzw. bewertet. Die Ergebnisse

<sup>645</sup> Vgl. Knobloch/Neckel (2005), S. 81.

<sup>646</sup> Vgl. ebd., S. 83.

<sup>647</sup> Die Trennung zwischen daten- und hypothesengetriebener Suche ist jedoch als idealisierende Systematik zu verstehen. Denn in der Praxis existieren keine vollkommen hypothesenfreien Datenanalysen, vgl. ebd., S. 83.

einer Datenanalyse müssen bestimmten Anforderungen genügen, um zur Lösung von CRM-Problemen beitragen zu können. Das besondere Augenmerk wird hier auf die Interessantheit der Data Mining-Ergebnisse gelegt.<sup>648</sup> „Erst die Interessantheit eines Ergebnisses macht eine entdeckte Information zu wertvollem Wissen.“<sup>649</sup> Die Interessantheit setzt sich aus verschiedenen Einzelkriterien zusammen. Generell kann ein Ergebnis nur dann als interessant bzw. positiv bewertet werden, wenn es gültig, neu, potentiell nützlich und leicht verständlich ist.<sup>650</sup> Alle diese Gütekriterien sind für die Interessantheit der Ergebnisse relevant.<sup>651</sup> Allerdings sind die Interessantheit bildende Kriterien nur schwer messbar und hängen von der konkreten Anwendung und den damit verbundenen Zielen ab.<sup>652</sup> Darüber hinaus ist die Interessantheit nicht objektiv, sondern stark vom jeweiligen Anwender abhängig.<sup>653</sup> Sie kann von Person zu Person ganz unterschiedlich wahrgenommen werden. So könnte z. B. ein Fachexperte an spezifischen Details eines Ergebnisses interessiert sein, während dasselbe Ergebnis jemand anderem kaum etwas bedeutet. Ein Ergebnis muss daher einerseits durch denjenigen, der mit der Data Mining-Durchführung betraut ist, und andererseits auch durch den Anwender dieses Ergebnisses beurteilt werden, um zu einer korrekten Ergebnisinterpretation zu gelangen.<sup>654</sup>

Beim KDD-Prozess verteilt sich der Zeitaufwand zu 20% auf die Definition und Festlegung der Aufgabenstellung, zu 60% auf die Auswahl der relevanten Datenbestände und Datenaufbereitung, zu 10% auf die Interpretation, Evaluation und Anwendung der Data Mining-Ergebnisse und nur zu 10% auf die Auswahl und Anwendung der Data Mining-Methoden.<sup>655</sup>

### **5.2.4 Aufgaben des Data Mining**

Im Rahmen von Data Mining-Projekten kann grundsätzlich zwischen zwei Aufgabenrichtungen unterschieden werden: die Vorhersage (prediction) und die Beschreibung (description).<sup>656</sup> Während sich die Prediktion darauf abzielt, aus den

---

<sup>648</sup> Vgl. Knobloch/Neckel (2005), S. 188.

<sup>649</sup> Nakhaeizadeh (1998), S. 23.

<sup>650</sup> Vgl. Nakhaeizadeh (1998), S. 22.

<sup>651</sup> Eine ausführliche Erklärung dieser vier Kriterien ist in: Nakhaeizadeh (1998), S. 22 sowie Knobloch/Neckel (2005), S. 189f. zu finden.

<sup>652</sup> Vgl. Nakhaeizadeh (1998), S. 22.

<sup>653</sup> Vgl. Knobloch/Neckel (2005), S. 191.

<sup>654</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 31.

<sup>655</sup> Vgl. Wilde (2001), S. 16.

<sup>656</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 43.

bestimmten Variablen der Datenbasis unbekannte oder zukünftige Werte anderer Variablen abzuleiten, fokussiert die Deskription auf die Entdeckung von in den Datenbeständen vorhandenen interpretierbaren Mustern, welche die Daten beschreiben.<sup>657</sup>



**Abbildung 36: Aufgabenstellungen des Data Mining**

Quelle: in Anlehnung an Zeller (2003), S. 126.

Im Folgenden werden die wesentlichen Aufgabentypen des Data Mining näher erläutert:

- *Datenbeschreibung und -zusammenfassung:*

Durch eine Datenbeschreibung und -zusammenfassung werden die wesentlichen Eigenschaften der Daten in kompakter Form beschrieben, um einen besseren Überblick über die Struktur dieser Daten zu gewinnen.<sup>658</sup> Die Datenbeschreibung kann manchmal das Hauptziel eines Data Mining-Projektes sein, stellt aber i. d. R. eher eine Teilaufgabe innerhalb des Data Mining-Prozesses dar, die insbesondere am Anfang des Prozesses von Bedeutung ist, da der Analytiker meistens in frühen Mining-Phasen keine genauen Vorstellungen über das Ziel der Analyse sowie über die Charakteristika der zu analysierenden Daten hat. Eine erste explorative Datenanalyse hilft dabei, die Daten besser zu verstehen.<sup>659</sup> Die Daten werden durch Methoden zusammengefasst, die eine Teilmenge dieser Daten kompakt beschreiben.<sup>660</sup> Dadurch werden sie überschaubar, was auch zu einer Postulierung interessanter Datensegmente führen kann, die dann gezielt weiteranalysiert

<sup>657</sup> (eigene Übersetzung) Fayyad/Shapiro/Smyth (1996b), S. 44.

<sup>658</sup> Vgl. Nakhaeizadeh (1998), S. 9.

<sup>659</sup> Vgl. Nakhaeizadeh (1998), S. 9.

<sup>660</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 45.

werden. Daher ist es sinnvoll, vor der Bearbeitung von anderen Data Mining-Zielen zunächst eine Datenbeschreibung und -zusammenfassung durchzuführen.<sup>661</sup>

- *Segmentierung (Gruppenbildung):*

Die Segmentierung (auch Clustering genannt) zielt ab auf die Gliederung der Datensätze in interessante und inhaltlich bedeutsame Gruppen (sogenannte Clusters), die durch gemeinsame Merkmalausprägungen beschrieben sind.<sup>662</sup> Die Elemente desselben Clusters sollen dabei hinsichtlich ihrer Eigenschaften möglichst ähnlich, die Elemente aus verschiedenen Clustern möglichst unterschiedlich sein. Durch die Gruppenbildung können die Daten effizient analysiert werden. Es ist wesentlich einfacher und meist auch sinnvoller, die Beziehungen und Abhängigkeiten innerhalb von sinnvollen Segmenten anstatt eines unsegmentierten Datenhaufens zu ermitteln (wie es z. B. bei der Warenkorbanalyse der Fall ist). Ein gutes Beispiel für die Gruppenbildung ist die Kundensegmentierung, bei der die Kunden z. B. anhand ihres Kaufverhaltens in homogene Gruppen zusammengefasst werden, die dann im Rahmen des CRM differenziert bearbeitet werden können.

- *Assoziationsanalyse:*

Die Assoziationsanalyse dient zur Beschreibung interessanter Beziehungsabhängigkeiten oder -zusammenhänge zwischen unterschiedlichen Merkmalsausprägungen von Datensätzen. Anders formuliert zielt die Assoziationsanalyse darauf ab, Assoziationsregeln zu erstellen, die häufig auftretende, in den Datenbeständen versteckte Regeln oder Muster beschreiben. Typisches Beispiel für eine Assoziationsanalyse ist eine Warenkorbanalyse, durch die komplementäre Beziehungen zwischen verkauften Artikeln z. B. in einem Supermarkt identifiziert werden.

- *Klassifikation:*

Bei der Klassifikation wird der zugrunde liegende Datenbestand vordefinierten, durch bestimmte Merkmale oder Eigenschaften beschriebenen Klassen zugeordnet. Diese Klassen können entweder vom Anwender vorweg definiert werden oder aus Segmentierung bzw. Clustering resultieren.<sup>663</sup> Für die Klassifikationsaufgabe wird ein Modell erstellt (auch Klassifikator genannt), das neue, bisher unbekannte

---

<sup>661</sup> Vgl. Nakhaeizadeh (1998), S. 10.

<sup>662</sup> Vgl. ebd., S. 7.

<sup>663</sup> Vgl. ebd., S. 8.

Datenobjekte in richtige Klassen einordnet.

Die Kreditwürdigkeitsprüfung der Bankkunden ist ein sehr gutes, weit verbreitetes Beispiel für die Klassifikation. Hierzu werden zwei Klassen - gute und schlechte Kunden - gebildet. Die bekannten Kunden der Bank werden jeweils einer der beiden Klassen zugewiesen. Aus dieser Trainingsmenge kann ein Klassifikator gemustert werden, den man dazu benutzt, neue potentielle Kunden entweder als gut oder schlecht zu klassifizieren, um sie dann entweder zu akzeptieren oder abzuweisen.

Die Klassifikation gilt als eine der wichtigsten Data Mining-Aufgaben, da sich viele Anwendungsprobleme darauf abbilden lassen.<sup>664</sup> Sie hat Verbindungen zu allen anderen Data Mining-Aufgaben.<sup>665</sup> So herrscht z. B. eine enge Verbindung zwischen der Klassifikation und Konzeptbeschreibung, die die Klassen oder Konzepten von Objekten verständlich beschreibt, wodurch neue Einsichten gewonnen werden können. Die Beschreibung der wesentlichen Merkmale von Konzepten ermöglicht, die verschiedenen Konzepte voneinander zu unterscheiden.<sup>666</sup> Daher können Konzeptbeschreibungen auch zur Klassifikation verwendet werden. Die Data Mining-Aufgabe Segmentierung fasst auch Datenobjekte zu Konzepten zusammen. Dies führt jedoch nicht ohne Weiteres zu einer verständlichen Beschreibung dieser Konzepte.<sup>667</sup> Zu den Methoden der Klassifikation zählen z. B. die logistische Regressionsanalyse, Diskriminanzanalyse, Klassifikationsbäume, Künstlichen Neuronalen Netze und genetischen Algorithmen.

- *Abweichungsanalyse:*

Die Abweichungsanalyse sucht automatisch nach Objekten, bei denen bestimmte Merkmalswerte von einer Norm oder einem erwarteten Wert abweichen. Solche Abweichungen können einerseits auf potenzielle Probleme mit den Daten hinweisen und eine entsprechende Datenbereinigung nach sich ziehen. Andererseits können diese Abweichungen aber auch ein Anzeichen für ein dem Anwender unbekanntes Phänomen sein, das detailliert analysiert werden muss.

---

<sup>664</sup> Vgl. Nakhaeizadeh (1998), S. 8.

<sup>665</sup> Z. B. können Prognoseaufgaben durch eine entsprechende Diskretisierung der Werte der Klassen in Klassifikationsaufgaben transformiert werden, vgl. ebd., S. 8.

<sup>666</sup> Eine Konzeptbeschreibung der Konzepte für bspw. treue und untreue Kunden beinhaltet die wesentlichen Merkmale, die treue Kunden von untreuen unterscheiden. Diese Kenntnis ermöglicht dem Anwender gezielte Maßnahmen für beide Kundengruppen einzuleiten, um z. B. treue Kunden noch stärker zu binden oder um untreue Kunden in treue umzuwandeln. Zudem können die Konzeptbeschreibungen der treuen und untreuen Kunden auf neu gewonnene Kunden angewendet werden, um deren zukünftigen Status zu prognostizieren, vgl. ebd., S. 8f.

<sup>667</sup> Vgl. ebd., S. 9.

Ein klassisches Beispiel für die Abweichungsanalyse ist die automatische Erkennung auffälliger Entwicklungen in Haushaltspanel-Daten, die kontinuierlich für eine Vielzahl von Artikeln anfallen.<sup>668</sup> Das Erkennen von Abweichungen ist mit dem Problem der dynamischen Umgebungen eng verbunden.<sup>669</sup> Da sich in vielen Anwendungen die Daten und die zugrunde liegenden Konzepte ständig ändern (ständige Erfassung neuer Daten über Kunden, neue Datenerfassungsmethoden und sich fortwährend änderndes Kundenverhalten), verlieren nach einiger Zeit die in einem Data Mining-Prozess entdeckten Zusammenhänge ihre Aktualität und daher müssen die Daten neu analysiert werden.

- *Regression und Vorhersage:*

Die Vorhersage bzw. Prognose versucht fehlende numerische Merkmalswerte eines Objekts zu ergänzen, die sich auf vergangene, gegenwärtige oder zukünftige Zeitpunkte beziehen (Z. B. die Prognose des Auftragsvolumens eines Kunden für die nächste Saison auf der Basis seiner bisherigen Kaufhistorie).

Die Prognose erweitert die Klassifikation um die zukünftige Zeitkomponente. Während es bei der Klassifikation die heutige Klassenzugehörigkeit bzw. Ausprägung einer zeitlich veränderlichen Größe symbolisch bestimmt wird, zielt die Prognose darauf ab, deren zukünftige Klassenzugehörigkeit oder Ausprägung numerisch zu ermitteln.<sup>670</sup> Die dafür geeigneten Data Mining-Methoden sind z. B. die Regressionsanalyse, Regressionsbäume, neuronalen Netze, Box-Jenkins-Methode und genetischen Algorithmen.

## **5.2.5 Methoden des Data Mining**

Für Fayyad, Shapiro und Smyth lassen sich die Zielsetzungen des Data Mining grundsätzlich in zwei Arten, nämlich die Vorhersage und Beschreibung unterscheiden.<sup>671</sup> Unter Anwendung bestimmter Data Mining-Methoden können beide Zielarten erreicht werden. Darüber hinaus kann man das allgemeine und übergeordnete Ziel des Data Mining als die „[...] effiziente Suche nach versteckten, aber potentiell nützlichen Informationen in großen Datenbeständen“<sup>672</sup> formulieren. Dies erfolgt auf der Grundlage eines induktiven maschinellen Lernens.<sup>673</sup>

---

<sup>668</sup> Vgl. Hippner/Wilde (2003), S. 219.

<sup>669</sup> Vgl. Nakhaeizadeh (1998), S. 10.

<sup>670</sup> Vgl. Nakhaeizadeh (1998), S. 9.

<sup>671</sup> Vgl. Fayyad/Shapiro/Smyth (1996b), S. 44.

<sup>672</sup> Vgl. Mertens/Wieczorrek (2000), S. 211.

<sup>673</sup> Vgl. Mertens/Wieczorrek (2000), S. 212 und Krahl/Windheuser/Zick (1998), S. 59f.

Induktion bedeutet die Ableitung von Informationen aus Daten. Damit weist sie auf die Entdeckung neuer, vorher nicht bekannter Zusammenhänge und Abhängigkeiten in den Daten hin.

Maschinelles Lernen bezeichnet „die Automatisierung eines Lernprozesses, der zum Ziel hat, auf einer Maschine eine Aufgabe beim nächsten Mal genauer, d. h. mit einer geringeren Fehlerwahrscheinlichkeit zu lösen als beim vorherigen Mal.“<sup>674</sup> Zudem steht es als ein Oberbegriff für die künstliche Generierung von Wissen aus Erfahrung: Ein künstliches System lernt aus Beispielen und kann nach der Beendigung der Lernphase diese Beispiele zu Regeln verallgemeinern (induktives Lernen). D. h. es lernt nicht einfach die Beispiele auswendig, sondern es erkennt die Gesetzmäßigkeiten in denen, so dass es dann diese Erfahrung auch für die Beurteilung unbekannter neuer Daten nutzen kann.

Grundsätzlich lassen sich folgende Lernverfahren unterscheiden:<sup>675</sup>

- *Induktives Lernen:* Bei diesem Lernverfahren werden in großen Datenbeständen neue Zusammenhänge, die vorher nicht bekannt waren, entdeckt. Die Voraussetzung hierfür ist das Vorhandensein vieler Daten, über die man geringes Wissen verfügt. Dieses Lernverfahren wird oft bei der Analyse von Kundentypen, Optimierung von Marketingaktionen oder Modellierung von komplexen Modellen des Devisen- und Aktienmarktes angewendet.
- *Deduktives Lernen:* Das Ziel des deduktiven Lernverfahrens ist, das in Wissensbasen gegebene Wissen zu analysieren, um es dann effizienter und automatisch einsetzen zu können. Zu diesem Zweck werden neue Erkenntnisse aus logischen Schlussfolgerungen (aus Wenn-dann Regeln) gewonnen. Dieses Lernverfahren eignet sich besonders zur Automatisierung von technischen Regelungsprozessen unter Verwendung von bekannten expliziten Regeln bzw. Expertenwissen.

Da die Data Mining-Anforderung, automatisch Trends und Verhaltensmuster vorherzusagen und bisher unbekannte Strukturen in Daten aufzudecken eher der Methodik des induktiven maschinellen Lernens entspricht, können Data Mining bzw. Data Mining-Systeme in den Kontext des induktiven maschinellen Lernens eingebettet werden.<sup>676</sup> Jedoch gibt es einige wesentliche Punkte, die im Folgenden zum besseren

---

<sup>674</sup> Krahl/Windheuser/Zick (1998), S. 59.

<sup>675</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 59 und Mertens/Wieczorrek (2000), S. 212.

<sup>676</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 59.

Verständnis und zur Abgrenzung beider Begriffe Data Mining und Induktives maschinelles Lernen erläutert werden müssen:<sup>677</sup>

- Induktives maschinelles Lernen behandelt Fragestellungen wie z. B. die Optimierung von Modellen physikalischer dynamischer Systeme oder die Erforschung des menschlichen Lernens. Data Mining bezieht sich dagegen auf das Finden verständlichen Wissens aus unstrukturierten Datenbeständen und adressiert damit die Anwender aus der realen Welt.
- Maschinelles Lernen behandelt typischerweise kleinere und idealisierte Datenmengen. Im Gegensatz dazu werden bei Data Mining sehr große und echte Datenbestände bearbeitet.
- Induktives maschinelles Lernen ist ein Prozess der Modellbildung. Die abgeleiteten Informationen stellen sich als ein Modell hinter den zugrunde liegenden Daten dar. Das Ergebnis dieses Modells ist i. d. R. eine Gruppierung ähnlicher Objekte, die zu Klassen zugeordnet sind, wobei auf der Grundlage dieser Klassen dann die Klassenzuordnung neu auftretender Objekte vorherbestimmt werden kann. Zudem könnte aber auch das automatische Aufdecken von in Daten versteckten speziellen Mustern, die sich nur aus einem kleinen, vorher nicht definierten Ausschnitt des Datenbestandes ergeben, als ein Ergebnis betrachtet werden.
- Induktives Lernen lässt sich in „überwachtes“ und „unüberwachtes“ Lernen unterteilen.

Basierend auf dem induktiven maschinellen Lernen weisen Data Mining-Systeme folgende Leistungsmerkmale auf:<sup>678</sup>

- Automatisierte Vorhersage von Trends, Verhalten und Mustern auf der Basis hinterlegter Daten und bereits bekannter Verhaltensschemata aus der Vergangenheit (überwachtes Lernen).
- Automatisierte Aufdeckung unbekannter Strukturen und Zusammenhänge in ungeordneten Datenbeständen.
- Preprocessing (Datenübernahme aus einem Data Warehouse, Datenbereinigung sowie analytische Grundfunktionalitäten wie z. B. das Erkennen von

---

<sup>677</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 59f.

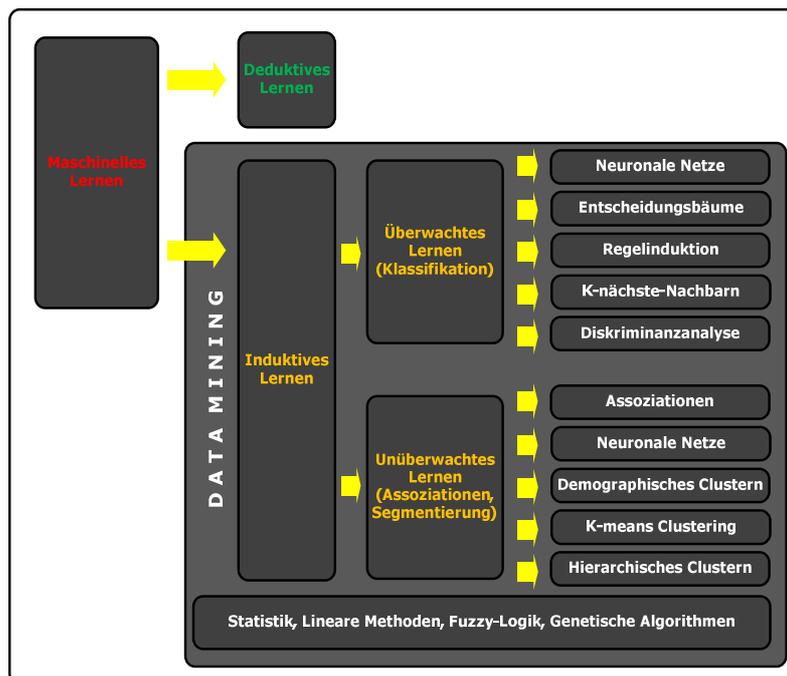
<sup>678</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 24 und Mertens/Wieczorrek (2000), S. 212.

Ausreißern) und Ergebnisaufbereitung (Visualisierung) durch entsprechende Zusatzkomponenten.

- Im Gegensatz zu SQL-Abfragen, Reportgeneratoren und OLAP-Werkzeugen beinhaltet Data Mining keine Analysevorgänge, die explizite Informationen erzeugen. Data Mining-Systeme haben implizite Informationen als Analyseergebnis.

Nun werden die Methoden von Data Mining kurz dargestellt. Da für diese Arbeit die Segmentierung und damit auch die Fuzzy-Logik im Vordergrund steht, wird es auf diese beiden im sechsten Kapitel detaillierter eingegangen.

Zunächst sollen die Methoden im Hinblick auf deren Lenkung differenziert werden.<sup>679</sup> Ausgehend von dieser Unterscheidung lassen sich dann zwei grundsätzliche Verfahren des induktiven Lernens zur Data Mining-Anwendung herauskristallisieren, nämlich überwachtes (Lernen aus Beispielen) und unüberwachtes Lernverfahren (Lernen durch Beobachtung), denen die entsprechenden Methoden zugeordnet werden können.



**Abbildung 37: Einbettung von Data Mining im maschinellen Lernen**

Quelle: Krahl/Windheuser/Zick (1998), S. 61.

Die Einbettung des Data Mining innerhalb des maschinellen Lernens sowie die Zuordnung der Methoden den Bereichen des überwachten und unüberwachten Lernens ist in der Abbildung 37 zu sehen.

<sup>679</sup> Vgl. Neckel/Knobloch (2005), S. 187.

Alle Data Mining-Ansätze haben gemeinsam, dass die zu bearbeitenden Aufgaben der Datenanalyse durch zwei Schritte gelöst werden müssen:<sup>680</sup>

So werden im ersten Schritt durch die Mustererkennung sämtliche Gruppierungen von Datenobjekten herausgearbeitet, die sich durch die Attribute ihrer einzuordnenden Objekte unterscheiden. Falls jedoch eine Klassifikation von Objekten bereits erstellt worden ist oder Hypothesen für Gruppierungen von Objekten vorliegen, können im zweiten Schritt des Data Mining die Strukturen der Klassen durch die Musterbeschreibung genau bestimmt werden. Die Verfahren der Datenmustererkennung (Unüberwachte Verfahren) dienen also dazu, ähnliche Datenobjekte zu Mustern und Gruppen zusammenzufassen, so dass diese Datenobjekte klassifiziert werden.<sup>681</sup> Ihr Ziel ist, interessante Beziehungen und Zusammenhänge in noch unstrukturierten Datenbeständen zu entdecken.<sup>682</sup> Zu den wichtigsten Verfahren der Datenmustererkennung zählen die Clusteranalyse, das Bayes- Verfahren und die Fuzzy-Datenanalyse.<sup>683</sup>

Die Verfahren der Datenmusterbeschreibung (Überwachte Verfahren) haben dagegen die Aufgabe, ein erzeugtes (deswegen auch bekanntes) Muster (Z. B. Kundengruppen) oder eine Hypothese über Zusammenhänge und Regelmäßigkeiten heranzuziehen und mittels Algorithmen möglichst genau zu beschreiben.<sup>684</sup> Dabei wird jedoch vorausgesetzt, dass dem jeweiligen Algorithmus nur Datenobjekte zur Verfügung gestellt werden, die jeweils einem bestimmten Muster entsprechen (Z. B. Kunden in einer bestimmten Kundengruppe).<sup>685</sup> Ansonsten würde das Data Mining versagen.

Überwachte Verfahren werden häufig nach unüberwachten Untersuchungen eingesetzt, um die Datenmustererkennung mit der Datenmusterbeschreibung zu kombinieren.<sup>686</sup> So könnte man z. B. zunächst eine Kundensegmentierung (unüberwacht) durchführen und für die ermittelten Kundensegmente anschließend Klassifizierungsregeln (Z. B. mit Hilfe eines Entscheidungsbaums) erzeugen lassen (überwacht). Zu den wichtigsten Verfahren der Musterbeschreibung gehören statistische Verfahren (Regressionsanalyse, Varianzanalyse usw.), Ansätze des maschinellen Lernens.<sup>687</sup>

---

<sup>680</sup> Vgl. Mertens/Wieczorrek (2000), S. 215.

<sup>681</sup> Vgl. Neckel/Knobloch (2005), S. 187.

<sup>682</sup> Vgl. Neckel/Knobloch (2005), S. 187 und Krahl/Windheuser/Zick (1998), S. 78.

<sup>683</sup> Vgl. Mertens/Wieczorrek (2000), S. 216.

<sup>684</sup> Vgl. Mertens/Wieczorrek (2000), S. 215.

<sup>685</sup> Vgl. Mertens/Wieczorrek (2000), S. 215.

<sup>686</sup> Vgl. Neckel/Knobloch (2005), S. 187.

<sup>687</sup> Vgl. Mertens/Wieczorrek (2000), S. 219.

- **Überwachte (Lern)Verfahren**

Bei diesen Verfahren werden Objekte anhand eines sogenannten „Klassifikators“ in unterschiedlichen Gruppen erfasst. Ein Klassifikator kann als ein System verstanden werden, welches Datenobjekte in vorgegebene Klassen einordnet.<sup>688</sup>

Grundsätzlich gibt es vier Typen von Klassifikatoren:<sup>689</sup> „Binär“ (Betrachtung von genau zwei unterschiedlichen Klassen, Beispiel: Kreditwürdig, Nicht-Kreditwürdig), „Mehrfach“ (Betrachtung von mehr als zwei, aber weiterhin einer beschränkten Anzahl von Klassen, Beispiel: Firmenratings von Aaa bis zu Caa), „Rekonstruktion einer Funktion“ (Beliebige Ausprägungen der Zielgröße innerhalb eines vorgegebenen Intervalls möglich, Beispiel: Wahrscheinlichkeit eines Kreditausfalls mit beliebigen Werten zwischen 0% und 100%), „Prognose“ (Zusätzlich zu den erstgenannten Typen wird noch eine zusätzliche zeitliche Komponente berücksichtigt, Beispiel: Aktien- und Devisenkurs-Vorhersagen).

Diese Klassifikatoren sind in der Lage, Regeln aufzustellen, mit deren Hilfe auch neue Objekte in bereits vorhandene Klassen eingeordnet werden können. Dazu wird allerdings das Vorhandensein bereits bekannter Fälle vorausgesetzt;<sup>690</sup> das heißt, zur Regelerzeugung werden solche Objekte herangezogen, deren Zuordnung zu Klassen bereits bekannt ist (Trainingsset). Das ausgewählte Modell muss nun so konfiguriert bzw. angepasst werden, dass die bekannten Fälle möglichst gut reproduziert werden können. Man spricht in diesem Zusammenhang vom Training des Modells.<sup>691</sup>

Der Trainingsprozess ist iterativ, wobei in jedem Schritt das aktuelle Modell auf die bekannten Daten angewendet und der noch vorhandene Klassifikationsfehler ausgewertet wird. Die Fehleranalyse führt dann zu Maßnahmen, die das Modell bzw. das System verbessern. Diese Verbesserung des Modells entspricht einem sogenannten Lernschritt. Das Modell lernt, die Trainingsdaten besser nachzubilden.

Gleichzeitig generiert der Klassifikator seine eigenen Regeln oder Strukturen. Nach Abschluss des Trainingsprozesses kann der Klassifikator auf noch nicht klassifizierte Daten angewendet werden. Das entspricht sogar dem eigentlichen Zweck des Lernprozesses. Somit dient das Gesamtsystem als „Generalisator“ für weitere Untersuchungen, da die Regeln oder Strukturen der Trainingsdaten auf weitere Daten generalisiert bzw. verallgemeinert werden können.

---

<sup>688</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 61.

<sup>689</sup> Vgl. ebd., S. 61.

<sup>690</sup> Vgl. ebd., S. 62.

<sup>691</sup> Vgl. ebd., S. 62.

Zur Qualitätskontrolle des Modells wird das Training durch eine Testmenge überprüft, welche die durch das Training ermittelten Regeln verifizieren soll. Die Testmenge beinhaltet weitere, bereits klassifizierte Objekte, die jedoch keinen direkten Einfluss auf die Modifizierung des Modells im Training haben.

Zusammenfassend kann man überwachtes Verfahren durch folgende Merkmale charakterisieren:<sup>692</sup>

- Das Ergebnis, das durch das Modell abgebildet werden soll, ist a priori bekannt.
- Die Datensätze des Analysedatenbestandes enthalten jeweils eine bekannte Variable, in der die Klassenzugehörigkeit kodiert ist.
- Durch das überwachte Verfahren werden die Modellparameter derart geändert, dass die durch das Modell generierten Klassifikationswerte mit den tatsächlichen Klassifikationswerten größtenteils übereinstimmen.
- Werden zu viele Objekte durch das Modell fehlerhaft klassifiziert, so müssen die Parameter angepasst werden.

- **Unüberwachte (Lern)Verfahren**

Bei unüberwachten Verfahren geht es darum, wie oben bereits erwähnt, interessante Strukturen in einem noch unstrukturierten Datenbestand zu entdecken. Dabei gibt es im Wesentlichen zwei Sichtweisen: Segmentierung und Assoziationen.<sup>693</sup>

Bei einer *Segmentierung* wird nach einer globalen Strukturierung der Daten gesucht.<sup>694</sup> Das Ziel ist dabei, die Daten in Cluster zu partitionieren, d. h. jeder Datensatz wird einem Segment zugeordnet. Es liegen zunächst nur die Input-Erkenntnisse vor, und die Output-Werte sind erst bei der Analyse zu ermitteln. Im Gegensatz zu überwachten Verfahren werden hier deshalb keine spezifischen Klasseneinteilungen vorgegeben. Es ist die Aufgabe des unüberwachten Verfahrens, die Ähnlichkeiten der vorhandenen Objekte selbstständig zu ermitteln und somit sämtliche Vorschläge für die Klasseneinteilungen zu erzeugen. Die Voraussetzung hierfür ist jedoch, dass alle Objekte des Datenbestandes den gleichen Informationsumfang haben sollen (z. B. Kundenstammdaten).

Zur Segmentierung wird das Cluster-Analyseverfahren angewendet.

---

<sup>692</sup> Vgl. Beekmann/Chamoni (2006), S. 267.

<sup>693</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 78.

<sup>694</sup> Vgl. ebd., S. 78.

Bei der Suche nach *Assoziationen* werden solche Objekte betrachtet, die einen vergleichbaren Informationsumfang besitzen (Z. B. im Fall einer Kassenbonnanalyse).<sup>695</sup> Die Untersuchungsgegenstände sind häufig auftretende (versteckte) Regeln oder Muster in den Daten (Z. B. immer wiederkehrende Kombinationen von Artikeln in Einkaufs- bzw. Warenkörben eines Supermarktkundens). Im Gegensatz zu globalen Aussagen bei der Segmentierung werden hier Aussagen über partielle Strukturen in den Daten gemacht.<sup>696</sup> Ein mögliches Ergebnis einer Kassenbonnanalyse ist z. B. folgende Formulierung: „An einem Freitag wird in 20% der Fälle, in denen Babywindeln gekauft werden, auch Bier gekauft.“ Man kann sofort erkennen, dass es kaum möglich ist, anhand dieser Erkenntnis eine sinnvolle Segmentierung der Kunden nach diesem Kriterium durchzuführen.<sup>697</sup> Dennoch liefern viele solcher Assoziationen wertvolles Wissen über das Kundenverhalten.

Abschließend lässt sich unüberwachtes Verfahren wie folgt zusammenfassen:<sup>698</sup>

- Es werden keine Vorgaben für das durch ein Muster dargestellte Ergebnis spezifiziert.
- Die Datensätze werden unabhängig von Ergebniserwartungen verarbeitet.
- Die Modellparameter werden beispielsweise dann angepasst, wenn die Muster sich wiederholen.

Nach dieser Unterscheidung zwischen den überwachten und unüberwachten Verfahren können nun die Data Mining-Methoden bezüglich der wichtigsten Aufgabenbereiche vorgestellt werden.

- **Klassifikationsmethoden**

Die Grundlage der Klassifikationsmethoden bildet ein Klassifikationsmodell, welches die Zuordnung von Elementen in vorgegebene Klassen abbildet.<sup>699</sup> Dieses Modell kann zur Prognose der Klassenzugehörigkeit von Datenobjekten mit unbekannter Klassenzugehörigkeit eingesetzt werden, indem ein solches Datenobjekt anhand seiner Merkmale einer von mehreren alternativen Klassen zugeordnet wird.<sup>700</sup> Die Erstellung

---

<sup>695</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 79f.

<sup>696</sup> Vgl. ebd., S. 80.

<sup>697</sup> Das Cluster-Analyseverfahren eignet sich hier nicht für die Analyse, da die nicht belegten Merkmale (Z. B. die nicht gekauften Artikel in den Einkaufskörben oder die nicht in Anspruch genommenen Produkte einer Bank) das Clustering verhindern würden, vgl. ebd., S. 80.

<sup>698</sup> Vgl. Beekmann/Chamoni (2006), S. 267.

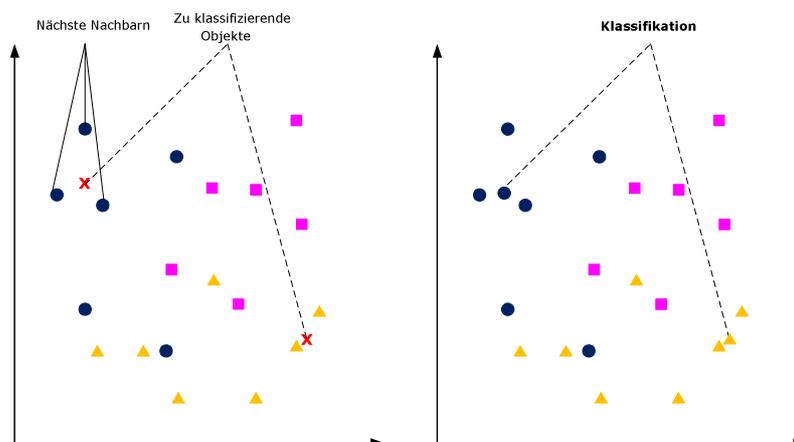
<sup>699</sup> Vgl. ebd., S. 264.

<sup>700</sup> Vgl. ebd., S. 264.

des Modells basiert auf einer bereits bestehenden Menge an Datenobjekten, deren Klassenzugehörigkeiten bereits bekannt sind. Ein typisches Beispiel für die Anwendung eines Klassifikationsmodells ist eine Kreditwürdigkeitsprüfung.<sup>701</sup> Hierbei wird aus einem existierenden Datenbestand ein Regelsystem generiert, so dass ein neuer Kunde, der einen Kredit beantragt, anhand seiner in das Regelsystem eingegebenen Kundendaten als kreditwürdig bzw. als nicht kreditwürdig klassifiziert wird.

In der Praxis werden für die Klassifikationsaufgabe im Rahmen des Data Mining unterschiedliche Methoden verwendet.<sup>702</sup> Diese reichen von Diskriminanzanalysen, K-Nächste Nachbarn-Verfahren über fallbasiertes Schließen, Entscheidungsbäume und Regelinduktionsverfahren bis hin zu Künstliche Neuronale Netze und Fuzzy-Logik basierte Verfahren.

Die *K-Nächste Nachbarn-Methode (KNN)* basiert auf dem Prinzip „Lernen durch Analogien“.<sup>703</sup> Es klassifiziert einen Datensatz auf der Grundlage der  $k$  nächsten Nachbarn.  $K$  ist dabei eine ganze, positive Zahl. Zu einem zu klassifizierenden Datensatz wird die Anzahl der  $k$ -ähnlichsten Datensätze ausgewählt und aus deren Klassifikation der neue Datensatz bewertet.



**Abbildung 38: 3-nächste-Nachbarn-Methode (k=3)**

Quelle: Krahl/Windheuser/Zick (1998), S. 75.

Möchte man z. B. einen neuen Bankkunden den Klassen „kreditwürdig“ oder „nicht-kreditwürdig“ zuordnen, versucht man aufgrund seiner Merkmalswerte (Z. B. Einkommen und Alter) die  $k$ -ähnlichsten Kunden zu finden. Falls die Mehrheit der ähnlichsten Kunden in der zugrunde liegenden Datenbasis zur Klasse kreditwürdig (oder nicht-kreditwürdig) zählt, so wird der neue Kunde auch als kreditwürdig (oder

<sup>701</sup> Vgl. Beekmann/Chamoni (2006), S. 265.

<sup>702</sup> Eine ausführlichere Beschreibung dieser Methoden findet sich z. B. bei Nakhaeizadeh (1998), S. 11-16, Krahl/Windheuser/Zick (1998), S. 64-78 und Gluchowski/Gabriel/Dittmar (2008), S. 197-201.

<sup>703</sup> Vgl. Nakhaeizadeh (1998), S. 12.

nicht-kreditwürdig) bewertet bzw. klassifiziert. Die Abbildung 38 veranschaulicht die Grundidee dieses Verfahrens für  $k=3$ .

Eine andere Methode, die zur Klassifikation eingesetzt werden kann, ist das *fallbasierte Schließen* (Case-Based-Reasoning, CBR).

Die Grundidee dieser Methode besteht darin, Beispiele bereits gelöster Probleme aus einem bestimmten Anwendungsbereich abzuspeichern, um sie später zur Lösung neuer, aber ähnlich gearteter Probleme heranziehen zu können.<sup>704</sup> Dazu werden alle Beispielprobleme zusammen mit ihrer Lösung in einem geeigneten Format dargestellt und in einer speziellen Datenbank, der Fallbasis, abgespeichert. Dabei wird ein Paar, bestehend aus Problembeschreibung und entsprechender Lösung, als Fall bezeichnet. Tritt ein neues Problem auf, soll dieses zunächst in dem vorgegebenen Format dargestellt werden. Dann können die Problembeschreibungen der in der Fallbasis vorhandenen Fälle durchsucht werden, um eine dem neuen Problem möglichst ähnliche Beschreibung zu finden. Dafür benötigt man ein Ähnlichkeitsmaß für die Problembeschreibungen. Der Fall mit der ähnlichsten Problembeschreibung wird genommen und seine Lösung als eine mögliche Lösung des neuen Problems ausgegeben. Eventuell ist es notwendig, die vorgeschlagene Lösung an die neuen Gegebenheiten anzupassen.

Ein Verfahren, das ebenfalls häufig bei der Klassifikation angewendet wird, ist die *Diskriminanzanalyse*.

Die Diskriminanzanalyse untersucht einen Datenbestand auf solche Attribute, die einen hohen Erklärungsgrad für eine bereits vorgegebene Klassifikation (Z. B. kreditwürdig, nicht-kreditwürdig) besitzen.<sup>705</sup> Das heißt, sie sucht die Unterschiede zwischen den vorgegebenen Gruppen und ermittelt die Attribute, die möglichst viel zur optimalen Trennung dieser Gruppen (Diskrimination) beitragen.<sup>706</sup> Das Ziel ist dabei, mit Hilfe der geschätzten Diskriminanzfunktion ein neues Datenobjekt, dessen Gruppenzugehörigkeit nicht bekannt ist, aufgrund seiner charakteristischen Merkmalsausprägungen einer Gruppe zuzuordnen.

So wird bei der univariaten Diskriminanzanalyse (UDA) das trennfähigste Attribut ausgewählt und der sogenannte „Cut-off-point“ bestimmt.<sup>707</sup> Die Klassifikation eines Datensatzes ist dann sehr leicht zu bestimmen, je nachdem, ob seine

---

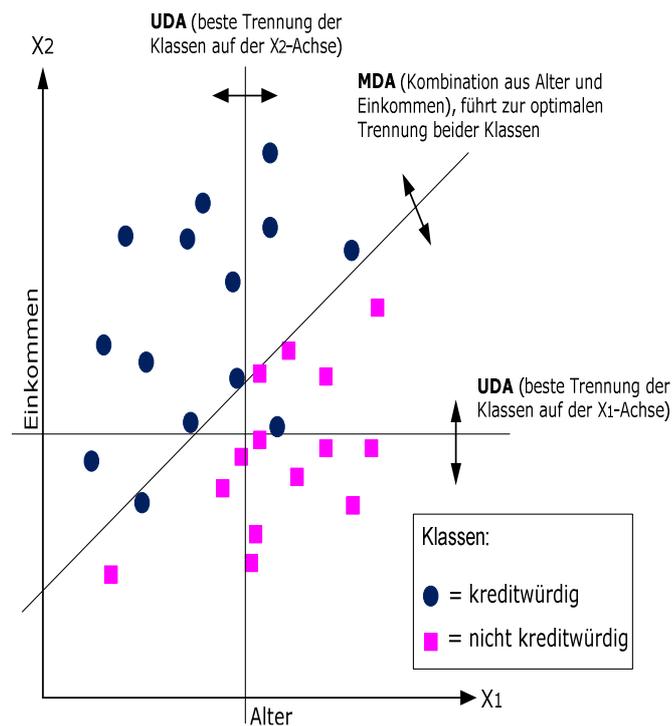
<sup>704</sup> Vgl. Dilger (2007), S. 27.

<sup>705</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 75.

<sup>706</sup> Vgl. Eckey/Kosfeld/Rengers (2002), S. 289.

<sup>707</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 75.

Attributausprägung oberhalb oder unterhalb dieses Cut-off-point liegt. Bei der multivariaten Diskriminanzanalyse (MDA) wird hingegen aus mehreren Attributen eine lineare Diskriminanzfunktion (auch Klassifikationsfunktion genannt) ermittelt.<sup>708</sup> Diese ergibt sich aus der gewichteten Summe aller ausgewählten, aussagefähigen Merkmale (Z. B. Alter, Einkommen von Kunden bei einer Kreditwürdigkeitsprüfung), die in einer linearen Funktion zusammengefasst sind. Es wird ein Cut-off-point bezüglich der Diskriminanzfunktion berechnet. Die Datensätze werden dann wie im univariaten Fall hinsichtlich deren Lage zum Cut-off-point klassifiziert.



**Abbildung 39: Diskriminanzanalyse**

Quelle: Krahl/Windheuser/Zick (1998), S. 76.

Die Diskriminanzanalyse wird vor allem für die Binärklassifikation (wie auch bei der Kreditwürdigkeitsprüfung der Fall ist) genutzt. Eine vereinfachte Version der Diskriminanzanalyse für Binärklassifikationsaufgaben ist die einer automatisierten Modellbildung mittels linearer Regressionen.<sup>709</sup> Hier wird die Diskriminanzfunktion selbsttätig hinsichtlich der Einflussgrößen und ihrer Gewichtung optimiert.

Eine *Regressionsanalyse* basiert auf einem Modell, mit dem eine abhängige, stetige Variable durch mehrere unabhängige Variablen erklärt wird.<sup>710</sup> Somit wird die Beeinflussbarkeit einer abhängigen Variablen von mehreren unabhängigen Variablen

<sup>708</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 75.

<sup>709</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 76.

<sup>710</sup> Vgl. Beekmann/Chamoni (2006), S. 265.

untersucht bzw. prognostiziert.<sup>711</sup> Ein solches Modell der Prognose zielt im Unterschied zu einem reinen Klassifikationsmodell auf die stetigen Variablen, so dass nicht mehr möglich wird, speziell von den einzelnen Klassen im Klassifikationsmodell auszugehen.<sup>712</sup>

Die Zielvariable ist bei der Prognose im Gegensatz zur Klassifikation stetig bzw. kontinuierlich. Diskretisiert man diese stetige Variable und verwendet man die gebildeten Intervalle als Klassenwerte, so kann das Klassifikationsmodell und damit alle Klassifikationsverfahren auch zur Prognose verwendet werden.<sup>713</sup>

In einer Diskriminanzanalyse wird eine nominalskalierte Variable durch eine oder mehrere metrisch skalierten Variablen erklärt. Dagegen gestattet die Regressionsanalyse, dass auch die erklärte Variable metrisch skaliert werden kann. Somit lässt sich leicht ermitteln, wie stark z. B. der Umsatz von Marketingausgaben abhängt. Einen Sonderfall stellt jedoch hier die Binärgruppierung der Diskriminanzanalyse, wobei die erklärte Variable binär ist und sich wie eine metrisch skalierte Variable behandeln lässt.<sup>714</sup> Dies macht deutlich, dass die Diskriminanz- und Regressionsanalysen im Wesentlichen übereinstimmen.

*Baumbasierte Verfahren* (Entscheidungsbäume) sind typische Modelle, die zur Klassifikation eingesetzt werden. Sie sind zuerst in den 60er Jahren unabhängig voneinander von den Disziplinen des maschinellen Lernens und der Statistik entwickelt worden. Sie wurden ständig von den Forschern des maschinellen Lernens verbessert und zählen heute in den meisten Data Mining-Werkzeugen zum Standard.<sup>715</sup> Im Gegensatz zur Diskriminanzanalyse geht man bei den baumbasierten Verfahren aus der Grundannahme vor, dass für jede Klasse (kreditwürdig, nicht-kreditwürdig)

<sup>711</sup> Im einfachsten Fall der linearen Regressionsanalyse betrachtet man zwei Variablen  $x$  und  $y$ , die vermutlich ungefähr in einem linearen Zusammenhang stehen:  $Y \approx a + bX$ . Dabei sind  $X$  als unabhängige und  $Y$  als abhängige Variable definiert. Es existieren von  $x$  und  $y$  je  $n$  Beobachtungen  $X_i$  und  $Y_i$ , wobei  $i$  von 1 bis  $n$  geht. Der funktionale Zusammenhang  $Y = f(x)$  zwischen  $x$  und  $Y$  kann nicht exakt festgestellt werden, da die Funktion  $f(x)$  von einer Störgröße  $\mathcal{E}$  überlagert wird.

Diese Störgröße ist als Zufallsvariable konzipiert, die nichterfassbare Einflüsse (menschliches Verhalten oder Messungenauigkeiten oder ähnliches) darstellt. Es ergibt sich also das Modell:  $Y = a + bX + \mathcal{E}$  oder genauer  $y_i = a + bx_i + \mathcal{E}_i$ .

Das Residuum  $\mathcal{E}_i$  gibt die Differenz zwischen der Regressionsgerade  $a + bx_i$  und den Messwerten  $y_i$  an. Des Weiteren bezeichnet man mit  $\gamma_i$  den Schätzwert für  $y_i$  und es gilt  $\gamma_i = a + bx_i$  und  $\mathcal{E}_i = y_i - \gamma_i$ . Um eine lineare Gerade als die Regressionsgerade zu definieren, benutzt man meistens die Methode der kleinsten Quadrate, die die Quadratsumme der Residuen minimiert, also die Abweichungen  $\mathcal{E}_i = y_i - \gamma_i$  der Beobachtungswerte  $(X_i, Y_i)$  von dieser Gerade:  $\sum \mathcal{E}_i^2 = \sum (y_i - (a + bx_i))^2 \rightarrow \min !$ . Diese minimierte Zielfunktion wird dann dazu verwendet, um die Regressionskoeffizienten  $a$  und  $b$  zu berechnen.

<sup>712</sup> Vgl. Beekmann/Chamoni (2006), S. 265.

<sup>713</sup> Vgl. Nakhaeizadeh (1998), S. 18, Beekmann/Chamoni (2006), S. 264-265

<sup>714</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 77.

<sup>715</sup> Vgl. Beekmann/Chamoni (2006), S. 268.

bestimmte Attribute oder Attributskombinationen verantwortlich sind. Diese Annahme erfordert eine systematische Untersuchung der Auswirkungen einzelner Attribute auf die zu erklärende Variable (Zielvariable).<sup>716</sup>

Die baumbasierten Techniken gelten als überwachte Verfahren, da sie den Datenbestand zunächst in eine Trainings- und eine Testmenge unterteilen.<sup>717</sup> Mit der Trainingsmenge werden dann die relevanten Regeln ermittelt, um diese anschließend durch die Testmenge zu verifizieren. Durch die sukzessive Aufspaltung der Trainingsmenge in Teilmengen werden die zu klassifizierenden Objekte zu homogenen Gruppen, die sich durch bestimmte Attributskombinationen darstellen, zusammengefasst. Für eine erste Aufteilung der Trainingsmenge in Untermengen wird ein solches Attribut ausgewählt, das sich für die Klassifikation der Objekte als äußerst informativ erweist.<sup>718</sup> Die weiteren Attribute sollen dann über ein zusätzliches Erklärungspotential verfügen. Diese Auswahl wird i. d. R. mit Hilfe eines Entropiemaßes durchgeführt:<sup>719</sup> Die Entropie gibt Auskunft über den potentiellen Informationsgehalt eines Attributs. Je kleiner die Entropie ist, desto höher ist der Informationsgehalt. Abhängig von dem Grad seines Informationsgehalts in Bezug auf eine Zielvariable findet sich dann ein Attribut eher weit oben (wenn der Informationsgehalt hoch ist) oder unten (wenn der Informationsgehalt gering ist) im Entscheidungsbaum.

Die Regel, nach welcher die Aufteilung der Trainingsmenge in Teilmengen durchgeführt wird, ist ein wesentliches Unterscheidungsmerkmal der baumbasierten Verfahren.

Nach den Aufteilungsregeln können die folgenden drei wichtigsten Klassen des Verfahrens differenziert werden:<sup>720</sup> Die erste Verfahrensklasse basiert auf den „Concept Learning Systems“ (CLS), woraus dann der bekannte ID3-Algorithmus und der C4.5-Algorithmus entwickelt wurden. Eine zweite Klasse entstammt dem Automatic-Interaction-Detection-Verfahren (AID), welches durch den THAID-Algorithmus und den bekannten CHAID-Algorithmus erweitert wurde. Als die dritte Klasse kann die CART-Gruppe (classification and regression trees) genannt werden, deren Algorithmen zwar nur binäre Aufteilungen zulassen, dafür jedoch auch für die Erstellung eines Regressionsmodells einsetzbar sind.

---

<sup>716</sup> Vgl. Nakhaeizadeh (1998), S. 14.

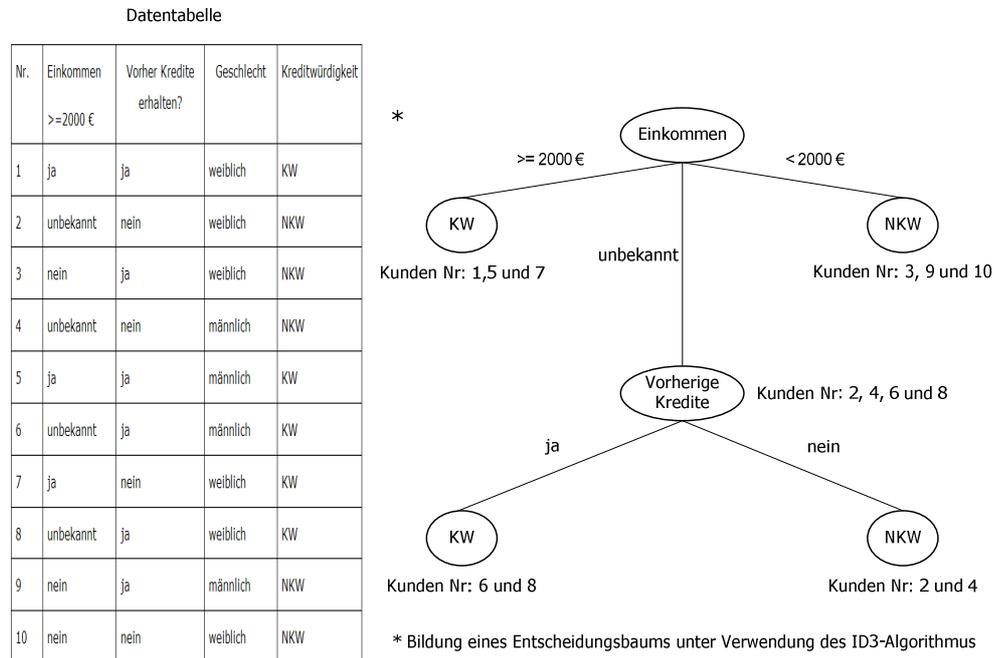
<sup>717</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 197.

<sup>718</sup> Vgl. Nakhaeizadeh (1998), S. 14 und Krahl/Windheuser/Zick (1998), S. 69.

<sup>719</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 71.

<sup>720</sup> Vgl. Beekmann/Chamoni (2006), S. 270.

Die Abbildung 40 zeigt ein Beispiel, wie mit Hilfe eines Entscheidungsbaums 10 Bankkunden nach ihrer Kreditwürdigkeit klassifiziert werden können.



**Abbildung 40: Entscheidungsbaum für die Kreditwürdigkeitsdaten**

Quelle: Nakhaeizadeh (1998), S. 14-15.

Die Datentabelle beinhaltet Informationen über die Kreditwürdigkeit dieser Kunden. Neben dem Einkommen und Geschlecht kann man aus der Tabelle auch Informationen darüber erhalten, ob der Kreditnehmer vorher schon Kredite erhalten hat oder nicht.

Das Ziel dieses Beispiels besteht darin, Regeln zu finden, welche aufgrund der ersten drei Attribute eine deutliche Trennung zwischen den kreditwürdigen und nicht-kreditwürdigen Kunden ermöglichen. Diese Regeln können dann zur Feststellung der Kreditwürdigkeit von neuen Bankkunden verwendet werden.

Das Einkommen ist hier das informativste Attribut, und davon ausgehend bildet man sukzessiv einen Entscheidungsbaum. Zunächst werden die 10 Datenobjekte (Kunden) in drei Untergruppen aufgeteilt. Da sich die Elemente in zwei von diesen Untergruppen bereits klassifiziert haben, wird dort der Baum nicht weiter gegliedert.

Bei der dritten Untergruppe gehören die Elemente zu zwei verschiedenen Klassen. Daher werden bei dieser Untergruppe die vier übriggebliebenen Objekte weiter in zwei Untergruppen aufgeteilt. Das geschieht durch die Auswahl des Attributs „vorherige Kredite“, das am informativsten ist. Die Elemente der neuen Untergruppen gehören nun jeweils zu einer Klasse, und aus diesem Grund wird das Verfahren beendet.

Aus dem generierten Entscheidungsbaum lassen sich vier Regeln ableiten:

- *Regel 1:* Wenn das Einkommen größer oder gleich € 2000 ist, ist der Kunde kreditwürdig.
- *Regel 2:* Wenn das Einkommen kleiner als € 2000 ist, ist der Kunde nicht kreditwürdig.
- *Regel 3:* Wenn das Einkommen unbekannt ist und der Kunde in der Vergangenheit schon Kredite erhalten hat, ist der Kunde kreditwürdig.
- *Regel 4:* Wenn das Einkommen unbekannt ist und der Kunde in der Vergangenheit keine Kredite erhalten hat, ist der Kunde nicht kreditwürdig.

Durch diese Entscheidungsregeln können die Informationen über Objekte und deren Klassen übersichtlich dargestellt werden. Die generierten vier Regeln ermöglichen im obigen Beispiel eine deutliche Trennung zwischen den kreditwürdigen und nicht kreditwürdigen Kunden. Außerdem können sie zur Klassifikation der neuen Kunden verwendet werden.

Im Gegensatz zum obigen Beispiel erhält man jedoch in realen Anwendungen sehr große Entscheidungsbäume, weil mehr Daten für die Analyse zur Verfügung stehen. Zudem sind diese Daten meistens verrauscht. Um die prognostische Leistung von solchen großen Entscheidungsbäumen zu gewährleisten, verwendet man das sogenannte „Pruningverfahren“.<sup>721</sup> Dieses Verfahren basiert auf dem Abschneiden weniger wichtiger Äste eines Entscheidungsbaums, was die Gefahr des sogenannten „Overfitting“ (Überanpassung) beseitigt und somit auch zu besseren Klassifikationsergebnissen führt.<sup>722</sup> Das Pruning dient ebenfalls dazu, den Baum so einfach wie möglich zu halten.<sup>723</sup>

Fasst man die oben angeführten Überlegungen zusammen, so wird es deutlich, dass die Entscheidungsbäume als Systeme zur Regelinduktion angesehen werden können. Im Fall von sehr großen Datenmengen ist eine Regelinduktion jedoch schwer durchzuführen, da die zufälligen Elemente (wie z. B. Missing Values oder Rauschen) den Entscheidungsbaum zur Übermodellierung zwingen und somit das Erkennen von

---

<sup>721</sup> Die mögliche Größe der Entscheidungsbäume kann sich negativ auswirken. Jede einzelne Regel ist zwar leicht abzulesen, den Gesamtüberblick zu haben ist jedoch sehr schwierig. Es wurden deshalb so genannte Pruning-Methoden entwickelt, welche die Entscheidungsbäume auf eine vernünftige Größe kürzen. Beispielsweise kann man die maximale Tiefe der Bäume beschränken oder eine Mindestanzahl der Objekte pro Knoten festlegen, vgl. Krahl/Windheuser/Zick (1998), S. 73.

<sup>722</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 198 und Nakhaeizadeh (1998), S. 16.

<sup>723</sup> Vgl. Beekmann/Chamoni (2006), S. 270.

scharfen Regeln erschweren. Wegen der Übergröße verliert man oft den Gesamtüberblick über den Entscheidungsbaum. Aus diesen Gründen wäre es denkbar, den Entscheidungsbaum als Wissensrepräsentation nur als einen Zwischenschritt zu benutzen und mit den entsprechenden (deduktiven) Verfahren eine effizientere Darstellung des Regelwerkes zu generieren.<sup>724</sup> Um zu den Regeln zu gelangen, können durch verschiedene Verfahren viele unterschiedliche Entscheidungsbäume generiert und häufig auftretende Regeln dabei extrahiert werden.<sup>725</sup> Die überlagerten Optimierungen liefern dann einen robusten, allgemeinen und korrekten Regelsatz.<sup>726</sup> Ein solches Verfahren, das aus einem anderen Verfahren hervorgeht und dessen Ergebnisse weiterverarbeitet, wird „hybrid“ genannt. Ein großer Vorteil von Entscheidungsbäumen ist, dass sie gut erklärbar und nachvollziehbar sind. Außerdem kann ihre Generierung schnell durchgeführt werden.

Die *regelbasierten* Verfahren versuchen dagegen, die Regeln direkt zu generieren, also nicht über einen Entscheidungsbaum.<sup>727</sup> Sie basieren auf Regeln der Art „Wenn A, dann B“. Solche Regeln stellen eher allgemeine Gesetze dar, aus welchen Schlussfolgerungen für konkrete Situationen gezogen werden sollen. Sie sind leicht verständlich und interpretierbar, was ihre Handhabbarkeit bei der Analyse sehr erleichtert. Zu den regelbasierten Algorithmen gehören vor allem die AQ-Familie und der CN2-Algorithmus.

*Künstliche neuronale Netze* entstammen dem Gebiet der Künstlichen Intelligenz und verfolgen das Ziel, durch die Nachbildung der Arbeitsweise des menschlichen Gehirns, welches aus vielen Neuronen und deren Verknüpfungen besteht, die Aufgaben des KDD (Knowledge Discovery in Databases) zu bearbeiten. Sie können für die Erstellung eines Klassifikationsmodells sowie eines Regressionsmodells und für die Segmentierung (Clustering) angewendet werden.<sup>728</sup>

Ein künstliches neuronales Netz besteht aus einer Menge von Verarbeitungselementen bzw. Neuronen und deren zugehörigen Verbindungen (biologisch: Synapsen). Die Verbindungen dienen dazu, die Ausgabewerte von Neuronen zu gewichten und diesen gewichteten Wert als Eingabewert an die verbundenen Neuronen weiterzuleiten. Ein Neuron besteht aus einer Eingabefunktion, einer Aktivierungsfunktion und einer Ausgabefunktion.

---

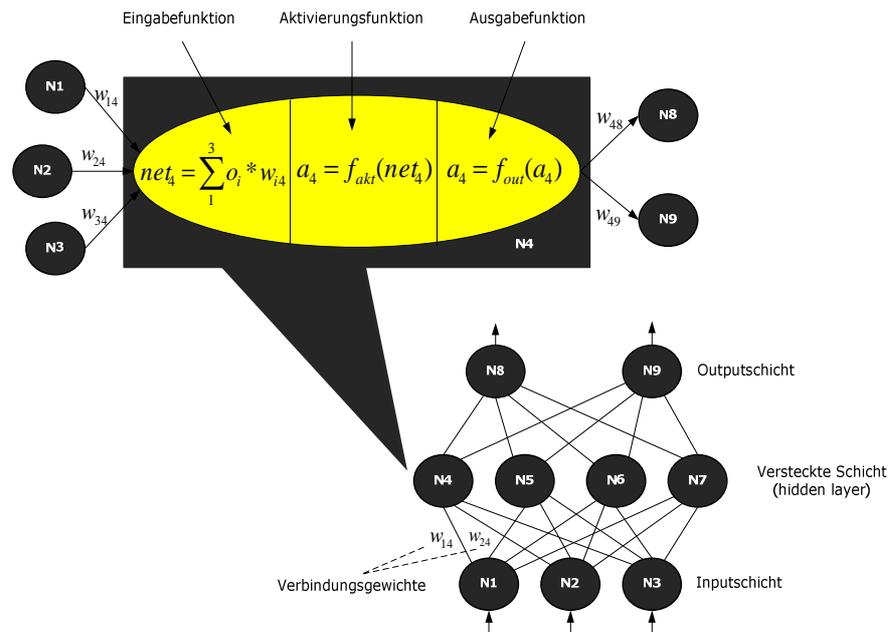
<sup>724</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 74.

<sup>725</sup> Nachteilig ist dabei aber, dass die Regeln in keinerlei Beziehungen zueinander stehen und dass widersprüchliche Regeln erzeugt werden können, vgl. Krahl/Windheuser/Zick (1998), S. 74.

<sup>726</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 74.

<sup>727</sup> Vgl. Nakhaeizadeh (1998), S. 16.

<sup>728</sup> Vgl. Beekmann/Chamoni (2006), S. 271.



**Abbildung 41: Aufbau eines Künstlichen Neuronales Netzes**

Quelle: Gluchowski/Gabriel/Dittmar (2008), S. 199.

Durch die Verbindungen erhält ein Neuron mehrere gewichtete Eingabewerte (Ein Eingabewert besteht aus einem Inputsignal  $o$  und einem dazugehörigen Verbindungsgewicht  $w$ ) von vorgelagerten Neuronen. Die Eingabefunktion  $net$  summiert alle diese Werte auf, und aus diesem neuen (Reiz-)Wert wird anschließend mit einer Aktivierungsfunktion  $f_{akt}$  ein Aktivierungszustand des Neurons  $a$  berechnet, der wiederum in eine Ausgabefunktion  $f_{out}$  mündet und den Ausgabewert des Neurons  $o$  bestimmt. Dieser Wert kann wieder über die Verbindungen an ein nachfolgendes Neuron als dessen Eingabewert weitergeleitet oder als Ausgabewert eines Netzes genutzt werden. Diese Eingabe-Ausgabe-Beziehungen zwischen den Neuronen ermöglichen eine Gruppierung der Verarbeitungseinheiten. Die Einheiten können in eine Eingabe- bzw. Inputschicht, eine Ausgabe- bzw. Outputschicht und eine oder mehrere Zwischenschichten bzw. versteckte Schichten (hidden layers) logisch angeordnet werden.

Die Entscheidung darüber, wann ein Neuron die Eingabewerte biologisch gesehen als einen Reiz empfindet und dementsprechend eine Reaktion auslöst, die dann den Ausgabewert bestimmt, wird durch die Aktivierungsfunktion gesteuert. Es gibt hauptsächlich zwei verschiedene Typen der Aktivierungsfunktionen:<sup>729</sup> Die Sigmoid-Funktionen und die Radial Basis-Funktionen. Während bei der Sigmoid-Funktion ein Neuron erst dann zu einer Reaktion veranlasst wird, wenn der aus summierten

<sup>729</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 65f.

Eingabewerten resultierende Input einen bestimmten Schwellenwert übersteigt, wird bei der Radial Basis-Funktion das Neuron bereits eine Reaktion veranlassen, wenn sich der Input in der Nähe eines bestimmten Schwerpunktes befindet. Während bisher in der Biologie kein Typ der Radial Basis-Funktion entdeckt werden konnte, ist seine Verwendung aus mathematischer Sicht sinnvoll, denn sie führt bei bestimmten Aufgaben, z. B. bei Prognoseaufgaben, zu besseren Ergebnissen.

Die Erstellung eines künstlichen neuronalen Netzes erfolgt durch die Vorgabe der Lernregeln und einer Grundstruktur (Topologie), welcher den Typ, die Anzahl und die Anordnung der Neuronen sowie die Verbindungen mit deren Gewichte umfasst. Die Lernregeln bzw. -Verfahren dienen dazu, ein künstliches neuronales Netz dazu zu bringen, für die bestimmten Eingangsmuster die zugehörigen Ausgabemuster zu erzeugen. Das geschieht grundsätzlich auf drei verschiedenen Wegen, nämlich durch überwachtes (supervised), unüberwachtes (unsupervised) und bestärktes Lernen (reinforcement learning).<sup>730</sup>

Beim überwachten Lernen wird dem künstlichen neuronalen Netz ein Eingangsmuster gegeben und die Ausgabe, die das neuronale Netz in seinem aktuellen Zustand produziert, mit dem Wert verglichen, den es eigentlich ausgeben soll. Durch den Vergleich der Soll- und Ist-Ausgabe kann auf die vorzunehmenden Änderungen der Netzkonfiguration geschlossen werden. Es ist jedoch nicht immer möglich, zu jedem Eingabedatensatz den passenden Ausgabedatensatz zum Trainieren des Netzes zur Verfügung zu stellen. Die Entscheidung darüber, welchen Aktivierungszustand sich das Netz übernehmen soll, um die bestmögliche Reaktion als Ergebnis zu feuern, könnte dann dem Netz aufgrund des fehlenden Ausgabedatensatzes schwer fallen. In diesem Fall soll dem Netz mitgeteilt werden, dass es mit der Wahl seines Aktivierungszustandes bzw. -Ergebnisses richtig oder falsch liegt, oder ob mit einer entsprechenden Korrektur des Ergebnisses die richtige Lösung erreicht werden kann (bestärktes Lernen). Das unüberwachte Lernen erfolgt ausschließlich durch die Eingabe der zu lernenden Muster. Das Künstliche Neuronale Netz verändert sich entsprechend den Eingabemustern von selbst.

Grundsätzlich kann man zwischen zwei Arten von künstlichen neuronalen Netzen unterscheiden:<sup>731</sup> die vorwärts gerichteten künstlichen neuronalen Netze (auch feed forward networks genannt), die als Regressions- oder Klassifikationsmodell verwendet werden und Kohonen-Netze (bekannt als selbstorganisierende Karten).

---

<sup>730</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 200.

<sup>731</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 200.

Feed Forward Networks sind dadurch gekennzeichnet, dass die Verbindung der Neuronen nur in einer Richtung, und zwar von der Eingabeschicht über die Zwischenschicht zur Ausgabeschicht (also feed forward), erlaubt ist.<sup>732</sup> Netze, in denen auch Rückkopplungen zwischen Neuronen vorkommen, werden als rekurrente Netze bezeichnet.

Die Arbeitsweise von Feed Forward Networks kann wie folgt zusammengefasst werden:

Die Daten der einzelnen Merkmale eines jeden Datensatzes werden zunächst in die Eingabeneuronen eingegeben. Diese verarbeiten die Eingabewerte, ermitteln den Aktivierungszustand und geben die Ausgabewerte über die gewichteten Verbindungen an die Neuronen der Zwischenschicht weiter. In der Zwischenschicht wiederholt sich dieser Vorgang bis die Werte anschließend an das Neuron der Ausgabeschicht weitergegeben werden. Der Ausgabewert soll den durch das künstliche neuronale Netz bestimmten Wert der Klassenzugehörigkeit repräsentieren. Im Falle einer Abweichung von der tatsächlich vorliegenden Klasseneinteilung werden die Gewichte der Verbindungen durch induktive Lernalgorithmen so modifiziert bzw. angepasst, dass anschließend der Ausgabewert mit dem vorab bekannten und korrekten Ergebnis übereinstimmt und demzufolge die korrekte Klassenzugehörigkeit angibt. Die Anpassung bzw. Fehlerkorrektur erfolgt von hinten nach vorne: Die aufgetretenen Differenzen werden an der Ausgabeschicht in das Netz eingespeist, wodurch die einzelnen Gewichte angepasst bzw. modifiziert werden können. Diese Modifizierung nennt man „Adaption der Gewichte“, der unüberwachte Gesamtlernprozess wird als „Backpropagation“ bezeichnet.<sup>733</sup> Durch Backpropagation können die Werte der erklärenden Variablen (Zielvariablen) durch den Ausgabewert möglichst gut erklärt werden.

Im Gegensatz zu Feed Forward Networks bestehen Kohonen-Netze (bekannt auch als selbstorganisierende Karten) aus einer Eingabe- und einer meist zweidimensionalen (manchmal auch dreidimensionalen) Ausgabeschicht. Die Ausgabeschicht (Feature Map) beinhaltet eine Reihe von Neuronen (Knoten), die in einer Gitterstruktur angeordnet sind. Ein Neuron besitzt weder eine Aktivierungsfunktion noch einen

---

<sup>732</sup> Feed Forward Networks werden auch als Netze ohne Rücksprünge bezeichnet, welche ebenwise Verbindungen aufweisen. Die ebenwise verbundenen Netze weisen mehrere Schichten auf, wobei die einzelnen Neuronen jeweils nur Verknüpfungen zu Neuronen in den benachbarten Schichten aufbauen. Treten dagegen auch Verbindungen zwischen Neuronen auf, bei denen Ebenen übersprungen werden, dann werden diese als short cut connections bezeichnet. Außerdem sind komplexere Netze vorhanden, in denen auch Rückkopplungen zulässig sind (rekurrente Netze). Eine graphische Darstellung verschiedener Netzwerktopologien ist in Gluchowski/Gabriel/Dittmar (2008), S. 201 zu finden.

<sup>733</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 200 und Beekmann/Chamoni (2006), S. 272.

Schwellenwert und repräsentiert nur die Ähnlichkeit der zugehörigen Gewichte zum Input. Die Eingabewerte entsprechen den Attributsausprägungen ausgewählter Attribute der einzelnen Datenobjekte. Bei der Eingabe eines entsprechenden Datensatzes werden durch Weitergabe der gewichteten Werte der Eingabeneuronen alle Ausgabeneuronen aktiviert. Im unüberwachten Trainingsprozess des Kohonen-Netzes werden die Gewichte der Eingabeneuronen mit den Gewichtsvektoren der Ausgabeneuronen verglichen. Anschließend wird dasjenige Ausgabeneuron mit möglichst hoher Aktivierung (d. h. das Ausgabeneuron, dessen Gewichtsvektor dem Eingabevektor am ähnlichsten ist) als Repräsentant eines Clusters ermittelt, in dem der Datensatz liegt. Dieses Ausgabeneuron passt seine Gewichte an den Input an. Daraufhin werden die umliegenden Neuronen in der Nachbarschaft des ermittelten Neurons adaptiert, um eine automatische Anordnung (=Cluster) ähnlich reagierender Neuronen zu erreichen. Die Nachbarschaftsbetrachtung kommt aus den biologischen Betrachtungen der menschlichen Hirnrinde. Es ist schon festgestellt worden, dass immer für einen Rezeptorenbereich sämtliche Gruppen von Neuronen zuständig sind. Das bedeutet, dass nie ein Neuron allein auf einen Reiz reagiert, sondern mit ihm auch seine Nachbarschaft. Durch die Kohonen-Netze ist es möglich, Cluster graphisch darzustellen, was die Anschaulichkeit des Verfahrens erhöht.

Die Besonderheit der künstlichen neuronalen Netze besteht darin, dass es nicht notwendig ist, ein theoretisches Modell bzw. einen funktionalen Zusammenhang zwischen verschiedenen Variablen explizit zu unterstellen.<sup>734</sup> Das Verfahren ist in der Lage, nichtlineare Strukturen in den Daten zu entdecken und zu modellieren.<sup>735</sup>

Ein künstliches neuronales Netz bietet einen höheren Freiheitsgrad an Lernkapazität, da die interne Modellierung und Konzeption des Netzes nicht so stark eingeschränkt wird.<sup>736</sup> Das heißt, es ist nicht an die einschränkenden Randbedingungen einer Wissensdarstellung in Form von scharfen Regeln gebunden. In diesem Zusammenhang können auch manche künstliche neuronale Netze (wie z. B. Radiale-Basis-Funktion-Netze/RBF-Netze) mit Fuzzy- Logik in Verbindung gebracht werden.<sup>737</sup> Künstliche neuronale Netze eignen sich besonders für die Verarbeitung von verrauschten, lückenhaften oder sogar widersprüchlichen Daten. Das deutet auf eine hohe Generalisierungsfähigkeit der künstlichen neuronalen Netze hin<sup>738</sup>: Bei

---

<sup>734</sup> Künstliche Neuronale Netze stellen im Gegensatz zu Regressions-, Faktoren- und Diskriminanzanalysen, welche auf linearen Modellen basieren, eine nichtlineare Modellierungsform dar, die mehr Flexibilität bietet, vgl. Krahl/Windneuser/Zick (1998), S. 68f.

<sup>735</sup> Vgl. Nakhaeizadeh (1998), S. 13.

<sup>736</sup> Vgl. Krahl/Windneuser/Zick (1998), S. 68.

<sup>737</sup> Vgl. ebd., S. 68.

<sup>738</sup> Vgl. ebd., S. 69.

entsprechendem Training erlernt ein Netz die wesentlichen Strukturen in den Daten und filtert die zufälligen Phänomene (Rauschen) heraus.

Die hohe Flexibilität und Generalisierungsfähigkeit sind wesentliche Vorteile der künstlichen neuronalen Netze, aber um in den Genuss dieser Vorteile zu gelangen, sind sehr aufwendige und komplizierte Trainingsvorgänge notwendig. Je mehr Freiheit in einem System steckt bzw. je flexibler ein System ist, desto mehr müssen potentielle Fehlerquellen im System berücksichtigt werden. Zwischen der Freiheit und der potentiellen Fehler eines Systems besteht also ein Trade-off. Da die Freiheit beim Modellbau stets Fehler hervorrufen, wäre es auf jeden Fall für einen guten Modellbau hilfreich, den Trainingsverlauf möglichst gut durch die Anzeige verschiedener Indikatoren zu überblicken.

- **Segmentierungs- bzw. Clusteringmethoden**

Clusteringmethoden werden verwendet, um Datensätze bezüglich eines Ähnlichkeitskonzeptes zu Gruppen (Cluster) so zusammenzufassen, dass die Datensätze innerhalb einer Gruppe möglichst ähnlich, hingegen die Datensätze aus unterschiedlichen Gruppen möglichst verschieden bzw. unähnlich sein sollen.<sup>739</sup> Durch das Clusterverfahren wird der Datenbestand segmentiert, indem Datengruppen gebildet werden. Diese Gruppen sind im Gegensatz zu den Klassen bei den Klassifikationsverfahren a priori nicht bekannt, d. h. die gebildeten Klassen sind das Ergebnis des Clusterverfahrens.<sup>740</sup>

Durch die Segmentierung des Datenbestandes in Gruppen soll Wissen über Ähnlichkeiten der Datenobjekte, die sich in verschiedenen Clustern befinden, entdeckt werden. Während dabei die Bildung von Clustern (z. B. durch gut entwickelte Methoden aus der multivariaten Statistik oder Fuzzy-Logik) relativ einfach ist, erweist sich die Interpretation der gefundenen Cluster oftmals als problematisch.<sup>741</sup>

Im CRM wird das Clusterverfahren häufig angewendet, um Kunden in Gruppen zu unterteilen und entsprechend dieser Gruppierung sie individuell anzusprechen.

Die Grundlage für die Anwendung des Clusterverfahrens ist die Definition von geeigneten Ähnlichkeitsmaßen sowohl zwischen je zwei Datensätzen als auch zwischen einzelnen Clustern.<sup>742</sup> Liegen Datensätze mit quantitativen (metrisch skalierten)

---

<sup>739</sup> Vgl. Beekmann/Chamoni (2006), S. 274.

<sup>740</sup> Vgl. ebd., S. 265.

<sup>741</sup> Vgl. ebd., S. 266.

<sup>742</sup> Vgl. ebd., S. 274.

Attributwerten vor, so wird bspw. die Euklidische Distanz (als Wurzel aus der Summe der Quadrate der jeweiligen Attributsdifferenzen zwischen zwei Datenobjekten) oder die City-Block-Metrik als Abstandsmaß gewählt. Im Falle von qualitativen (nominal skalierten) Attributwerten kann als Ähnlichkeitsmaß die Anzahl der übereinstimmenden Attributwerte zweier Datensätze benutzt werden. Ein derartiges Ähnlichkeitsmaß wird bspw. von dem im IBM-Intelligent-Miner implementierten Demographischen Clustering-Verfahren verwendet. Diesem Verfahren liegt folgende Idee zugrunde: Paare von Datensätze werden über die Werte der einzelnen Attribute miteinander verglichen. Die Anzahl der ähnlichen Attribute bemisst den Ähnlichkeitsgrad, wogegen die Anzahl der unähnlichen Attribute den Unähnlichkeitsgrad angibt.

Die Clusterverfahren lassen sich nach *hierarchischen* und *partitionierenden* Verfahren unterscheiden.<sup>743</sup> Bei dem hierarchischen Clustering, wie man schon am Namen erkennt, wird eine ganze Hierarchie möglicher Segmentierungen ermittelt. Es gibt zwei mögliche Vorgehensweisen:<sup>744</sup>

Ausgehend von einer Aufteilung, bei der jedes Datenobjekt als eigener Cluster verstanden wird, ähnliche Cluster schrittweise zu größeren Clustern zusammengefasst werden, bis eine geeignete Aufteilung des Gesamtdatenbestandes erreicht ist. Dieses Vorgehen nennt man agglomeratives hierarchisches Clustering. Der bekannteste Vertreter dieses Verfahrens ist das Nearest Neighbour-Verfahren. Bei dem divisiven hierarchischen Clustering gehört dagegen jedes Datenobjekt zunächst zu einem Cluster. Sukzessiv werden danach neue (Teil-)Cluster erzeugt, die zu einer sinnvolleren Clusterbildung führen, bis nur noch Cluster, bestehend aus Einzelobjekten übrig bleiben. Bei diesem Vorgehen wird also der Datenbestand in immer feinere Cluster aufgeteilt.

Das hierarchische Clustering ist handhabbar und äußert flexibel, da es durch die verschiedenen Hierarchiestufen mehrere mögliche Lösungen präsentieren kann. Die Hierarchiestufen unterscheiden sich in der Anzahl der Cluster. Es ist damit möglich, nach der Durchführung eines solchen Verfahrens eine beliebige Hierarchiestufe auszuwählen und zu verwenden. Eine optimale Lösung für die Clusterbildung kann daher leicht herausgefunden werden. Außerdem muss die Anzahl der Cluster vor der Anwendung des Verfahrens nicht festgelegt werden. Das ist ein wesentlicher Vorteil gegenüber dem partitionierenden Clusterverfahren.<sup>745</sup> Nachteilig ist, dass wegen der

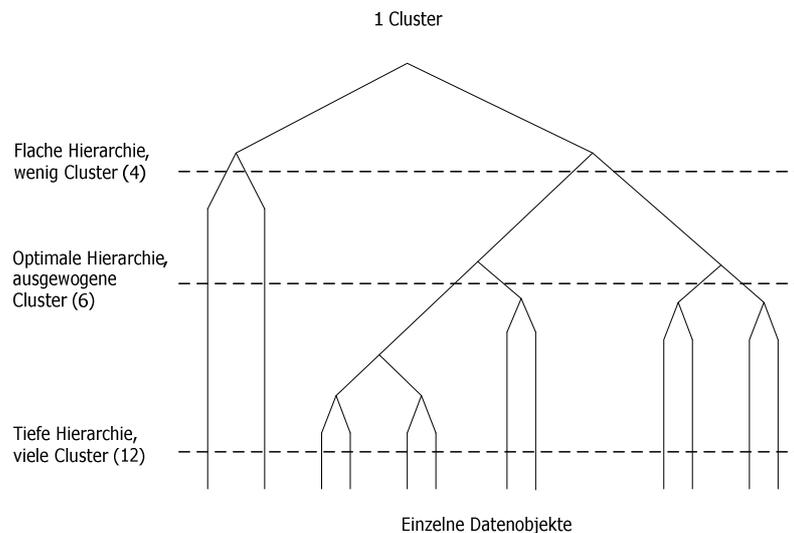
---

<sup>743</sup> Vgl. Beekmann/Chamoni (2006), S. 274.

<sup>744</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 88, Gluchowski/Gabriel/Dittmar (2008), S. 196.

<sup>745</sup> Vgl. Beekmann/Chamoni (2006), S. 275.

fehlenden Iteration nicht möglich ist, anfangs getätigte falsche Clusterbildungen rückgängig zu machen.



**Abbildung 42: Hierarchisches Clustering**

Quelle: Krahl/Windheuser/Zick (1998), S. 88.

Die hierarchischen Clusterverfahren sind für die Anwendung auf große Datenbestände schlecht geeignet. Daher werden in den meisten Data Mining-Werkzeugen überwiegend partitionierende Clusterverfahren verwendet.<sup>746</sup>

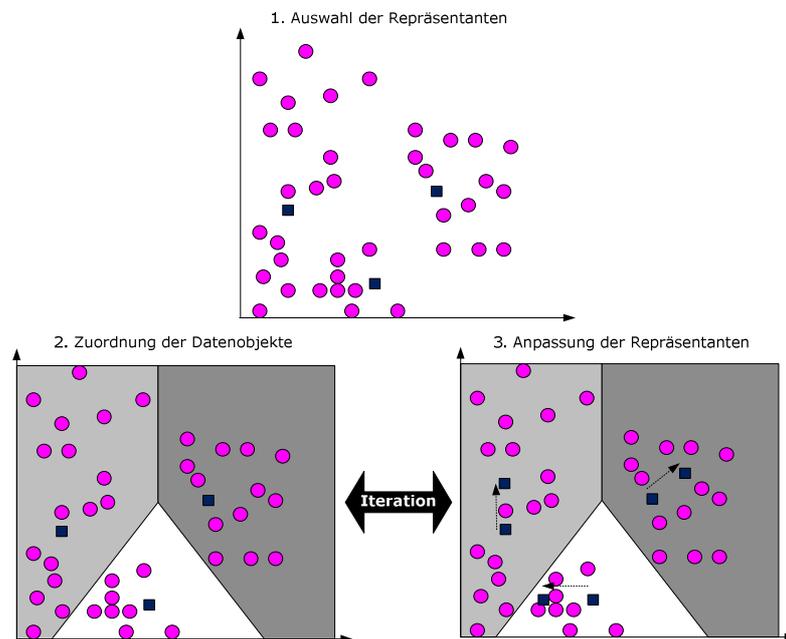
Bei den partitionierenden Clusterverfahren wird die Clusteranzahl fest vorgegeben, so dass jeder Datensatz sukzessiv in die Cluster eingeteilt wird, deren Ähnlichkeit zum Datensatz am größten ist.<sup>747</sup> Der bekannteste Vertreter des partitionierenden Verfahrens ist das K-Means-Verfahren. K-Means-Clustering ist ein Segmentierungsverfahren, das im Wesentlichen in zwei Schritten durchgeführt wird:

Zunächst wird aus dem Datenbestand eine bestimmte Anzahl (k-Stück) von Datenobjekten zufällig ausgewählt, die in der Folge als Clusterrepräsentanten und als Zentrum jeweils eines Clusters anzusehen sind. Im zweiten Schritt wird jedes weitere (k+1)-te Datenobjekt demjenigen Cluster zugeordnet, dessen Repräsentanten sowie dessen Zentrum die größte Ähnlichkeit aufweist. Die Zuordnung aller Datenobjekte zu den Clustern erfolgt durch ein geeignet gewähltes Ähnlichkeitsmaß. Die Repräsentanten werden dann nach der Zuordnung der neuen Datenobjekte angepasst. Das heißt, die Clusterzentren müssen neu berechnet werden, und zwar als Mittelwert aller Datenobjekte, die sich im selben Cluster befinden. Dieses Verfahren, das Clusterzentren jedes Mal neu berechnet und anschließend Datenobjekte erneut

<sup>746</sup> Vgl. Beekmann/Chamoni (2006), S. 275.

<sup>747</sup> Vgl. ebd., S. 275.

zuordnet, wird für jedes Datenobjekt iterativ wiederholt, bis schließlich alle Objekte auf die k-Cluster verteilt sind.



**Abbildung 43: 3-means Clustering**

Quelle: Krahl/Windheuser/Zick (1998), S. 87.

Problematisch bei dem K-Means-Verfahren ist, dass die Anzahl der zu bestimmenden Cluster fest vorzugeben ist.<sup>748</sup> Eine optimale Anzahl kann aber a priori nicht bekannt sein. Um diese zu bestimmen, müssen mehrere Durchläufe mit verschiedenen Werten für k durchgeführt und die Ergebnisse miteinander verglichen werden. Außerdem muss der Anwender selbst die Bedeutung der gebildeten Cluster bestimmen. Häufig treten jedoch Erklärungs- bzw. Interpretationsprobleme auf.

Des Weiteren gibt es zahlreiche Erweiterungen des K-Means-Verfahrens durch die Anwendung von Fuzzy-Logik. Hierbei werden keine scharfen, sondern prozentuale Clusterzugehörigkeiten betrachtet. Das heißt, ein Datenobjekt wird gleichzeitig mehreren Clustern nach einem bestimmten Prozentsatz zugeordnet, wobei der Prozentwert dann als Zugehörigkeitsgrad interpretiert werden kann. Dieses Verfahren wird als Fuzzy-K-Means (oder als Fuzzy-C-Means)-Verfahren bezeichnet.

- **Entdeckung von Abhängigkeiten: Assoziationsanalysen**

Bei der Entdeckung von Abhängigkeiten besteht die Aufgabe darin, Abhängigkeiten zwischen Merkmalen oder einzelnen Ausprägungen dieser Merkmale eines Datenbestandes bzw. einer Teilmenge dieses Bestandes zu erfassen. Dabei werden a

<sup>748</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 87.

priori keine Abhängigkeiten angenommen, wie etwa bei der Regression, sondern sie sollen selbstständig entdeckt werden.<sup>749</sup>

Die Abhängigkeiten werden zunächst gemäß einem Kriterium ohne eine kausale Erklärung angezeigt. Danach werden sie einzeln geprüft, um sicherstellen zu können, dass solche Abhängigkeiten tatsächlich vorliegen. Es existieren zahlreiche unterschiedliche Methoden, die zur Entdeckung von Abhängigkeiten eingesetzt werden können.<sup>750</sup> Eine Gruppe von Methoden, welche als Output Assoziationsregeln erzeugen, ist aber besonders wichtig: Durch die sogenannten Assoziationsanalysen werden Regeln generiert, welche die Beziehungen zwischen den in den Datensätzen eines Datenbestandes vorkommenden Elementen (Items) beschreiben. Grundsätzlich kann man zwischen drei Beziehungstypen unterscheiden:<sup>751</sup>

- *Verknüpfungen*: Das Aufdecken von Verknüpfungen zielt darauf ab, gemeinsam auftretende Ereignisse oder Vorfälle (bspw. wiederholte Käufe gleicher Artikelkombinationen) zu identifizieren (Durch eine Kassenbondatenuntersuchung werden die zusammen gekauften Artikel im Supermarkt identifiziert. Diese aufgedeckten Verknüpfungen zwischen den Artikeln können dann dazu verwendet werden, bestimmte Artikel zu einem erfolgversprechenden Einzelprodukt zu bündeln).
- *Abhängigkeiten*: Über den Aussagegehalt von Verknüpfungen hinaus beschreibt die Abhängigkeitsbeziehung zudem die Stärke und Richtung der Verbindung und legt diese in Form von Abhängigkeitsregeln (Wenn-Dann-Beziehungen) mit den zugehörigen bedingten Wahrscheinlichkeiten fest.
- *Sequenzen*: Eine Sequenz deutet darauf hin, dass einzelne Ereignisse häufig mit einem gewissen zeitlichen Verzug aufeinander folgen. Im obigen Beispiel der Kassenbon-Datenuntersuchung würde das bedeuten, dass kein Kassenbon unabhängig von den anderen im Mittelpunkt der Betrachtung steht, sondern eine ganze Sequenz von Kassenbons, die zu ein und demselben Kunden gehören. Damit wird das Kaufverhalten eines Kunden über einen Zeitraum hinweg analysiert, wodurch z. B. Folgekäufe wie erst Computer, dann Software und dann noch weitere Hardware aufgedeckt werden können.<sup>752</sup>

---

<sup>749</sup> Vgl. Beekmann/Chamoni (2006), S. 266.

<sup>750</sup> Ausführlich zu diesen Methoden, siehe z. B. Nakhaeizadeh (1998), S. 16f.

<sup>751</sup> Vgl. Gluchowski/Gabriel/Dittmar (2008), S. 202.

<sup>752</sup> Das Ziel von Sequenzanalysen ist, zusammengehörige, aber zeitlich versetzte Transaktionen zu untersuchen und interessante Kaufmuster aufzudecken. Die Ergebnisse können dabei helfen, geeignete Zeitpunkte für Werbeaktivitäten für Zusatzprodukte oder -leistungen bei bestehenden Kunden zu ermitteln, vgl. Gluchowski/Gabriel/Dittmar (2008), S. 204.

Eine Assoziationsregel könnte wie folgt lauten: *Wenn Item A vorkommt, (dann) tritt auch Item B auf.*

Dabei repräsentieren die Items die Ausprägungen von Attributwerten eines Datensatzes, und die Regel wird durch diesen Datensatz unterstützt, falls der Datensatz die Items der Regel beinhaltet. Der Bedingungsteil (Wenn-Teil) der Regel wird auch als Prämisse oder Antezedens bezeichnet, der Dann-Teil dagegen als Konklusion bzw. Implikation oder Sukzedens. Die algorithmische Umsetzung der Assoziationsanalyse zur Aufdeckung von Assoziationen basiert auf der Häufigkeitsbetrachtung von Attributkombinationen. Dazu werden zwei grundlegende Maße verwendet, mit denen man die Assoziationsregeln bewerten kann:<sup>753</sup>

Der Träger (*Support*) einer Regel (bestehend aus einer Attributmenge) gibt an, wie prozentual häufig die dieser Regel zugrunde liegenden Attribute gemeinsam innerhalb des gesamten Datenbestandes vorkommen. Je größer der Supportwert ist, desto relevanter ist die Regel.

$$\text{Supp} (A > B) = \frac{\text{Anzahl der Transaktionen, die A und B enthalten}}{\text{Anzahl aller Transaktionen}}$$

Die Konfidenz (*Confidence*) verwendet man für die Berechnung der Stärke einer Assoziationsregel. Sie gibt prozentual an, wie oft bei Zutreffen auf A tatsächlich auch B betroffen ist.

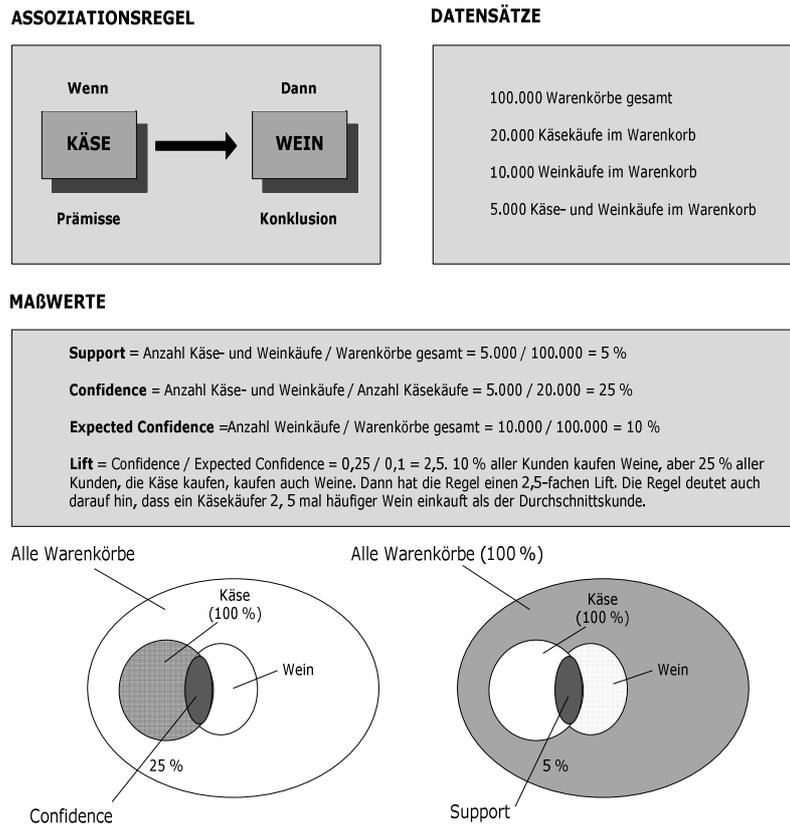
$$\text{Conf} (A > B) = \frac{\text{Anzahl der Transaktionen, die A und B enthalten}}{\text{Anzahl der Transaktionen, die A enthalten}}$$

Ein hoher Konfidenzwert deutet auf eine ausgeprägte Abhängigkeit zwischen den betrachteten Items hin.

Ferner kann man die Anzahl der Transaktionen, in denen der Konklusionsteil der Regel auftritt, mit der Gesamtzahl aller Transaktionen in Verbindung bringen. Somit ergibt sich als weitere Maßgröße die erwartete Konfidenz (*Expected Confidence*). Aus der Division von Confidence und Expected Confidence lässt sich anschließend der sogenannte Lift berechnen. Der Lift einer Regel gibt an, um wieviel höher sich die berechnete Konfidenz im Vergleich zur erwartenden Konfidenz darstellt und damit wieviel häufiger die Konklusion (die rechte Regelhälfte) bei erfüllter Prämisse (die linke Regelhälfte) eintritt als in der Grundgesamtheit.

---

<sup>753</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 81 und Gluchowski/Gabriel/Dittmar (2008), S. 202f.



**Abbildung 44: Zentrale Maßgrößen der Assoziationsanalyse**  
 Quelle: in Anlehnung an Gluchowski/Gabriel/Dittmar (2008), S. 203.

Eine mögliche Assoziationsregel, die durch die Maße Träger und Konfidenz bewertet wird, ist z. B.:

*In 75% der Fälle, in denen Item A in einem Datensatz enthalten ist, tritt auch Item B auf. Insgesamt erscheint diese Regel in 10% aller Datensätze im Gesamtbestand. Anders formuliert: In 75% der Fälle, in denen ein Computer gekauft wurde, wird ein Drucker gekauft. Diese Situation lässt sich bei 10% aller Transaktionen beobachten.*

In einer umfangreichen Datenbasis treten beliebig viele Assoziationen wie diese auf. Daher stellt sich nun die Frage, wie man entscheiden kann, ob alle diese Assoziationsregeln für die Analyse tatsächlich sinnvoll sind.

Ausgehend von dieser Fragestellung müssen zur Generierung von Assoziationsregeln zunächst Mindestwerte der gesuchten Regeln definiert werden. Durch die Vorgabe von minimalen Träger- und Konfidenzwerten kann man leicht herausfinden, ab wann eine Assoziation für die Analyse als besonders interessant und nützlich angesehen werden kann. Anschließend werden nur solche Regeln generiert, deren Träger- bzw. Konfidenzwerte die vorgegebenen Mindestwerte überschreiten. Die folgende Abbildung 45 zeigt ein Beispiel für die Regelgenerierung.

Im ersten Schritt werden die Attributkombinationen

- {Deodorant, Zahnbürste, Zahncreme, Zahnseide},
- {Zahnbürste, Zahncreme, Zahnseide} und
- {Zahncreme, Zahnseide}

mit Trägerwerten 5%, 10% bzw. 25% gefunden. Als minimaler Trägerwert ist 5% vorgegeben. Um die Gültigkeit der Regeln

- {Deodorant, Zahnbürste} > {Zahncreme, Zahnseide}
- {Zahnbürste} > {Zahncreme, Zahnseide}

zu überprüfen, muss noch die Konfidenz berechnet werden:

- Conf ({Deodorant, Zahnbürste} > {Zahncreme, Zahnseide}) =  $5/25 = 20\%$
- Conf ({Zahnbürste} > {Zahncreme, Zahnseide}) =  $10/25 = 40\%$

Bei einer geforderten minimalen Konfidenz von 30% wird nur die zweite Regel akzeptiert.

**Abbildung 45: Generierung einer Regel in der Assoziationsanalyse**

Quelle: in Anlehnung an Krahl/Windheuser/Zick (1998), S. 82.

Vorteile der Assoziationsanalyse sind die leichte Bedienbarkeit und Anschaulichkeit. Der Aufwand der Modellbildung ist eher gering, und die Ergebnisse sind leicht verständlich. Mit der Assoziationsanalyse können sehr umfangreiche Transaktionsdatenbestände mühelos analysiert werden. Sein wesentlicher Nachteil liegt jedoch darin, dass der Benutzer aufgrund der hohen Anzahl an generierten Regeln leicht den Überblick verliert und daher die Ergebnisse nicht adäquat auswerten kann.

### 5.2.6 Statistik und andere Methoden, die im Data Mining ihre Anwendung finden

Einige Methoden und Verfahren, die im letzten Kapitel beschrieben wurden, wie z. B. K-Nächste-Nachbarn-Technik, Diskriminanzanalyse, K-Means-Clustering und hierarchisches Clustering basieren eher auf klassischen statistischen Untersuchungen der vorhandenen Datenbasis. Daher ist es nicht sinnlos, diese Methoden und Verfahren eher dem Gebiet der Statistik zuzuordnen. Ein isolierter Einsatz nur moderner Methoden wird für komplexe Fragestellungen in Knowledge Discovery in Databases bzw. Data Mining nicht ausreichend sein.<sup>754</sup> Denn zum einen ist eine umfangreiche Vorverarbeitung der Daten (Preprocessing)<sup>755</sup> notwendig, um überhaupt mit einem Data Mining-Verfahren ansetzen zu können. Dieses Preprocessing wird von empirischen Untersuchungen unterstützt. Und zum anderen werden diese Daten nach der Vorverarbeitung im Zuge eines Data Mining-Projektes in einer neuen Datenbasis

<sup>754</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 89.

<sup>755</sup> Preprocessing umfasst Data Cleansing, Selection, Transformation, Normalization usw.

zusammengeführt (Aggregation), deren statistische Auswertung interessante Informationen zum Vorschein bringt. In diesem Zusammenhang kann Statistik nicht vollständig von KDD bzw. Data Mining getrennt werden. Sie ist ein wesentlicher Bestandteil davon, welcher einerseits die klassischen Analysemethoden dem Data Mining zur Verfügung stellt, andererseits die Voraussetzungen für die Anwendbarkeit der modernen Verfahren im Data Mining schafft.<sup>756</sup>

Jedoch gibt es ein Paar wichtige Punkte, die für die Abgrenzung von Data Mining zur statistischen Datenanalyse herangezogen werden können:<sup>757</sup>

In der Statistik werden Hypothesen formuliert und diese dann auf ihre Richtigkeit hin geprüft. Dazu müssen die Daten i. d. R. viele mathematische Voraussetzungen erfüllen, was aber in der Realität selten vorkommt. Häufig werden zur Analyse nur Stichproben aus den Datenbeständen gezogen, wodurch die allgemeine Gültigkeit der Analyseergebnisse verringert wird. Beim Data Mining hingegen ist die Fragestellung a priori nicht genau bekannt, und erst nach der Analyse ergeben sich konkrete Fragestellungen. Außerdem sind die zu analysierenden Datenmengen umfangreicher und i. d. R. ohne Voraussetzungen wählbar.

Ein wichtiger Bestandteil der Statistik sind sogenannte *lineare Methoden*. Diese verwendet man, falls zwischen den betrachteten Größen lineare Zusammenhänge bestehen. Ein linearer Zusammenhang zweier Variablen kann folgendermaßen definiert werden:  $X = a \times Y + b$ . Die Variable  $X$  wird hier durch die Variable  $Y$  erklärt, wobei  $a$  und  $b$  die Parameter des Modells repräsentieren. Stellt man die Abhängigkeit graphisch dar, erhält man eine Gerade (das heißt, eine lineare Abhängigkeit der Variablen  $X$  und  $Y$ ). Ein linearer Zusammenhang mehrerer Variablen hat z. B. die Form:  $V = a \times W + b \times X + c \times Y + d \times Z + e$ . In diesem Fall wird die Variable  $V$  durch die Variablen  $W$ ,  $X$ ,  $Y$  und  $Z$  erklärt. Wichtige Verfahren, die auf solchen linearen Methoden basieren, sind die Regressionsanalyse, die Faktorenanalyse und die Diskriminanzanalyse.

Die linearen Methoden erweisen sich aber bei realen Anwendungen als problematisch:<sup>758</sup> Große Veränderungen der Modellparameter bzw. der Variablen können eine übermäßige Reaktion des Modells hervorrufen, während kleine Veränderungen das Modell sehr wenig sogar kaum beeinflussen. Zur realistischen Anwendung dieses Verfahrens müssen daher gewisse Modellvoraussetzungen

---

<sup>756</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 89.

<sup>757</sup> Vgl. ebd., S. 89.

<sup>758</sup> Vgl. ebd., S. 90.

getroffen werden, welche aber in realen Problemen i. d. R. nicht vorliegen. Diese Diskrepanz zwischen Realität und deren Abbild ist bei den Verfahren der Statistik größer als bei den modernen Verfahren des Data Mining.

Ein weiteres Themengebiet, welches keine eindeutige Abgrenzung vom Data Mining erlaubt, ist die *Fuzzy-Logik*. Sie wird oft in einigen Expertensystemen und anderen Anwendungen der Künstlichen Intelligenz (KI) eingesetzt.

Fuzzy-Logik (englisch: fuzzy logic) ist eine Theorie, welche vor allem für die Darstellung des menschlichen (und damit unscharfen) Wissens entwickelt wurde. Sie ist eine Verallgemeinerung der zweiwertigen Booleschen Logik um einen weiteren unscharfen (englisch: fuzzy) Wahrheitswert, der zwischen wahr (1) oder falsch (0) liegt. Beispielsweise kann der Wahrheitswert den Wert 0,6 annehmen, so dass damit auch unscharfe Angaben wie „ein bisschen“, „ziemlich“ oder „stark“ mathematisch behandelt werden können. Daher arbeiten Fuzzy-Logik-unterstützte Programme näher am menschlichen Denken als übliche Programme.

Die Grundlage der Fuzzy-Logik sind die sogenannten unscharfen Mengen. Im Gegensatz zu traditionellen Mengen (im Kontext der Fuzzy-Logik auch scharfe Mengen genannt), in denen ein Element in einer vorgegebenen Grundmenge entweder enthalten oder nicht enthalten sein kann, kann ein Element in einer unscharfen Menge auch ein wenig enthalten sein. Der Grad an Zugehörigkeit wird meist durch eine Zugehörigkeitsfunktion (membership function)  $\mu$  beschrieben, die den Elementen einer Grundmenge eine reelle Zahl zwischen 0 und 1 zuordnet.

Der Fuzzy-Logik-Ansatz wird dem deduktiven Zweig des maschinellen Lernens zugeordnet.<sup>759</sup> Dieser Ansatz alleine ist daher kein Teil des Data Mining, aber findet im Zusammenspiel mit anderen Verfahren (Neuro-Fuzzy-Methoden, Fuzzy-Clusteranalyse) seine Anwendung im Data Mining. So erzeugt z. B. die Fuzzy-Clusteranalyse nicht wie die klassische Clusteranalyse scharfe Abgrenzungen von Klassen und deren Objekten, sondern erlaubt unscharfe Zuordnungen, wobei es in Grenzbereichen zu Überschneidungen von Klassen kommt. Objekte können gleichzeitig zwei nicht klar voneinander getrennten Klassen zugeordnet werden; sie gehören mit unterschiedlichen Zugehörigkeitsgraden zu verschiedenen Klassen.

Eine weitere weit verbreitete Analysemethoden sind *Genetische Algorithmen*. Genetische Algorithmen können dem Bereich der Evolutionären Algorithmen zugeordnet werden. Evolutionäre Algorithmen stellen eine besondere Klasse von

---

<sup>759</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 91.

stochastischen Optimierungsverfahren dar, deren Konzeption und Strategie von den Prinzipien biologischer Evolution inspiriert ist.<sup>760</sup> Durch Nachbildung evolutionstheoretischer Erkenntnisse wie Selektion, Kreuzung und Mutation versuchen Genetische Algorithmen, aus einer Menge zufällig bereitgestellter Lösungsmöglichkeiten eine (zumindest) nahezu optimale Lösung zu entwickeln.<sup>761</sup> Dabei werden zunächst eine Menge (Population) von Lösungskandidaten bzw. -Möglichkeiten (Individuen) zufällig erzeugt und diejenigen ausgewählt, die einem bestimmten Gütekriterium am besten entsprechen (Auslese). Deren Eigenschaften (Parameterwerte) werden dann miteinander kombiniert (Kreuzung-Rekombination) und leicht verändert (Mutation), um eine neue Population von Lösungskandidaten (eine neue Generation) zu erzeugen. Nach der Erzeugung jeder Generation wird die Güte der Lösung und deren Verbesserung zur vorherigen Generation verglichen. Dieser Prozess wiederholt sich iterativ solange, bis eines der Abbruchkriterien erreicht ist. Ein typischer Genetischer Algorithmus umfasst folgende Schritte:

1. Initialisierung: Erzeugen einer ausreichend großen Menge unterschiedlicher Lösungskandidaten („Individuen“). Diese bilden die erste Generation.
2. Evaluation: Für jeden Lösungskandidaten der aktuellen Generation wird anhand einer Zielfunktion (auch Fitness-Funktion genannt) ein Wert bestimmt, der seine Güte angibt.
3. Selektion: Zufällige Auswahl von Lösungskandidaten aus der aktuellen Generation. Dabei werden Lösungskandidaten mit besseren Zielfunktionswerten mit einer höheren Wahrscheinlichkeit ausgewählt.
4. Kreuzung bzw. Rekombination: Die Daten („Genom“) der ausgewählten Individuen werden gemischt und daraus neue Individuen erzeugt („Vermehrung“).
5. Mutation: Die Parameterwerte werden zufälligen Abänderungen unterworfen, um sicherzustellen, dass die Suche nach der optimalen Lösung nicht auf den anfänglichen Vorrat der Parameter beschränkt bleibt („Inzucht-Gefahr“).
6. Aus den neuen Individuen und eventuell, je nach Gestaltung des Verfahrens, auch aus den Individuen der aktuellen Generation werden die Mitglieder der neuen Generation ausgewählt, die dann zur aktuellen Generation wird.

---

<sup>760</sup> Vgl. Küsters (2001), S. 120.

<sup>761</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 93.

7. Wenn ein Abbruchkriterium erfüllt ist, wird der beste gefundene Lösungskandidat als Ergebnis ausgegeben und der Algorithmus beendet. Sonst wird er mit dem zweiten Schritt fortgesetzt.

Mit Hilfe genetischer Algorithmen gelangt man so von Generation zu Generation zu immer besseren Lösungen. Sie lassen sich überall dort einsetzen, wo für die Gesamtheit aller Lösungen ein Gütemaß existiert, mit welchem die Qualität einer jeden Lösung beurteilt werden kann. Genetische Algorithmen werden eher nur zur Optimierung innerhalb anderer Verfahren angewendet.<sup>762</sup> So werden sie z. B. für Klassifikationsaufgaben verwendet, um den Klassifikationsfehler zu optimieren.

### **5.3 PROZESSMODELL FÜR DIE IMPLEMENTIERUNG EINER ERFOLGREICHEN DATA MINING-LÖSUNG IM CRM**

Data Mining stellt alleine kein autonomes System dar, das selbstständig in einem beliebigen Datenbestand Auffälligkeiten und Zusammenhänge findet, sondern dabei handelt es sich vielmehr um einen interaktiven Prozess, in den oft mehrere Abteilungen eines Unternehmens einzubinden sind.<sup>763</sup> Die Integration von Data Mining im Unternehmen ist eine Herausforderung, da Unternehmen ganz unterschiedliche Charakteristiken bezüglich ihrer Geschäftstätigkeit, Aufbauorganisation und internen Prozessen aufweisen. Diese Vielfalt an Unterschiedlichkeiten zwischen Unternehmen führt auch zu verschiedenen CRM-Strategien, die jeweils nach firmenspezifischen Anforderungen unterschiedlich konzipiert und erarbeitet sind. Vor diesem Hintergrund müssen bei der Implementierung einer Data Mining-Lösung für CRM zum einen alle firmenspezifischen Anforderungen berücksichtigt, zum anderen eine prozessartige Vorgehensweise verfolgt werden. Diese Vorgehensweise kann anhand eines Referenzmodells CRISP-DM erklärt werden (siehe Abb. 46).

CRISP-DM (Cross Industry Standard Process for Data Mining) ist eine von der europäischen Kommission und den Unternehmen SPSS, NCR, DaimlerChrysler und OHRA geförderte Initiative zur Bildung eines standardisierten Prozessmodells, das weitgehend technologie- und branchenunabhängig ist.<sup>764</sup> Das Prozessmodell dient als Referenz für die Beschreibung eines typischen Data Mining-Kreislaufes und umfasst alle relevanten Phasen mit ihren Aufgaben und den Beziehungen zwischen ihnen. Der Data Mining-Kreislauf im Prozessmodell besteht aus sechs Phasen, wobei nicht alle

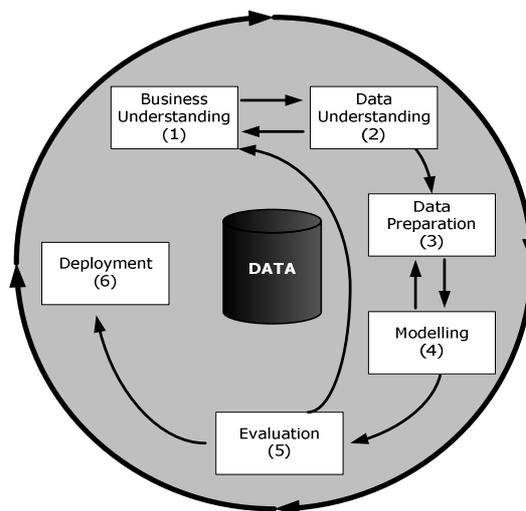
---

<sup>762</sup> Vgl. Krahl/Windheuser/Zick (1998), S. 93.

<sup>763</sup> Vgl. Meyer/Weingärtner/Döring (2001), S. 122.

<sup>764</sup> Vgl. CRISP-DM 1.0 (2000), S. 3f.

möglichen Beziehungen zwischen den Phasen beschrieben werden können, weil das mit den Zielen, den Hintergrundinteressen und vor allem mit dem jeweiligen Datenmaterial zusammenhängt. Die Reihenfolge der Phasen ist nicht starr, sondern eher flexibel in Abhängigkeit von den Ergebnissen der Phasen, die als Nächstes im Prozessablauf ausgeführt sind;<sup>765</sup> d. h. falls das ermittelte Ergebnis einer Phase nicht zufriedenstellend ist, kann der Prozess in die vorigen Phasen rückgekoppelt werden.



**Abbildung 46: Phasen des Referenzmodells CRISP-DM**

Quelle: in Anlehnung an CRISP-DM 1.0 (2000), S. 13.

Der äußere Kreislauf symbolisiert die zyklische Natur von Data Mining. So kann der Data Mining-Prozess die Phasen erneut durchlaufen, nachdem eine Data Mining-Lösung erfolgreich umgesetzt worden ist. Die gefundene Lösung und die Erfahrungen bzw. die Erkenntnisse, die während des Prozessdurchlaufs gewonnen werden, können neue, meistens fokussiertere Geschäftsfragen auslösen. Im CRM-Kontext bedeutet dies, dass jede nachfolgende Data Mining-Analyse von vorangegangenen Analysen profitiert, sich ständig an neue Gegebenheiten anpasst und die Güte der Informationen über Kunden verbessert.

Die Phasen des Kreislaufs können im CRISP-DM Modell wie folgt zusammengefasst werden:<sup>766</sup>

1. **Business Understanding:** Den Ausgangspunkt eines Data Mining-Prozesses bildet ein spezifisches Geschäftsproblem. Die aus den CRM-Prozessen resultierende Fragestellung muss unter Berücksichtigung der vorhandenen Datenbasis in ein konkretes Data Mining-Ziel transformiert werden (Z. B.

<sup>765</sup> Vgl. CRISP-DM 1.0 (2000), S. 13.

<sup>766</sup> Vgl. CRISP-DM 1.0 (2000), S. 13, Holland (2004a), S. 134-137 und Meyer/Weingärtner/Döring (2001), S. 122-126.

Realisierung von Cross Selling-Potentialen über die Bestimmung von Kaufwahrscheinlichkeiten für ein bestimmtes Kundensegment). Die Ziele und Erfordernisse des Data Mining-Projektes werden aus dem geschäftlichen Blickwinkel der Anwender untersucht und festgelegt.

2. **Data Understanding:** In dieser Phase des Data Mining-Prozesses soll sich der Analyst mit den Daten, die dem konkreten Geschäftsproblem zugrunde liegen, vertraut machen, indem er zunächst die möglichen relevanten Rohdaten sammelt und beschreibt und daraufhin die zugehörigen Datenfelder hinsichtlich ihres Lösungsbeitrags zur Problemdefinition und der Datenqualität überprüft. Für die Überprüfung und Bestimmung der Rohdatenbasis, welche dann für die nachfolgende Datenaufbereitung zur Verfügung steht, können statistische Analysen verwendet werden. Zusätzlich sollen Hypothesen formuliert werden, um bereits versteckte Informationen in den Daten aufzudecken. Falls kein zentraler Datenbestand für Data Mining existiert, ist es durchaus möglich, die relevanten Daten aus verschiedenen Quellsystemen zu beschaffen oder komplett durch Marktforschung zu erheben.
3. **Data Preparation:** Die Datenaufbereitung stellt innerhalb des Data Mining-Prozesses die zeit- und arbeitsintensivste Phase dar. Nach der Selektion der relevanten Daten als Rohdatenbestand müssen nun diese Daten bereinigt und in eine für die Analyse geeignete Form und Struktur gebracht werden: Zuerst werden fehlende Werte und sogenannte Ausreißer aus dem Datenbestand herausgenommen. Daraufhin werden zusätzlich benötigte Attribute dem Datenbestand hinzugefügt. Im nächsten Schritt sollen die Daten durch Aggregation verdichtet werden, da viele Detailinformationen im Datenbestand nur kumuliert für die Weiterverarbeitung im Data Mining-Prozess benötigt werden. Folglich werden diese Daten in das benötigte Format des jeweiligen Analyseverfahrens überführt. Am Ende erhält man aus dem anfänglichen Rohdatenbestand (initial raw data) eine Gruppe von Daten (final data set), auf die spezielle Data Mining-Techniken und -Werkzeuge in der Modellierungsphase angewendet werden können.
4. **Modelling:** Innerhalb dieser Phase kommen die eigentlichen Data Mining-Algorithmen zum Einsatz. In Abhängigkeit vom ausgehenden Geschäftsproblem ist die geeignete Modellierungstechnik auszuwählen und anzuwenden. Dabei werden innerhalb der Modellierung die beeinflussenden Parameter so lange iterativ angepasst, bis das Modell die gewünschte Güte erreicht hat.

5. **Evaluation:** In der Evaluierungsphase wird die Güte der Data Mining-Modelle beurteilt. Unter Einbezug der Anwender des Modells werden die bisherigen Schritte, die für den Entwurf des Modells ausgeführt sind, und die aus dem Modell resultierenden Ergebnisse überprüft, um sicherzustellen, dass dieses Modell die geschäftlichen Ziele in richtiger und genauer Art und Weise erreicht. Eine wichtige Aufgabe dabei ist es, festzustellen, ob geschäftliche Sachverhalte vorliegen, die nicht ausreichend berücksichtigt worden sind. Am Ende dieser Phase wird eine Entscheidung über den Gebrauch der Data Mining-Ergebnisse erzielt.
  
6. **Deployment:** Auf Basis der Ergebnisse und Erkenntnisse der Evaluierungsphase wird in der Deployment-Phase die Umsetzung der Data Mining-Lösung in einzelnen Schritten geplant. Das Ergebnis der Deployment-Phase kann so einfach wie die Erstellung eines Schlussberichts oder so komplex wie die Implementierung eines sich wiederholenden unternehmensübergreifenden Data Mining-Prozesses sein. Im komplexeren zweiten Fall müssen die laufende Überwachung (monitoring) und Aufrechterhaltung (maintenance) der Data Mining-Ergebnisse in die Planung mit einbezogen werden. In vielen Fällen werden die Deployment-Schritte im Unternehmen eher von Anwendern ausgeführt und nicht von Analysten. Sie müssen durchaus verstanden haben, welche Aktionen auszuführen sind, um die ihnen von Analysten mitgeteilten Data Mining-Ergebnisse gebräuchlich machen zu können. Bei der Berichterstellung ist es daher zu beachten, dass die Ergebnisse in einer leicht verständlichen Form erfasst werden. Deployment kann im CRM-Kontext als die Integration des Data Mining-Prozesses in die CRM-Prozesse verstanden werden. Im Sinne eines effizienten CRM-Konzepts werden die im Data Mining-Prozess gewonnenen Informationen über Kunden an die CRM-Systeme verteilt, deren primäre Aufgabe darin besteht, Kundenkontakte zu managen. Dabei werden nicht nur Informationen über bestimmte Kaufwahrscheinlichkeiten der Kunden berücksichtigt, sondern auch die von Kunden bevorzugten Kommunikationskanäle für einen möglichen Dialog mit dem Unternehmen.

# KAPITEL 6

---

## FUZZY-LOGIK

### 6.1 EINFÜHRUNG

*„Genauigkeit ist nicht die Wahrheit“*

*[Henri Matisse, französischer Maler]*

*„Unschärfe Mengen...bringen Computer dazu,  
ähnlich wie Menschen zu denken“*

*[Lotfi A. Zadeh, Gründer der Theorie von Fuzzy-Mengen]*

Ein grundlegendes Problem von Verfahren und Methoden, die für Knowledge Discovery in Databases und Data Mining eingesetzt werden, besteht darin, dass die Verwender mit großen Mengen von unterschiedlichen Daten bzw. Informationen umgehen und trotz deren Komplexität und Unsicherheit immer noch gute Ergebnisse erzielen sollen. Die Daten werden für die Analyse meistens aus unterschiedlichen Quellen bzw. Datenbanken und Informationssystemen erhoben. Dieser erhobene Datenbestand verbirgt aber unvollständige und sogar fehlerhafte Daten in sich, welche die Durchführung einer effektiven und vor allem korrekten Analyse erschweren. Widersprüchliche, fehlende Daten oder sogenannte Ausreißer müssen daher vor der Durchführung einer Analyse entsprechend behandelt werden. So werden im Prozess des Data Cleansing widersprüchliche Daten und Ausreißer aus dem Datenbestand entfernt, bzw. fehlende Daten werden interpoliert. Das Löschen oder Entfernen von Informationen aus dem Datenbestand können aber kritische Folgen in CRM haben. Dadurch wird z. B. die Entdeckung neu entwickelnder Trends oder die Identifizierung der Kunden mit besonders interessanten Profilen deutlich beeinträchtigt.

Die notwendigen Entscheidungen in CRM stützen sich oft auf unsichere und unvollständige Informationen. Demzufolge erzeugen auch Analysen, die mit solchen lückenhaften Daten durchgeführt sind, einen hohen Grad an Unsicherheit in den

Ergebnissen. Die Kombination der modernen Datenhaltungskonzepten und -techniken mit dem CRM ermöglicht Unternehmen ihre Geschäftsprozesse möglichst gut zu optimieren und somit den zu verarbeitenden Datenbestand klar und konsistent zu halten, dennoch entstehen aufgrund der enormen Informationsflut immer wieder inkonsistente und unvollständige Daten. Um mit solchen vagen Daten umgehen zu können und die daraus resultierende Unsicherheit und Komplexität der Analysevorgänge handhabbar zu machen, sind revolutionäre Methoden und Ansätze nötig. Die fuzzy-theoretische Erweiterung der klassischen Logik und Methoden könnte hier eine wirksame Abhilfe für die robuste und zuverlässige Datenanalyse schaffen.

## **6.2 WARUM FUZZY-LOGIK?**

Jeder kennt die klassische Logik eines Computers. Entweder ist ein Zustand wahr oder falsch, eins oder null. Halbwahrheiten gibt es bei dieser Logik nicht, und die Welt des Computers hat somit eine klare Trennung zwischen schwarz und weiß, gut und böse und so weiter.

Anders als Computer denken die Menschen aber viel differenzierter. Neben schwarz und weiß können wir auch jegliche Zwischenstufen bzw. Graustufen wahrnehmen. Das Problem ist jedoch, egal ob man in Graustufen oder wie bei Computern gar in exakten Farben denkt, dass man nie ganz sicher sein kann, ob etwas wirklich so ist, wie man es haben möchte. Die Tatsache ist immer mit einer gewissen Unsicherheit bzw. Vagheit behaftet. Davon ausgehend, trennt sich die Fuzzy-Logik von der Vorstellung, dass alles exakt bzw. scharf, gut oder schlecht sein muss. Es dürfen auch die Werte dazwischen liegen – zu 60% schwarz und zu 40% weiss.

Der Begriff „Fuzzy“ kommt aus dem englischen Sprachraum und bedeutet in etwa „unscharf“, „unsicher“, „fusselig“, „struppig“, „verschwommen“. So könnte man vielleicht sagen, dass der Begriff „Fuzzy-Logik“ eine unscharfe Logik bezeichnen soll. Aber eine solche Bezeichnung ist irreführend, da sich eine erfolgreich eingesetzte Logik durch Schärfe und Präzision auszeichnet und sich daher die Begriffe „unscharf“ und „Logik“ auszuschließen scheinen. Allerdings deutet der Begriff „Fuzzy“ in Verbindung mit dem Begriff „Logik“ auf eine gezielt angewandte Unschärfe hin, die sich bei realen Systemen als besonders vorteilhaft herausstellen kann. Denn diese Systeme weisen oft eine hohe Komplexität auf, die präzise Modellierungen und sichere Prognosen von vornherein aussichtslos erscheinen lässt.<sup>767</sup> In dieser Hinsicht stellt die Fuzzy-Logik keine schlampige und willkürliche Logik dar, sondern sie ist „[...] eine präzise Theorie des Unpräzisen. Sie nimmt in Kauf, dass wir die Gesetze, die komplexe

---

<sup>767</sup> Vgl. Biewer (1997), S. 28.

Systeme regieren, nicht mit letzter Präzision kennen, und versucht das Beste daraus zu machen. Und kommt dabei oft zu präziseren Ergebnissen als klassische Methoden, die ein Problem scheinbar exakt erfassen, aber an den auch für sie notwendigen Vereinfachungen scheitern.“<sup>768</sup>

## 6.3 GRUNDLAGENDE BEGRIFFE

Die Kenntnis der klassischen Mengentheorie ist der Schlüssel zum Verständnis der Fuzzy-Theorie. Außerdem bedarf die erfolgreiche Anwendung der Fuzzy-Logik einiger Festlegungen und Definitionen. In diesem Kapitel werden die wichtigsten davon kurz erläutert, um eine Grundlage für die weiteren Kapitel zu schaffen.

### 6.3.1 Klassische Mengenlehre und Logik

Unter klassischer Logik wird in der Regel die zweiwertige (duale/boolesche) Logik mit ihren Eigenschaften und Regeln zum logischen Schlussfolgern verstanden. Den Kern der klassischen Logik bildet die Verknüpfung logischer Aussagen. Da die abendländische mathematisch-philosophische Tradition von Schwarz-Weiß-Denken geprägt ist, kennt die klassische Logik nur wahre und falsche Aussagen. Sie beruht auf Axiomen, also unmittelbar einleuchtenden Grundsätzen. Beispiele für Axiome sind:

- *Satz vom ausgeschlossenen Widerspruch:*  $\neg (p \wedge \neg p)$ . Die Aussage  $p$  kann nicht gleichzeitig mit ihrer Satzverneinung also dem Gegenteil zutreffen, d. h. es gibt keine Aussage, die sowohl wahr als auch falsch ist („Es ist nicht der Fall, dass es (zugleich und am selben Ort) regnet und nicht regnet“).
- *Satz vom ausgeschlossenen Dritten:*  $p \vee \neg p$ . Jede Aussage ist wahr oder falsch („Die Erde ist rund, oder es ist nicht der Fall, dass die Erde rund ist“).

Die klassische Logik, abstammend von dem griechischen Wort „logos“ (was soviel heißt wie Wort, Gedanke, Sinn, Denken oder auch Vernunft) wurde als Lehre des vernünftigen Schließens von Aristoteles begründet und entwickelte sich über mehrere Stufen bis zu der heute üblichen mathematischen Logik.<sup>769</sup> Aus mengentheoretischer Sicht ist die klassische Logik dadurch gekennzeichnet, dass die Zugehörigkeit jedes Elementes zu einer Menge eindeutig durch „0“ oder „1“ bzw. durch „ja“ oder „nein“ definiert werden kann. Mit Hilfe klassischer mathematischer Definitionen werden Dinge, die bestimmte Eigenschaften haben, in Mengen eingeteilt. So kann eine klassische (scharfe) Menge als Zusammenfassung von Elementen angesehen werden,

---

<sup>768</sup> Drösser (1994), S. 9f.

<sup>769</sup> Vgl. Wikipedia, Begriff: Logik, [Abruf: 10-09-2008].

die mindestens eine gemeinsame Eigenschaft besitzen.<sup>770</sup> Wenn diese Eigenschaft nicht vorhanden ist, gehört das Element nicht zur Menge.

Georg Cantor begründete die Mengenlehre<sup>771</sup> in den Jahren 1874–1897. Statt des Begriffs *Menge* benutzte er anfangs Wörter wie „Inbegriff“ oder „Mannigfaltigkeit“; von Mengen und *Mengenlehre* sprach er erst später. 1895 formulierte er folgende Mengendefinition:

„Unter einer ‚Menge‘ verstehen wir jede Zusammenfassung  $M$  von bestimmten wohlunterschiedenen Objekten  $m$  unserer Anschauung oder unseres Denkens (welche die ‚Elemente‘ von  $M$  genannt werden) zu einem Ganzen.“<sup>772</sup>

Das Attribut „bestimmt“ drückt hier aus, dass es eindeutig entscheidbar ist, ob ein Objekt zu einer Menge gehört oder nicht. „Wohlunterschieden“ bedeutet, dass es eindeutig entscheidbar ist, ob zwei beliebige Elemente der Menge gleich oder ungleich sind. Daraus folgt, dass jedes Element nur einmal in der Menge enthalten ist.

Gestützt auf diese Definition kann die Ansammlung von zusammengehörenden Objekten aus einer Grundmenge  $\Omega$  als Menge  $M$  bezeichnet werden. Die einzelnen Objekte stellen die Elemente  $x$  der Menge  $M$  dar. Eine Menge wird durch seine Eigenschaften charakterisiert. Dazu wird meistens eine Grundmenge, eine Variable für die Elemente und eine Aussageform (Bedingungen) angegeben.<sup>773</sup>

Die so charakterisierte Menge  $M$  beinhaltet dann genau die Elemente der Grundmenge, für die die Aussageform nach Ersetzung der Variablen eine wahre Aussage ergibt, also die in der Aussageform beschriebenen Bedingungen ( $B$ ) erfüllt sind:

$$M = \{ x \mid x \in \Omega, B(x) \}.$$

Gegeben sei eine Menge  $A$ , bestehend aus den diskreten Werten  $x_1, x_2, x_3$  und  $x_4$ . Mit Hilfe einer zweiwertigen Funktion  $f_A$  (auch *Indikatorfunktion* oder *charakteristische Funktion* genannt) kann auch die Zugehörigkeit eines Elementes  $x$  zu dieser Menge durch einen der beiden zulässigen Werte „0“ und „1“ einfach festgestellt werden.<sup>774</sup>

---

<sup>770</sup> Vgl. Zimmermann (1993), S. 8.

<sup>771</sup> Vgl. Wikipedia, Begriff: Mengenlehre, [Abruf: 10-09-2008].

<sup>772</sup> Cantor (1966), S. 282, zit in: Biewer (1997), S. 48.

<sup>773</sup> Vgl. Biewer (1997), S. 49.

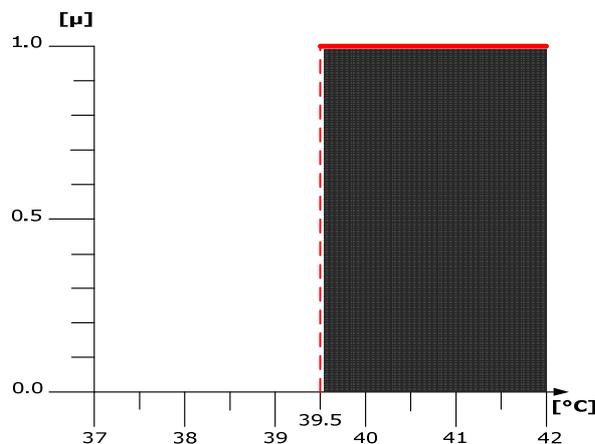
<sup>774</sup> Vgl. Biewer (1997), S. 50. und Schulte (1993), S. 18.

Die charakteristische Funktion der Menge  $A$ ,  $f_A(x): \Omega \rightarrow \{0, 1\}$ , liefert für jedes Element der Grundmenge  $\Omega$  den Wert eins, falls es zur Menge  $A$  gehört, und den Wert null, falls es nicht dazugehört:

$$f_A(x) = \begin{cases} 1, & \text{falls } x \in A \\ 0, & \text{sonst} \end{cases}$$

Da es eindeutig entscheidbar ist, ob ein Objekt zu einer Menge gehört oder nicht, handelt es sich hier um eine scharfe Menge (Crisp-Set) bzw. Mengeneinteilung.

Zur Veranschaulichung des Prinzips einer scharfen Mengeneinteilung kann das darunterliegende Beispiel dienen. Dieses Beispiel zeigt, wie man den Ausdruck „starkes Fieber“ als eine Zahlenmenge im klassischen Sinne definiert. Dazu stellt man zunächst einen Grenzwert fest, nach dem eine scharfe Einteilung der Körpertemperaturen in zwei unterschiedliche Gruppen „starkes Fieber“ und „nicht starkes Fieber“ erfolgen soll. Wählt man z. B. die Temperatur 39,5 C, so gelangt man zu folgender Funktionsdefinition:



**Abbildung 47: „starkes Fieber“ als klassische Menge**

Quelle: in Anlehnung an Biewer (1997), S. 56.

$$\Omega = \{x \mid x \in \mathfrak{R}, 32 \leq x \leq 44\}$$

$$\forall x \in \Omega: \mu_{SF}(x) = \begin{cases} 1, & \text{falls } x \geq 39,5 \\ 0, & \text{sonst} \end{cases}$$

Man erhält so eine Stufenfunktion mit einer Sprungstelle an der Abszisse  $x = 39,5$  °C. Diese Funktion besagt, dass die Werte der Körpertemperaturen von 37 °C bis 39,5 °C

einen Zugehörigkeitswert von 0.0 besitzen, mit der Bedeutung, dass die Temperaturen 37 °C bis 39,5 °C zur Menge „nicht starkes Fieber“ gehören. Hingegen sind ab der 39,5 °C Grenze alle Temperaturwerte als „starkes Fieber“ zu klassifizieren. Diese Eigenschaft der scharfen Grenzziehung ist charakteristisch für die klassische Mengenlehre: ein Element gehört entweder ganz ( $\mu = 1.0$ ) oder gar nicht ( $\mu = 0.0$ ) zu einer Menge.

Die scharfe Mengeneinteilung scheint zwar leicht anwendbar zu sein, aber die Entscheidung darüber, wie der Grenzwert sein soll stellt ein wesentliches Problem für die realitätsnahen Überlegungen dar. Tatsächlich könnten die Temperaturen, die nahe am Grenzwert liegen, wie z. B. 39,4 °C oder 39,6 °C von jemandem sowohl als starkes oder nicht starkes Fieber wahrgenommen werden. Und wie man den Grenzwert für starkes Fieber festlegen mag, ist es auf jeden Fall willkürlich und unplausibel, eine Temperatur als „nicht starkes Fieber“ einzustufen, obwohl deren Wert diese Grenze auch nur wenig unterschreitet. Deswegen wäre es durchaus sinnvoll, die scharfen Grenzen der Mengen in weiche bzw. gleitende Übergänge umzuwandeln, so dass in den Randbereichen geringfügige Schwankungen der Merkmalsausprägungen von Objekten sich als unterschiedliche Grade ihrer Zugehörigkeit zu einer Menge widerspiegeln können.<sup>775</sup>

### **6.3.2 Unscharfe Mengenlehre und Logik**

„Die Fuzzy-Logik beschäftigt sich hauptsächlich mit der Quantifizierung von vagen bzw. unscharfen Begriffen der natürlichen Sprache und dem Schlussfolgern über Aussagen mit diesen Begriffen.“<sup>776</sup> Sie beruht auf dem Konzept der Unscharfen Mengen bzw. Fuzzy-Mengen (engl.: fuzzy sets). „Fuzzy set theory provides a strict mathematical framework [...] in which vague conceptual phenomena can be precisely and rigorously studied.“<sup>777</sup> Die Theorie von Fuzzy-Mengen beinhaltet die klassische Mengenlehre und erweitert ihre Prinzipien. Smithson definiert die unscharfe Mengenlehre (engl.: fuzzy set theory) als „[...] an extension of classical set theory [...] that provides a mathematical framework for handling categories that permit partial membership (or membership in degree)“<sup>778</sup>.

Die unscharfe Mengenlehre ist eine Verallgemeinerung der klassischen Mengenlehre, indem sie die Definition gradueller Zugehörigkeiten zu Mengen ermöglicht.<sup>779</sup> Analog

---

<sup>775</sup> Vgl. Biewer (1997), S. 55.

<sup>776</sup> Stein/Lettmann (2005), S. 10.

<sup>777</sup> Zimmermann (1991), S. 6.

<sup>778</sup> Smithson (2006), S. 1.

<sup>779</sup> Vgl. Zimmermann (1993), S. 8.

dazu kann die Fuzzy-Logik als eine Verallgemeinerung der klassischen Logik angesehen werden, denn mittels dieser können graduelle Wahrheitswerte und vage Aussagen wahrgenommen werden.<sup>780</sup>

Die Fuzzy-Mengen ermöglichen die Definition sogenannter linguistischer Variablen, die Begriffe mit weichen Übergängen von „trifft zu“ zu „trifft nicht zu“ modellieren können. Sie zeichnen sich dadurch aus, dass ihre Werte keine Zahlen, sondern sprachliche (linguistische) Terme sind. Dies gilt für viele Begriffe, wie sie in der Alltagssprache verwendet werden und durchaus zur Verständigung und Entscheidungsfindung genügen. Sie sind häufig nicht scharf abgegrenzt und können mit Fuzzy-Mengen angemessen modelliert werden. Beispiele für diese Begriffe sind z. B.: „kaum“, „ausreichend“, „ziemlich“, „sehr“ u. a.

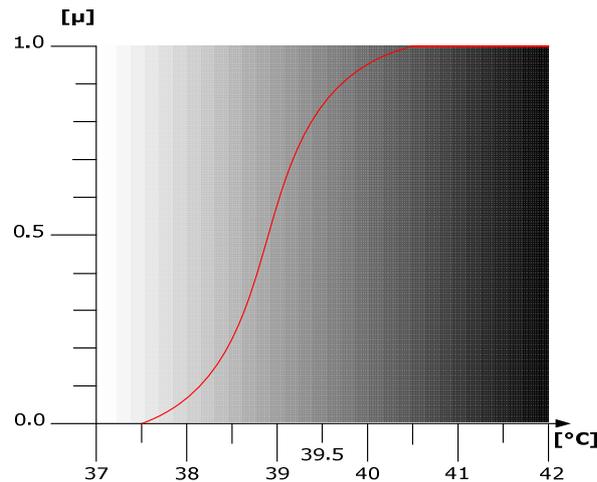
Wie im vorigen Kapitel bereits erwähnt, zerfällt in der Booleschen Logik nach dem Satz der Zweiwertigkeit die Menge aller Aussagen disjunkt in die Menge aller wahren Aussagen und die Menge aller falschen Aussagen. Es gibt also genau zwei Wahrheitswerte: „wahr“ und „falsch“. Dieses binäre System eindeutiger Zugehörigkeiten hat sich in vielen Bereichen mathematischer Berechnungen bewährt. Für manche Anwendungen sind jedoch gleitende, abgestufte Übergänge zwischen „Zugehörigkeit“ oder „Nichtzugehörigkeit“ zu einer Menge, zwischen Zutreffen und Nichtzutreffen eines Begriffs wünschenswert. Ein gutes Beispiel dafür stellt der Ladezustand eines Akkus dar. Ein Akku hat nicht nur die Ladezustände „leer“ und „voll“, sondern besitzt beliebig viele Zwischenzustände. Mit einer binärlogischen Beschreibung könnte ein Batterieladegerät nicht arbeiten. Das Ladegerät würde immer um die scharfe Grenze zwischen „leer“ und „voll“ hin- und herwechseln und der Akku somit nie einsatzbereit sein. In solchen Fällen kann die Fuzzy-Logik Abhilfe schaffen, da sie nicht nur die beiden Aussagen „leer“ und „voll“ kennt, sondern auch alle dazwischenliegenden Zustände, wie „etwas“, „ungefähr“, „ziemlich“ u. a. So erlaubt sie für das obige Beispiel die Aussage „der Akku ist ziemlich voll“. Diese Unterteilung des Ladezustands in beliebig viele Zwischenzustände zwischen von „leer“ bis „voll“ kommt der plausiblen Denkweise des Menschen näher als die binäre Logik.

Um die Idee der Fuzzy-Logik mathematisch umzusetzen, ist es notwendig, die Bildmenge der charakteristischen Funktion von zwei Werten,  $\{0, 1\}$ , auf mehrere Werte zu erweitern. Der Gründer von Fuzzy-Mengen-Theorie Lotfi A. Zadeh hat als Bildmenge das Intervall  $[0,1]$  vorgeschlagen, also einen Wertebereich, der die klassischen Werte „null“ und „eins“ als Randpunkte einschließt und zugleich unendlich

---

<sup>780</sup> Vgl. Zimmermann (1993), S. 8.

viele und beliebig feine Graduierungen abzubilden erlaubt. Durch die Verwendung einer stetigen Funktion lassen sich dann die sachlich angemessener erscheinenden „gleitenden“ Übergänge problemlos erzielen, wie in der folgenden Abbildung zu sehen ist. Geringfügige Temperaturänderungen im Bereich 39,5 °C machen sich dann nicht mehr in einer sprunghaften Antwort („starkes Fieber“ oder „nicht starkes Fieber“), sondern vielmehr in geringfügigen Verschiebungen der Zugehörigkeitsgrade zu den Mengen „starkes Fieber“ und „nicht starkes Fieber“ geltend.



**Abbildung 48: „starkes Fieber“ als Fuzzy-Menge**  
Quelle: in Anlehnung an Biewer (1997), S. 56.

Eine unscharfe Menge  $\tilde{M}$  (Fuzzy-Menge) wird charakterisiert durch eine Funktion  $\mu$  von einer Grundmenge  $\Omega$  in das reelle Einheitsintervall<sup>781</sup>:

$$\mu_{\tilde{M}}(x): \Omega \rightarrow [0, 1].$$

Die charakteristische Funktion  $\mu$  wird Zugehörigkeitsfunktion (engl.: membership function) genannt, ihre Werte heißen Zugehörigkeitsgrade (engl.: grades of membership). Die Menge aller über einer Grundmenge bildbaren Fuzzy-Mengen wird durch Großbuchstaben mit einer Tilde ( $\tilde{\phantom{X}}$ ) darüber bezeichnet. Ist z. B.  $X$  eine Menge von Objekten, die hinsichtlich einer unscharfen Aussage zu bewerten sind, so heißt:

$$\tilde{A} = \{ (x, \mu_{\tilde{A}}(x)) \mid x \in X \}$$

eine unscharfe Menge auf  $X$ .<sup>782</sup>

<sup>781</sup> Vgl. Biewer (1997), S. 56.

<sup>782</sup> Vgl. Zimmermann (1991), S. 12 und (1993), S. 8.

Hierbei ist  $\mu_A : X \rightarrow \mathfrak{R}$  eine reellwertige Funktion. Die Zugehörigkeitsfunktion kann als eine verallgemeinerte charakteristische Funktion angesehen werden. Sind ihre Werte auf das Intervall von 0 bis 1 beschränkt, spricht man von einer „normierten unscharfen Menge“.<sup>783</sup> Gewöhnliche Mengen (im Sinne Cantors) lassen sich als spezielle Fuzzy-Mengen begreifen, in denen alle Elemente den Zugehörigkeitsgrad eins aufweisen.<sup>784</sup> Enthält eine Fuzzy-Menge genau ein Element mit einem Zugehörigkeitsgrad größer als null, heißt sie Fuzzy-Einermenge (engl.: fuzzy singleton).<sup>785</sup>

### 6.3.3 Zugehörigkeitsfunktion

Die unscharfen Mengen werden über ihre Zugehörigkeitsfunktionen beschrieben, daher stellt die Zugehörigkeitsfunktion einen zentralen Begriff in der Fuzzy-Logik dar. Seine Interpretation wurde mit der Definition der Fuzzy-Mengen im vorigen Kapitel bereits angedeutet. Der Grad, in dem der Wert einer Basisgröße (z.B. ein Temperaturmesswert) den einzelnen Termen („kalt“, „kühl“, warm oder „heiß“) einer linguistischen Variablen (Raumtemperatur) entspricht, wird als Zugehörigkeitsgrad bezeichnet. Ist die Basisvariable eine kontinuierliche Variable, so wird dieser Grad an Zugehörigkeit durch eine mathematische Funktion beschrieben. Diese Zugehörigkeitsfunktion  $\mu_A(x)$  sagt aus, ob ein Wert  $x$  zu einer Umgebung von  $x_0$  hinzugehört ( $\mu_A(x) \neq 0$ ) oder nicht ( $\mu_A(x) = 0$ ). Je nach Wert von  $\mu_A(x)$  ist die Zugehörigkeit von  $x$  zu  $x_0$  groß ( $\mu$  nahe 1) oder klein ( $\mu$  nahe 0). Dies kann an einem Beispiel „Raumtemperatur“ veranschaulicht werden. An diesem Beispiel wird erneut der Vorteil von unscharfen gegenüber scharfen Mengen deutlich, da die stetige Zugehörigkeitsfunktion keine scharfen Grenzen aufzeigt und damit die Temperatur eines Raumes genauer beschreiben kann (siehe Abb. 49).

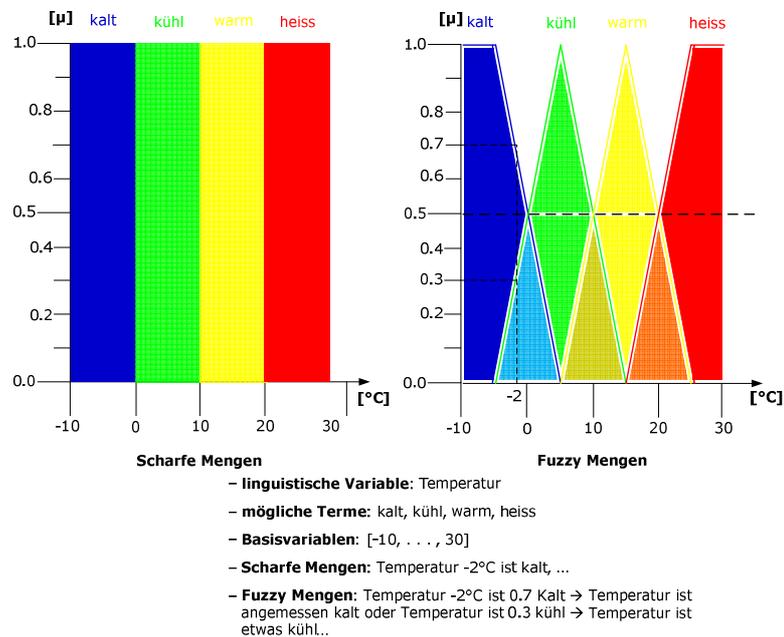
Die linguistischen Variablen, wie z. B. die Aussage „Temperatur ist kalt“, werden in mehrere genauer spezifizierende Aussagen - die „linguistischen Termen“ oder auch „Fuzzy-Mengen“ - unterteilt. In dem gewählten Beispiel bietet sich die Aufteilung in „Temperatur ist angemessen kalt“, „Temperatur ist etwas kühl“, „Temperatur ist ganz heiß“ u. a. an. Diese linguistischen Terme überschneiden sich in den meisten Fällen. Der Übergang zwischen kalter und kühler Temperatur ist nicht an einem winzigen

<sup>783</sup> Vgl. Zimmermann (1993), S. 8.

<sup>784</sup> Vgl. Biewer (1997), S. 56.

<sup>785</sup> Vgl. Biewer (1997), S. 56. In der Fuzzy-Set Theorie werden zur Repräsentation von scharfen einelementigen Mengen sogenannte *Singletons* verwendet. Das sind Funktionen, deren Funktionswerte nur an einer einzigen Stelle von 0 unterscheiden.

Temperaturwechsel auszumachen, sondern vollzieht sich dadurch, dass der Erfüllungsgrad (der Wahrheitsgehalt) des einen Terms langsam abnimmt, wogegen der Erfüllungsgrad des anderen weiter zunimmt.



**Abbildung 49: Scharfe und Fuzzy-Mengen am Beispiel „Raumtemperatur“**

Quelle: eigene Darstellung

Die Terme einer linguistischen Variablen werden also durch ihre jeweiligen Zugehörigkeitsfunktionen definiert, indem jedem Wert der kontinuierlichen Basisgröße ein Zugehörigkeitsgrad des entsprechenden Terms zwischen 0 und 1 zugeordnet wird. Die Zugehörigkeitsfunktionen der Terme sind durch bestimmte Definitionspunkte (auch Stützstellen genannt) eindeutig festgelegt.

Die Bestimmung dieser Definitionspunkte ist aber problematisch. Denn es existieren keine einheitlichen Regeln, nach denen solche Punkte definiert werden können. Mit der Bestimmung der Stützstellen kann man beliebige parametrische Zugehörigkeitsfunktionen erhalten, die einen dreieckförmigen ( $\lambda$ -Typ), trapezförmigen ( $\pi$ -Typ), s-, z- oder glockenförmigen Funktionsverlauf aufweisen.<sup>786</sup> Als Sonderform existieren zudem sogenannte Singletons (Zugehörigkeitsfunktion mit nur mit einem Stützpunkt), die dazu verwendet wird, um Rechenzeit zu sparen.<sup>787</sup> Empirische Untersuchungen können für die Bestimmung des Funktionsverlaufs herangezogen werden, jedoch wird dadurch nur zum Teil möglich, die geeignete Form der Zugehörigkeitsfunktion für eine spezielle Anwendung herauszufinden. Das

<sup>786</sup> Verschiedene Möglichkeiten der Konstruktion von Zugehörigkeitsfunktionen sind in: Biewer (1997), S. 57f. und Bothe (1995), S. 27f. zu finden.

<sup>787</sup> Vgl. Tizhoosh (1997), S. 20.

bedeutet wiederum, dass sehr oft bei der Fuzzy-Modellierung die Formen der Zugehörigkeitsfunktion verändert und adaptiert werden müssen, um das optimale Ergebnis zu erzielen.<sup>788</sup>

Die Form der Zugehörigkeitsfunktion hat einen besonderen Einfluss auf das Verhalten der Fuzzy- Modellierung.<sup>789</sup> Die Wahl der passenden Zugehörigkeitsfunktion für eine bestimmte Problemstellung ist daher sehr wichtig und eine Sache der Erfahrung. In der Regel wird jede Basisgröße bzw. -variable durch Zugehörigkeitswerte zu zwei Termen einer linguistischen Variablen zugeordnet. Außerdem ergibt die Summe der Zugehörigkeitswerte immer eins.<sup>790</sup>

### **6.3.4 Unbestimmtheit, Unsicherheit, Unschärfe, Wahrscheinlichkeit, Möglichkeit**

Bevor im nächsten Kapitel die Grundlagen der Fuzzy-Mengen-Theorie erläutert werden, sollen zunächst einige wesentliche Begriffe kurz erläutert und voneinander abgegrenzt werden.

Manchmal ist es unmöglich festzustellen, ob bestimmte Elemente zu einer Menge gehören oder nicht. Ab wann kann man z. B. einen Hügel als einen Berg bezeichnen oder einen Teich als einen See. Die klassische Logik und Mengenlehre beschäftigten sich eine lange Zeit mit dieser ganz leicht scheinenden Klassifizierungsaufgabe. Eine mögliche Antwort für diese Klassifizierung wäre, dass ein Hügel ab einer bestimmten Höhe ein Berg ist oder eine Kurzgeschichte ab einer bestimmten Seitenzahl – z. B. 100 – ein Roman. Diese Klassifizierung führt jedoch zu Absurditäten, denn ein literarisches Werk mit 101 Seiten ist dann ein Roman, mit 99 Seiten eine Kurzgeschichte. Wenn der Verfasser aber eine größere Schrift verwendet, wird die Kurzgeschichte zum Roman. Die Kurzgeschichte und den Roman voneinander scharf abzugrenzen, scheint daher nicht so leicht zu sein.

Auch in der Realität ist es nicht selten, dass man nicht mit voller Sicherheit sagen kann, welcher Gegenstand zu welcher Klasse oder Menge gehört (Wann kann ein Stuhl als ein Kunstobjekt angesehen werden?). Man spricht dann von *Unbestimmtheit*. Ein Beispiel dafür wäre z. B. das Sorites (Haufen-)Paradoxen, das man von der Antike kennt: Gegeben sei ein Sandhaufen. Nimmt man ein Sandkorn weg, bleibt immer noch

---

<sup>788</sup> Vgl. Tizhoosh (1997), S. 20.

<sup>789</sup> Vgl. Tilli (1992), S. 230.

<sup>790</sup> Das ist aber keine zwingende Voraussetzung, sondern dient nur der besseren Veranschaulichung; es besteht auch die Möglichkeit, Fuzzy-Mengen so zu definieren, dass die Summe aller Funktionswerte auf einer Basisvariablen ungleich eins ist. In diesem Fall liegt eine Doppelbenennung von Zuständen vor bzw. eine mehrfache Abbildung der Basisgröße auf 1, siehe dazu: Kempf (2001), S. 10.

ein Sandhaufen. Daran ändert sich nichts, auch wenn man ein zweites Sandkorn wegnimmt, dann ein drittes, viertes usw. Am Ende schrumpft er auf nur mehr ein Sandkorn. Kann man ihn nun noch als Sandhaufen bezeichnen?

In der klassischen Mengenlehre Cantors löst man diese Problematik mittels Definitionen: Man stellt die Grenze zwischen zwei Zuständen einfach fest. So bildet eine bestimmte Anzahl von Sandkörnern einen Sandhaufen; nimmt man davon eines weg, handelt es sich nicht mehr um einen Sandhaufen. Cantors Mengen sind scharf und exakt definiert, und kein Element kann sich auf die Grenzlinie – welche die Welt in ein „Innen“ und „Außen“, in „wahr“ und „falsch“ unterteilt – zwischen den Mengen stellen. Wenn aber eine Menge keine exakten Grenzen besitzt, versagen dann die Regeln der klassischen Mengenlehre und damit auch die klassische Logik, indem sie zu unlogischen Schlüssen gelangt.

Dagegen löst die unscharfe Mengenlehre die Haufenparadoxe ohne Probleme auf, indem sie bestimmte Grade von Zugehörigkeit zu einer Klasse zulässt; also etwa tausend Sandkörner sind zu einem gewissen Grad mehr Sandhaufen als hundert. Mit jedem Sandkorn, das entfernt wird, verringert sich der Zugehörigkeitsgrad zur Klasse „Haufen“. Er fällt von 1.0 über beispielsweise 0.7 und 0.3 kontinuierlich bis zum Wert 0. Unscharfe Mengen erlauben so eine graduelle Abstufung der Wahrheit jenseits der strikten Wahrheitsalternativen „1“ (= ein Haufen) und „0“ (= kein Haufen) von Cantor Mengen. „Wahr“ und „Falsch“ bzw. „1“ und „0“ fungieren dann als Grenzwerte eines breiten Spektrums von Wahrheitswerten. Somit scheint das Konzept der unscharfen Mengen adäquat zur Bestimmung des Wahrheitsgehalts einer Aussage in *Unsicherheit* (engl.: uncertainty) anzuwenden.

Das Konzept der Unschärfe als Teilaspekt von Unsicherheit bildet die Grundlage und Ausgangsbasis für die Fuzzy-Mengen-Theorie. Deswegen soll im Folgenden der Begriff *Unschärfe* (engl.: fuzziness) und sein Zusammenhang mit der Unsicherheit beleuchtet werden.

Die Unsicherheit entsteht durch den Mangel an präzisen und vollständig bestätigten Informationen oder aufgrund widersprüchlicher Informationen.<sup>791</sup> Der Begriff „Unsicherheit“ hat aber einen breiten semantischen Inhalt.

So zählt das Webster's New Twentieth Century Dictionary folgende Bedeutungen für den Begriff „unsicher“ auf:<sup>792</sup>

---

<sup>791</sup> Vgl. Schulte (1993), S. 28.

<sup>792</sup> Vgl. Klir/Folger (1998), S. 138.

- Not certainly known; questionable; problematical.
- Vague; not definite or determined.
- Doubtful; not having certain knowledge; not sure.
- Ambiguous.
- Not steady or constant; varying.
- Liable to change or vary; not dependable or reliable.

Klir und Folger klassifizieren in ihrem Buch „Fuzzy Sets, Uncertainty and Information“ diese Bedeutungen des Begriffs „unsicher“ in zwei Kategorien: *Vagheit* (engl.: vagueness), *Verschwommenheit* oder *Unschärfe* (engl.: fuzziness); interpretiert als die Schwierigkeit scharfe und präzise Unterscheidungen vorzunehmen, und *Mehrdeutigkeit* (engl.: ambiguity), assoziiert mit one-to-many Relationen, also Situationen, in denen die Wahl zwischen zwei oder mehreren Alternativen unspezifiziert bleibt.<sup>793</sup> Diese beiden Formen von Unsicherheit sind mit verschiedenen Konzepten verbunden.

<b>Arten von Unsicherheit</b>			
Konzept	Teilkonzepte	Theorie	Arten
<b>Vagueness</b>	Fuzziness (Unschärfe) Haziness (Verschwommenheit) Cloudiness (Unklarheit) Unclearness (Undeutlichkeit) Indistinctiveness (Ununterscheidbarkeit) Sharplessness (Unscharfeheit)	„ [...] the concept of a fuzzy set provides a basic mathematical framework for dealing with vagueness.“ <sup>794</sup>	
<b>Ambiguity</b>	Nonspecificity (Ungenauigkeit) One-to-many Relation (1 zu n Beziehung) Variety (Vielfalt) Generality (Allgemeinheit) Diversity (Diversität) Divergence (Divergenz)	„ [...] the concept of a fuzzy measure provides a general framework for dealing with ambiguity.“ <sup>795</sup>	<i>nonspecificity in evidence</i> (erscheinbare Ungenauigkeit) <i>dissonance in evidence</i> (erscheinbare Dissonanz) <i>confusion in evidence</i> (erscheinbare Unordnung)

**Tabelle 7: Konzepte von Unsicherheit**

Quelle: eigene Darstellung, in Anlehnung an Klir/Folger (1998), S. 139.

Die Tabelle 7 gibt eine Übersicht über diese Konzepten: Sie macht deutlich, dass

<sup>793</sup> Vgl. Klir/Folger (1988), S. 138.

<sup>794</sup> Klir/Folger (1988), S. 139.

<sup>795</sup> Klir/Folger (1988), S. 139.

unscharfe Mengen (engl.: fuzzy sets) und fuzzy-Maße (engl.: fuzzy measures) zwei verschiedene Konzepte von Unsicherheit darstellen, die unter den Namen Vagueness und Ambiguity zusammengefasst werden können.

Zur Veranschaulichung dieser Verschiedenheit beider Arten von Unsicherheit kann folgendes Beispiel dienen:<sup>796</sup>

Eine beliebige Person, deren Alter nicht genannt wird, soll in eine Altersgruppe durch Schätzung eingeteilt werden. Die Evidenz des Alters, die für die sichere Zuteilung dieser Person in eine Altersgruppe der Zwanzigjährigen, Dreißigjährigen, Vierzigjährigen oder Fünfzigjährigen notwendig würde, könnte mangelhaft sein. Dabei ist zu beachten, dass zwar die Altersmengen scharf (crisp) und exakt sind - also es besteht keine Vagheit über deren Grenzen - jedoch die Schätzung, nach der die Zuteilung erfolgen wird, nicht sicher ist. Der Grad an Sicherheit nimmt zu, je mehr Evidenz für ein bestimmtes Alter dieser Person besteht. Erst wenn man einen absoluten Beweis für das exakte Alter der Person findet, wäre es möglich, sie mit einer Sicherheit von 1 aus dem Intervall  $[1,0]$  einer scharfen Menge (Altersgruppe) zuzuteilen. Demzufolge bekommt die Altersgruppe mit dem höchsten Schätzwert den Grad „1“ und die anderen den Grad „0“. Ein *fuzzy measure* ordnet also jeder scharfen Menge einen graduellen Wert zu, der den Grad der Evidenz oder die subjektive Sicherheit in Bezug auf die Zugehörigkeit eines Elementes zu dieser exakten Menge indiziert.

Im Gegensatz dazu würde dieses Zuteilungsproblem im Kontext von *fuzzy sets* so aussehen, dass das Alter der Person bekannt ist, allerdings wäre die Menge, der sie zugeteilt werden soll, unscharf, das heißt, es muss festgestellt werden, zu welchem Grad diese Person in verschiedenen Altersgruppen repräsentiert ist. Soll eine Person, deren Alter bspw. 30 Jahre beträgt, der Menge „junge Menschen“ zugeordnet werden oder der Menge „Menschen mittleren Alters“?

Außer der obigen Unterscheidung sind in der Literatur noch andere Unterscheidungen zu finden. Beispielsweise nennt Kruse zwei unterschiedliche Ansätze von Vagheit. Nach ihm werden die Fuzzy-Mengen einerseits zur Darstellung impräziser und widersprüchlicher Information über einen unbekanntem, aber existierenden Messwert verwendet (epistemische Interpretation), andererseits werden sie dort angewendet, wo die Unschärfe auf der Nichtunterscheidbarkeit bzw. Nichtmessbarkeit von Daten beruht.<sup>797</sup>

---

<sup>796</sup> Vgl. Klir/Folger (1988), S. 12.

<sup>797</sup> Vgl. Kruse /Gebhardt/ Klawonn (1993), S. 44.

Weiters unterscheidet Zimmermann zwischen drei Arten von Unsicherheit:<sup>798</sup>

1. *zufällige (stochastische) Unsicherheit*: Diese Unsicherheit tritt auf, wenn man nicht sicher ist, ob ein Ereignis eintritt. Das Eintreten eines Ereignisses ist also mit einer gewissen Unsicherheit behaftet, kann jedoch mit Hilfe der Wahrscheinlichkeitstheorie adäquat modelliert werden. Ein gutes Beispiel hierfür ist das Würfeln einer „6“ mit der Wahrscheinlichkeit 1/6. Über eine Vielzahl von Ereignissen kann man mit statistischen Verfahren zu einer allgemeinen und meist genaueren Beschreibung des Eintretens eines bestimmten Ereignisses kommen.
2. *lexikale (linguistische bzw. sprachliche) Unsicherheit*: Die lexikalische Unsicherheit beschreibt die Unsicherheit in der Interpretation einer Aussage. Als „warm“ kann z. B. eine Person eine Temperatur von 15 °C empfinden, eine andere aber vielleicht eine Temperatur von 25 °C.

„The meaning of a word might even be well defined, but when using the word as a label for a set, the boundaries within which objects do or do not belong to the set become fuzzy or vague.”<sup>799</sup> Das bedeutet, dass die Bedeutung der Ausdrücke kontextabhängig ist. So hat z. B. der Ausdruck „große Steine“ eine andere Bedeutung im Juwelierjargon als im Bergbau. Dieser inhaltliche Unterschied stellt jedoch für die menschliche Kommunikation kein Problem dar, da die Menschen in der Lage sind, aus dem Zusammenhang die Bedeutung von Wörtern und Sätzen zu erkennen. Spricht z. B. ein Mensch in Moskau von warmen Wetter, meint er wahrscheinlich 12 °C; bei einem Menschen in Kairo wäre eher 30 °C gemeint. Ein Computer ist aber nicht dazu fähig, die eigentliche Bedeutung von „warm“ aus dem Kontext abzuleiten. Menschliches Wissen kann in einer der menschlichen Verwendung entsprechenden Form vom Computer nur dann benutzt werden, wenn es inhaltlich definiert ist.

3. *informationale Unsicherheit*: Diese Art der Unsicherheit ist auf einen Informationsüberfluss zurückzuführen. Der Überfluss bedeutet hier umfassende und damit übermäßige Information, die dazu notwendig ist, um ein System zu beschreiben. Allerdings ist ein Mensch nicht in der Lage, diese große Menge an Informationen gleichzeitig wahrzunehmen und zu verarbeiten: „The complete description of a real system often would require by far more detailed data than a human being could ever recognize simultaneously, process and

---

<sup>798</sup> Vgl. Zimmermann (1995), S. 3-7.

<sup>799</sup> Vgl. Zimmermann (1991), S. 4.

understand.“<sup>800</sup> Ein System oder sogar ein Begriff lässt sich durch sogenannte Deskriptoren (beschreibende Eigenschaften) definieren, die jedoch eher subjektiv bestimmt sind. Diese Subjektivität und übermäßige Information bei der Beschreibung eines Begriffes („Kreditwürdigkeit“, „Vertrauenswürdigkeit“ oder „Sympathie“) führt dann zur Unsicherheit.

Die stochastische Unsicherheit lässt sich auf der Grundlage der Wahrscheinlichkeitstheorie modellieren. Für die linguistische und informationale Unsicherheit kann man die Theorie von Fuzzy-Mengen heranziehen.

Eine letzte und wichtige Unterscheidung betrifft die Begriffe *Wahrscheinlichkeit* und *Möglichkeit*. Man verwechselt die Möglichkeit oft mit der Wahrscheinlichkeit, jedoch werden diese Begriffe in der Fachliteratur ganz unterschiedlich gebraucht.

Um ein besseres Gefühl für die Unterschiede zwischen beiden Begriffen zu geben, ist es sinnvoll, zunächst die Unschärfe und Wahrscheinlichkeit näher zu betrachten: Beide beschreiben das Phänomen der Unbestimmtheit numerisch. Die Wahrscheinlichkeitstheorie behandelt jedoch Ja/Nein-Entscheidungen, beruht auf Unkenntnis und ist ihrer Natur nach statistisch. Die Theorie der Unschärfe dagegen handelt von Abstufungen, setzt Unkenntnis nicht voraus und ist ihrem Wesen nach unstatistisch.<sup>801</sup> Wie der Fuzzy-Logik-Theoretiker Bart Kosko bei seinem bekannten „Apfelexperiment“<sup>802</sup> schon bemerkt hat, verschwindet die Wahrscheinlichkeit mit der Information<sup>803</sup>, das heißt, je mehr Information über das einzutretende Ereignis vorhanden ist, desto weniger unbestimmt bzw. wahrscheinlicher ist es. Man spricht also bei Vorhandensein absolut vollständiger Information nicht mehr von der Wahrscheinlichkeit<sup>804</sup> und kann alle Fragen in Bezug auf dieses Ereignis präzise durch „Ja“ und „Nein“ beantworten. Im Gegensatz zur Wahrscheinlichkeit kann aber die Unschärfe auch bei vollständiger Information bestehen.<sup>805</sup>

Vor der Entwicklung der Fuzzy-Logik war die einzige Möglichkeit, die Unbestimmtheit zu beschreiben, die von Blaise Pascal (1623-1663) erfundene Wahrscheinlichkeitsrechnung – also wenn man also keine exakte Aussage über ein Ereignis machen kann, dann kann man wenigstens eine exakte Zahl über die Wahrscheinlichkeit seines Eintretens berechnen.

---

<sup>800</sup> Vgl. Zimmermann (1991), S. 3

<sup>801</sup> Vgl. McNeill/Freiberger (1996), S. 81.

<sup>802</sup> Siehe dazu: McNeill/Freiberger (1996), S. 81.

<sup>803</sup> Vgl. Drösser (1994), S. 78.

<sup>804</sup> denn diese setzt Unkenntnis voraus; das ist auch der Grund, warum sie sich mit der Zukunft beschäftigt.

<sup>805</sup> Vgl. McNeill/Freiberger (1996), S. 81.

Bei der klassischen Wahrscheinlichkeitsrechnung (engl.: probabilistic) werden Aussagen über zukünftige Ereignisse, die „zufällig eintreten können“, getroffen.<sup>806</sup> Die relative Häufigkeit bestimmt dabei die Wahrscheinlichkeit. Die Wahrscheinlichkeitsrechnung ist unter identischen Bedingungen beliebig oft wiederholbar und drückt sich folgendermaßen aus:

$$\text{Wahrscheinlichkeit} = \frac{\text{Anzahl der günstigen Ereignisse}}{\text{Anzahl aller möglichen Ereignisse}}$$

Die Theorie der Unschärfe hingegen betrachtet in erster Linie die Möglichkeit. Der Unterschied zwischen Wahrscheinlichkeit (engl.: probability) und Möglichkeit (engl.: possibility) kann mit den Worten Zadeh's wie folgt zum Ausdruck gebracht werden: „Was möglich ist, mag nicht wahrscheinlich sein, und was unwahrscheinlich ist, mag nicht unmöglich sein.“<sup>807</sup> Die Möglichkeit eines Ereignisses ist immer größer als die Wahrscheinlichkeit, was wiederum bedeutet: was völlig unwahrscheinlich ist kann trotzdem vollkommen möglich sein.<sup>808</sup> Außerdem ist die Möglichkeit ein Maß dafür, wie leicht ein Ergebnis eintreten kann.<sup>809</sup> Die Möglichkeit handelt also davon, ob ein Ereignis stattfinden kann, die Wahrscheinlichkeit hingegen davon, ob es stattfinden wird.

Zum besseren Verständnis für den Unterscheid zwischen Wahrscheinlichkeit und Möglichkeit kann folgendes Beispiel von Klir und Folger dienen: Wahrscheinlichkeit bezieht sich auf die Zahl der Menschen, die an einem bestimmten Tag Auto fahren, Möglichkeit dagegen auf die Zahl jener, die überhaupt zu irgendeinem Zeitpunkt Auto fahren können.<sup>810</sup>

Die Möglichkeit eines Ereignisses ist wie die Wahrscheinlichkeit eine Zahl zwischen 0 und 1, nur gehorcht die Möglichkeit anderen Gesetzen. Das Rechnen mit Möglichkeiten (engl.: possibilistic), das von Zadeh in 1978 als eine neue Theorie entwickelt wurde, lässt sich genauso formalisieren wie jenes von Wahrscheinlichkeiten.

Allerdings unterscheiden sich die Regeln:<sup>811</sup>

Für voneinander unabhängige Ereignisse A und B gilt in der

---

<sup>806</sup> Vgl. McNeill/Freiberger (1996), S. 77

<sup>807</sup> Drösser (1994), S. 80.

<sup>808</sup> Vgl. Drösser (1994), S. 81.

<sup>809</sup> Lotfi A. Zadeh definierte die Möglichkeit als „den Grad der Mühelosigkeit“, mit der ein Ereignis stattfinden könnte, vgl. McNeill/Freiberger (1996), S. 101.

<sup>810</sup> Vgl. McNeill/Freiberger (1996), S. 101.

<sup>811</sup> Vgl. Drösser (1994), S. 81

Wahrscheinlichkeitsrechnung:  $\text{Prob}(A \wedge B) = \text{Prob}(A) \times \text{Prob}(B)$ .

Beispiel: Die Wahrscheinlichkeit beim Würfeln zwei Sechsen hintereinander zu werfen:  
 $1/6 \times 1/6 = 1/36$ .

Für die Möglichkeit gilt hingegen:  $\text{Poss}(A \wedge B) = \min(\text{Poss}(A), \text{Poss}(B))$ .

Die Möglichkeit einer Kette von Ereignissen ist also die Möglichkeit des schwächsten Gliedes. Die Regel besagt auch, dass die Möglichkeitsverteilungen im Wesentlichen genauso gehandhabt werden wie die Zugehörigkeiten bei Fuzzy-Mengen.<sup>812</sup>

## 6.4 FUZZY-SET-THEORIE

### 6.4.1 Grundlagen

Die Fuzzy-Set-Theorie, oder auch unscharfe Mengenlehre, wurde erstmals 1965 von Lotfi A. Zadeh, ehemaligem Professor für Elektrotechnik an der Universität von Kalifornien in Berkeley, in einer Publikation erwähnt. Diese Theorie kann als eine Vereinigung verschiedener Disziplinen wie Systemtheorie, Philosophie der Vagheit, Mengenlehre, mehrwertige Logik und Linguistik angesehen werden.<sup>813</sup>

Der Hauptgegenstand dieser Theorie ist der Umgang mit unscharfen Mengen.

Zu Beginn waren viele Wissenschaftler gegenüber der neuen Herangehensweise zu Mengenlehre und Logik eher skeptisch eingestellt. Dieser Paradigmenwechsel in der Wissenschaft war als unpräzise und wissenschaftlich unseriös erachtet worden.<sup>814</sup> Trotzdem führten die ersten Anwendungen in den 70ern zu einer langsamen Akzeptanz der Fuzzy-Set-Theorie. In den darauffolgenden Jahren verstärkte sich diese Theorie durch ihre fortwährende Akzeptanz, und anfangs der 90er Jahre wurde schließlich begonnen, sie auch an Universitäten als vollständige Lehre anzubieten. Heute findet die Fuzzy-Theorie in vielen Bereichen Anwendung, wie bspw. in der Regelungstechnik und der Künstlichen Intelligenz-Forschung.

Die Fuzzy-Theorie versucht auf Basis der Informations- und Wissensrepräsentation die Komplexität, welche sich unmittelbar aus der Verfügbarkeit großer Mengen von Information und der damit verbundenen Unsicherheit resultiert, zu vereinfachen und einen Weg für die adäquate Verwendung von Information zu finden.<sup>815</sup>

---

<sup>812</sup> Vgl. Drösser (1994), S. 81.

<sup>813</sup> Vgl. McNeill/Freiburger (1996), S. 33.

<sup>814</sup> Vgl. McNeill/Freiburger (1996), S. 16.

<sup>815</sup> Vgl. Klir/Folger (1988), S. 1ff. und 231f.

Mit Hilfe der Fuzzy-Theorie lässt sich menschliches (und damit meist unscharfes bzw. unpräzises) Wissen darstellen. Fuzzy-Wissenssysteme können unscharfes Wissen mathematisch behandeln und damit realitätsnäher repräsentieren, analysieren, klassifizieren und simulieren. Sie verarbeiten im Gegensatz zu herkömmlichen Systemen - die nur binäre Werte verarbeiten können – auch Zwischenwerte, wodurch die Möglichkeit entsteht, unscharfe, verbale Angaben wie z. B. „ein bisschen“, „ziemlich“, „sehr“ numerisch modellieren zu können. Die Fuzzy-Logik-unterstützte Computerprogramme sind damit in der Lage, Information näher am menschlichen Denken zu verarbeiten als übliche Programme.<sup>816</sup> Demnach sind die Anwendungsgebiete<sup>817</sup> sehr breit gefächert und dienen der Wissensrepräsentation, Datenanalyse, Mustererkennung und Entscheidungsunterstützung. In der Literatur häufig genannte Anwendungen sind Expertensysteme (z. B. medizinische Diagnosesysteme), Fuzzy-Regelsysteme, Mustererkennungssysteme (Fuzzy Clustering) und Entscheidungsunterstützungs- (Decision Support) und Optimierungssysteme (Operation Research).<sup>818</sup>

Die folgende Tabelle zeigt eine mögliche Klassifizierung der Fuzzy-Logik-basierten Anwendungen nach Zimmermann.

<b>Anwendungsbereiche</b>	<b>Teilbereiche</b>	<b>Fuzzy-Techniken</b>
Mathematik (Verallgemeinerung der traditionellen Mathematik)	Algebra, Logik, Graphentheorie, Topologie,...	Fuzzy-Logik, Approximate Reasoning
Algorithmen	Clustering-Methoden, Steuerungsalgorithmen, mathematische Programmierung,...	Fuzzy Clustering, fuzzy lineares Programmieren, fuzzy dynamisches Programmieren
Standardmodelle	Transportwesen-Modelle, Inventurkontrollmodelle, Instandhaltungsmodelle,...	Standard Operations Research Modelle
„Real-world-Probleme“	Expertensysteme, Regelungstechnik	Expertensysteme, Fuzzy Regelsysteme (in denen die Fuzzy-Set-Theorie weitgehend angewendet wird, und sie ist am nächsten zu realen Anwendungen)

**Tabelle 8: Anwendungsbereiche der Fuzzy-Set-Theorie**

Quelle: eigene Darstellung, in Anlehnung an Zimmermann (1991), S. 129.

Im folgenden Kapitel wird eine Einführung in die wichtigsten mathematischen Grundlagen der Fuzzy-Set-Theorie gegeben.

<sup>816</sup> Vgl. Zimmermann (1991), S. 179.

<sup>817</sup> Siehe dafür: Biewer (1997), S. 30ff.

<sup>818</sup> Vgl. Zimmermann (1991), S. 171-332.

### 6.4.2 Mathematische Grundlage der Fuzzy-Set-Theorie

Eine klassische Menge  $M$  ist eine Sammlung von Elementen  $x$  aus einer Grundmenge  $\Omega$ . Jedes Element  $x$  kann einer beliebigen Teilmenge  $A$  ( $A \subseteq \Omega$ ) angehören, oder nicht. Die Aussage „ $x$  gehört zu  $A$ “, kann also entweder wahr (1) oder falsch (0) sein. Die der klassischen Menge zugehörigen Elemente werden mit der charakteristischen Funktion  $\mu_A(x): \Omega \rightarrow \{0, 1\}$  beschrieben, wobei „1“ die Zugehörigkeit von Element  $x$  zur jeweiligen Teilmenge  $A$  beschreibt, und „0“ die Nicht-Zugehörigkeit.<sup>819</sup>

Dann gilt:

$$\chi = \begin{cases} 1, & \text{falls } x \in A \\ 0, & \text{sonst.} \end{cases}$$

In der klassischen Mengenlehre wird also von einer eindeutigen und erkennbaren Zugehörigkeit bzw. Nicht-Zugehörigkeit eines Elements zu einer Menge ausgegangen.

In der unscharfen Mengenlehre erlaubt dagegen die charakteristische Funktion (Zugehörigkeitsfunktion) einen Grad an Zugehörigkeit des Elements  $x$  aus dem Grunduniversum bzw. der Grundmenge  $\Omega$  zu einer Fuzzy-Menge  $\tilde{A}$ , der innerhalb eines Intervalls  $[0,1]$  verläuft:

$$\mu_{\tilde{A}}(x): \Omega \rightarrow [0, 1].$$

$\mu_{\tilde{A}}(x)$  wird Zugehörigkeitsfunktion genannt und stellt den Grad bzw. Zugehörigkeitswert (engl.: membership value) dar, mit dem ein Element  $x$  in der unscharfen Menge  $\tilde{A}$  enthalten ist. Ist z. B.  $X$  eine Menge von Objekten, die hinsichtlich einer unscharfen Aussage zu bewerten sind, so heißt:

$$\tilde{A} = \{ (x, \mu_{\tilde{A}}(x)) \mid x \in X \}, \text{ eine unscharfe Menge auf } X. \quad ^{820}$$

Die Fuzzy-Mengen sind gut geeignet, vage Konzepte zu repräsentieren. Die

<sup>819</sup> Vgl. Zimmermann (1991), S. 11.

<sup>820</sup> Vgl. Zimmermann (1991), S. 12 und (1993), S. 8.

Grundmenge kann sowohl kontinuierlich als auch diskret sein, wobei im ersten Fall eine Darstellung einer unscharfen Menge  $\tilde{A}$ , als Liste von zwei Tupeln erfolgt. Dabei enthält jedes Tupel ein Element der Grundmenge und seinen Zugehörigkeitswert zu  $\tilde{A}$ .

Im zweiten Fall werden die unscharfen Mengen durch kontinuierliche Funktionen modelliert. Als Beispiel dient die Modellierung des menschlichen Alters. Die Frage ist dabei, zu welchem Grade jemand mit 45 Jahren als „alter Mensch“ zu bezeichnen ist.

Sei nun  $x$  die Zahl, die ein Mensch als Alter annehmen kann. Dann wird die Menge  $X$  des möglichen Alters eines Menschen folgendermaßen beschrieben:

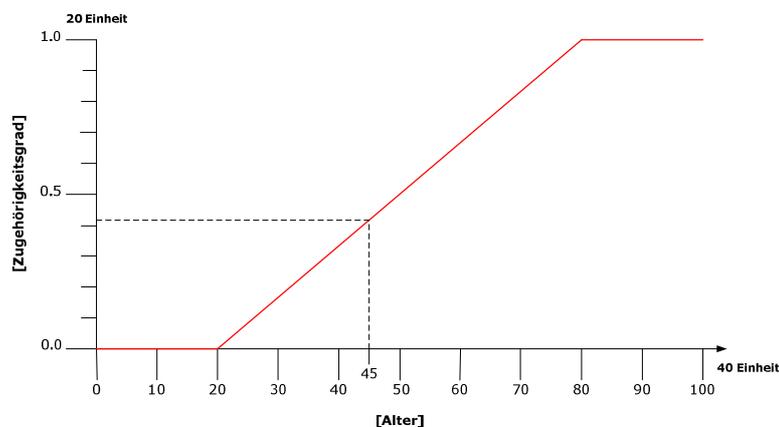
$$X = \{x \mid x \in \mathfrak{R}, 0 \leq x \leq 110\}.$$

Die verschiedenen Darstellungsformen der Fuzzy-Menge  $\tilde{A}$  sind dann:<sup>821</sup>

1. Die unscharfe Menge  $\tilde{A}$  „alte Menschen“ kann mittels diskreten Wertepaaren (Element:  $x$ , Zugehörigkeitsgrad:  $\mu_{\tilde{A}}(x)$ ) festgelegt werden:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}, \text{ d. h. ; } \tilde{A} = \{(26, 0.1), (62, 0.7), (44, 0.4), (38, 0.3), (56, 0.6), (74, 0.9), (68, 0.8), (32, 0.2), (50, 0.5)\}.$$

2. Die graphische Darstellung sieht folgendermaßen aus:



**Abbildung 50: Zugehörigkeitsfunktion für die Fuzzy-Menge „alter Mensch“**

Quelle: eigene Darstellung

<sup>821</sup> Vgl. Zimmermann (1991), S. 12.

3. Die parametrische Darstellung der Zugehörigkeitsfunktion  $\mu_{\tilde{A}}(x)$  hat folgende Schreibweise:

$$\forall x \in X : \mu_{\tilde{A}}(x) = \left\{ \begin{array}{l} 1, \text{ falls } x \geq 80 \\ \frac{5x - 40}{6}, \text{ falls } 20 < x < 80 \\ 0, \text{ falls } x \leq 20 \end{array} \right\}$$

Eine unscharfe Menge ist die Verallgemeinerung einer klassischen Menge und die Zugehörigkeitsfunktion eine Verallgemeinerung der charakteristischen Funktion.<sup>822</sup> In der Fuzzy-Set-Theorie kann ein Element auch eine Zugehörigkeit von 0.0 oder 1.0 (als Extremwerte) zu einer unscharfen Menge besitzen, und somit kann die unscharfe Menge den gleichen Charakter wie eine klassische Menge aufweisen. Allerdings erfordert diese Herangehensweise die exakte Definition von Zugehörigkeitsfunktionen bzw. -graden von Elementen zu einer Fuzzy-Menge.

Die folgenden Erweiterungen und Verallgemeinerungen der Fuzzy-Mengen dienen der Berücksichtigung des Umstandes, dass der Denkprozess des Menschen meist keinen exakt definierten Zugehörigkeitsfunktionen bzw. -graden folgt.<sup>823</sup>

Eine Verallgemeinerung von Fuzzy-Mengen kann durch *L-Fuzzy-Mengen* vorgenommen werden. Obwohl die Zugehörigkeit meist mit Werten aus dem Intervall  $[0,1]$  repräsentiert wird, besteht auch die Möglichkeit, sie durch jede beliebige Menge mit gänzlicher oder teilweiser Ordnung darzustellen. Die Elemente dieser Menge beinhalten dann keine numerischen Werte, sollen aber einer gänzlichen oder teilweisen Ordnung unterliegen, die verschiedene Stärken von Zugehörigkeit repräsentieren kann.<sup>824</sup> Diese verallgemeinerte Zugehörigkeitsfunktion hat die Form:

$$\mu_{\tilde{A}} : X \rightarrow L$$

$L$  bezeichnet eine beliebige Menge, die zumindest teilweise geordnet sein muss.<sup>825</sup> Diese Art von unscharfen Mengen werden L-Fuzzy-Mengen genannt. Die wichtigste L-Fuzzy-Menge ist  $L = [0,1]^n$ .

Die Bezeichnung  $[0,1]^n$  ist die Schreibweise des Kartesischen Produkts in

<sup>822</sup> Vgl. Zimmermann (1991), S. 13.

<sup>823</sup> Vgl. Zimmermann (1991), S. 24.

<sup>824</sup> Vgl. Klir/Folger (1988), S. 11

<sup>825</sup> Vgl. Klir/Folger (1988), S. 11.

Kurzform:  $\underbrace{[0,1] \times [0,1] \times \dots \times [0,1]}_{n\text{-Mal}}$ .

Indem man für  $L$  statt des Einheitsintervalls  $[0,1]$  das Kartesische Produkt wählt, können Konzepte mit mehreren Unschärfe-Dimensionen modelliert werden.<sup>826</sup> Obwohl die Menge mit Zugehörigkeitswerten  $[0,1]$  eine totale Ordnung<sup>827</sup> besitzt, sind Mengen mit  $[0,1]^n$  für  $n \geq 2$  sind nur partial<sup>828</sup> geordnet.

Die Unangemessenheit der Unterscheidung reellwertiger Zugehörigkeitsgrade wie 0,774 und 0,773 zur kognitiven Modellierung kann als wesentliche Motivation für die Verwendung von L-Fuzzy-Mengen betrachtet werden. Diese Unangemessenheit erweist sich sowohl in der Wissensakquisition als auch in der Anwendung bei der Interpretation der Zugehörigkeitsgrade als problematisch. Gegenüber dieser Problematik einer quantitativen Übermodellierung ist in vielen Fällen eine Modellierung adäquater, in der eine lediglich qualitativ abgestufte Skala von Zugehörigkeitsgraden zugrunde gelegt wird.

Eine Erweiterung der Fuzzy-Mengen stellen sogenannte *Fuzzy-Mengen vom Typ 2* dar.

Im Gegensatz zu den gewöhnlichen Fuzzy-Mengen, die Typ-1-Fuzzy-Mengen sind und ihre Zugehörigkeitswerte aus reellen Nummern bestehen, besitzen Typ-2-Fuzzy-Mengen als Zugehörigkeitswerte Typ-1-Fuzzy-Mengen aus dem Intervall  $[0,1]$ . Ein Beispiel dafür wäre z. B.: Wenn eine Typ-2-Fuzzy-Menge „intelligent“ definiert ist, dann könnten die Zugehörigkeitsgrade, die den Elementen zugewiesen sind, Typ-1-Fuzzy-Mengen wie „unterdurchschnittlich“, „durchschnittlich“, „überdurchschnittlich“, „außergewöhnlich“, „genial“ usw. sein. Auf diese Weise wird der Unterschied zwischen

<sup>826</sup> Vgl. Biewer (1997), S. 72.

<sup>827</sup> In der Mathematik sind Ordnungsrelationen Verallgemeinerungen der "kleiner-gleich"-Beziehung. Sie ermöglicht es, dass Elemente einer Menge miteinander verglichen bzw. sortiert werden können. Eine Ordnungsrelation ist formal eine zweistellige Relation:  $R \subseteq (X, X)$  auf einer Menge  $X$  mit bestimmten Eigenschaften.

Ist eine Menge  $M$  mit einer Ordnungsrelation  $R$  gegeben, dann nennt man das Paar  $(M, R)$  eine *geordnete Menge*.

Meist bevorzugt man an Stelle der Schreibweise  $(a, b) \in R$  die sogenannte Infix-Notation  $a R b$ . Außerdem wird für Ordnungsrelationen in den seltensten Fällen ein Symbol wie  $R$  verwendet. Stattdessen verwendet man häufig das Symbol „ $\leq$ “.

Eine *Totalordnung*, *Anordnung* oder *lineare Ordnung* ist eine Halbordnung, die zudem eine totale Relation ist, das heißt für je zwei beliebige Elemente  $a, b$  der Grundmenge mit  $a \neq b$  ist stets mindestens eine der beiden Relationen  $a R b$  oder  $b R a$  erfüllt. Der Begriff *linear* orientiert sich an der Vorstellung, die ganze Menge in einer Sequenz  $\dots Ra_1 Ra_2 Ra_3 Ra_4 \dots$  aufzuzählen. Eine totale Ordnung auf einer abzählbaren Menge liefert eine Anordnung der Elemente in einer bestimmten Reihenfolge, zum Beispiel die Anordnung der Buchstaben A bis Z im lateinischen Alphabet. Die Reihenfolge der Buchstaben ist willkürlich festgelegt, und jede andere Reihenfolge wäre ebenfalls eine Ordnung, vgl. Wikipedia, Begriff: Ordnungsrelation, [Abruf: 11-09-2008].

<sup>828</sup> Eine Relation  $R \subseteq (X, X)$ , welche reflexiv, antisymmetrisch und transitiv ist, nennt man eine (Halb-) *Partialordnung*, vgl. Wikipedia, Begriff: Ordnungsrelation, [Abruf: 11-09-2008].

verschiedenen Zugehörigkeitsgraden verwischt.<sup>829</sup> Diese Annahme wird durch *Typ-m-Fuzzy-Mengen* verallgemeinert, bei denen  $m > 1$  ist. Je tiefer die Verschachtelung dieser Mengen ist, desto schwieriger wird es, diese Mengen zu messen oder zu visualisieren und somit auch auf deren Grundlage Berechnungen durchzuführen.<sup>830</sup>

Um die Unsicherheit und Variabilität zu erfassen, die bei der Festlegung von Zugehörigkeitsfunktionen auftreten können, kann man die Forderung nach der Präzision einer Zugehörigkeitsfunktion lockern, indem man Zugehörigkeitsgrade durch Teilintervalle statt einzelne Nummern des Einheitsintervalls  $[0,1]$  beschreibt.<sup>831</sup> Fuzzy-Mengen dieser Art werden dann *intervallwertige Fuzzy-Mengen* genannt. Das Konzept der intervallwertigen Fuzzy-Mengen können zu L-Fuzzy-Mengen erweitert werden, indem man  $[0,1]$  durch eine Menge partieller Ordnung  $L$  ersetzt und somit voraussetzt, dass für jedes Element  $x \in X$  die unscharfe Zugehörigkeit des Elementes  $\mu_A(x)$  einem Segment von Elementen totaler Ordnung in  $L$  entspricht.<sup>832</sup>

Trotz Verallgemeinerung und Erweiterung<sup>833</sup> ist eine bestimmte Anordnung der Elemente in einer Fuzzy-Menge unumgänglich. Sie stellt eine Mindestanforderung für die Anwendung der Fuzzy-Set-Theorie dar.

#### 6.4.2.1 Festlegung der Zugehörigkeitsfunktion

Um die Fuzzy-Set-Theorie anwenden zu können, sollen folgende Schritte gesetzt werden:<sup>834</sup>

1. Präzise Spezifizierung der Domäne  $X$  .
  - Was ist das Grunduniversum der Objekte unter Beobachtung?
2. Den Objekten  $x$  aus dem Grunduniversum  $X$  sind Zugehörigkeitsgrade  $\mu(x)$  in unscharfen Mengen zuzuordnen.
  - Welche Bedeutungen/Eigenschaften repräsentieren die unscharfen Mengen?
  - Was soll der Zugehörigkeitsgrad repräsentieren?

<sup>829</sup> Vgl. Klir/Folger (1988), S. 13.

<sup>830</sup> Vgl. Zimmermann (1991), S. 24 und Klir/Folger (1988), S. 13.

<sup>831</sup> Vgl. Biewer (1997), S. 73 und Klir/Folger (1988), S. 14.

<sup>832</sup> Vgl. Klir/Folger (1988), S. 14.

<sup>833</sup> Außer den oben genannten Typen zur Verallgemeinerung und Erweiterung existieren noch andere Typen von Fuzzy-Mengen. Siehe dazu: Zimmermann (1991), S. 25ff.

<sup>834</sup> Vgl. Smithson (2006), S. 18.

Eine sorgfältige und exakte Konzeptualisierung der verwendeten Mengen ist essentiell. Dazu muss zunächst eine Quantifizierung des Datenmaterials vorgenommen werden und eine einheitliche Skala festgelegt werden, um die Daten vergleichbar zu machen.<sup>835</sup> Der zweite Schritt erfordert, dass die unscharfe Menge eine Bedeutung aufweist, die einen bestimmten Zugehörigkeitsgrad eines Objektes begründet.

Laut Smithson ist eine Zugehörigkeitsfunktion „ [...] a function for an attribute A over some space of objects (which may or may not to be numerical)  $\Xi$  mapping to the unit interval  $[0, 1]$  [...]“.<sup>836</sup> Als ein Index für die Mengenzugehörigkeit zeigt diese Funktion den Grad an, zu dem ein Objekt  $x$  mit der Eigenschaft A ein Element einer bestimmten Menge ist.<sup>837</sup> Jedem Objekt  $x$  wird zunächst ein Zugehörigkeitsgrad zugeordnet, und somit wird nur eine einzige Dimension berücksichtigt. Die Untersuchung mehrerer Eigenschaften dieses Objekts erfordert dann genauso viele Zugehörigkeitsfunktionen wie Dimensionen.

Generell ist der Zugehörigkeitsgrad latent, das heißt, er lässt sich nicht direkt messen, und er beinhaltet eine stark vom Kontext abhängige Interpretation. Diesbezüglich muss der Kontext so genau wie möglich spezifiziert werden. Zum Beispiel hat die Bezeichnung „groß“ für die Klassifizierung der Steine im Kontext von Juweliergeschäft eine andere Bedeutung als im Kontext von Bergbau. Dieser Aspekt muss bei der Festlegung der Zugehörigkeitsfunktion bzw. bei der Konzeptualisierung der unscharfen Mengen berücksichtigt werden.

Für die Festlegung der Zugehörigkeitsfunktion finden sich in der Literatur verschiedene Konzepte und Herangehensweisen.<sup>838</sup> Die Zugehörigkeitsfunktion kann subjektiv (intuitiv, d. h. über subjektive Überzeugung) oder empirisch (über wissenschaftliche Studien) festgelegt werden.<sup>839</sup> Häufig werden Trapezfunktionen, Dreiecke oder Glockenkurven zur Beschreibung von unscharfen Mengen verwendet. Die Wahl solcher Funktionsformen beschleunigt und vereinfacht das Verfahren, mit mehr oder weniger großen Einbußen. Geringfügige Variationen des Funktionsverlaufs haben normalerweise keine signifikanten Auswirkungen, weil die Operationen über Fuzzy-Mengen im Allgemeinen sehr robust sind.<sup>840</sup> Die Trapezfunktion gibt die Aussage „etwa zwischen  $x$  und  $y$ “ wieder, während die Dreiecksfunktion und Glockenkurve den

---

<sup>835</sup> Vgl. Smithson (2006), S. 18.

<sup>836</sup> Vgl. ebd., S. 19.

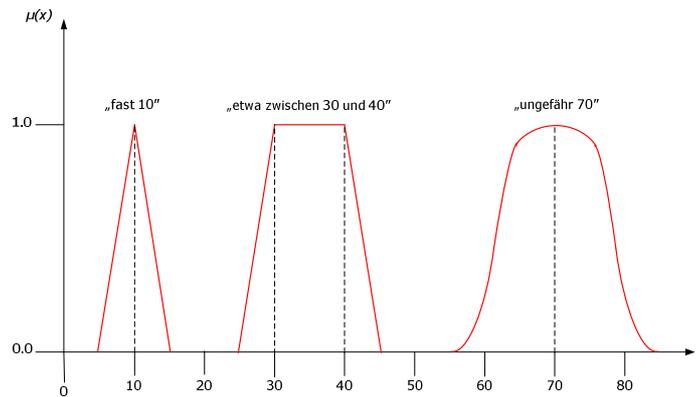
<sup>837</sup> Vgl. ebd., S. 7.

<sup>838</sup> Darstellungen verschiedener Methoden zur Festlegung und Validierung von Zugehörigkeitsfunktionen bieten Dubois/Prade (1980), S. 256ff. und (1988), S. 19ff., Zimmermann (1991), S. 344 ff.

<sup>839</sup> Vgl. Pofahl/Viertel (2006), S. 31 und Biewer (1997), S. 60f.

<sup>840</sup> Vgl. Biewer (1997), S. 61.

Aussagen „fast  $x$ “ und „ungefähr  $x$ “ nahe kommen.



**Abbildung 51: Beispiele für Zugehörigkeitsfunktionen**

Quelle: eigene Darstellung, in Anlehnung an TU Graz, Fuzzy-Logik, [Abfrage: 10-09-2008]

Smithson stellt vier Ansätze zur Festlegung der Zugehörigkeitsfunktion vor:<sup>841</sup> Je nach Forschungsziel kommt dann entweder die *formalistische Interpretation* (Die Festlegung der Zugehörigkeitsfunktion erfolgt ausschließlich mathematisch, indem bestimmte Stützvariablen der Zugehörigkeitsskala bzw. -funktion zugrunde gelegt werden), die *probabilistische Interpretation* (Die Festlegung der Zugehörigkeitsfunktion basiert auf der Wahrscheinlichkeitstheorie), die *entscheidungstheoretische Interpretation* (Die Festlegung der Zugehörigkeitsfunktion basiert auf dem Nutzen der Feststellung, dass  $x$  zu  $A$  gehört und bezieht sich auf den Wahrheitsgrad dieser Feststellung) oder die *axiomatische Messtheorie* (Die Festlegung der Zugehörigkeitsfunktion basiert auf der Annahme, dass die numerische Zugehörigkeitsfestlegung quantitativ ist in dem Sinne, dass sie sich wie partielle Anteile verhalten. Die quantitative Struktur der Zugehörigkeit kann in einer Menge von qualitativen axiomatischen Bedingungen zusammengefasst werden, die empirisch dargestellt werden können) zur Anwendung.

Diese Verschiedenheit von Methoden und Konzepten weist darauf hin, dass zur Bestimmung der Zugehörigkeitsfunktion keine einheitlichen Regeln gelten. Welche Methode angewendet wird, hängt von der subjektiven Entscheidung des jeweiligen Betrachters ab. Es ist sinnvoll, dass der Betrachter diejenige Methode auswählt, welche ihm angemessen erscheint. So existieren in der Fuzzy-Theorie nur methodische Richtlinien, deren Auslegung mehr oder weniger der Interpretation des Anwenders dieser Theorie überlassen bleibt.

Die Anforderungen, die für eine erfolgreiche Anwendung des Konzepts unscharfer Mengen vonnöten sind, können wie folgt zusammengefasst werden:

<sup>841</sup> Vgl. Smithson (2006), S. 21ff.

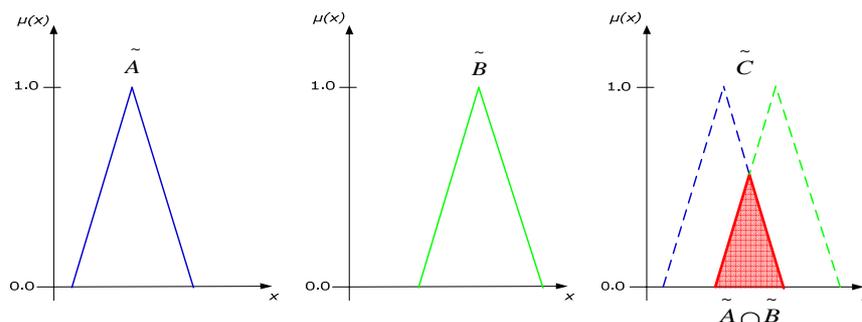
- Definition eines Grunduniversums  $X$ .
- Definition der Elemente  $x$  aus dem Grunduniversum  $X$ .
- Definition der unscharfen Menge bzw. Mengen (durch ihre Eigenschaften bzw. Bedeutungen).
- Bestimmung der Grenzen der unscharfen Menge bzw. Mengen (Ober- und Untergrenze).
- Jedem Element  $x$  wird ein Zugehörigkeitsgrad in den unscharfen Mengen zugeordnet und damit die Form der Zugehörigkeitsfunktionen festgelegt.

Ist das Konzept unscharfer Mengen anwendbar, so lassen sich, analog zur klassischen Mengenlehre, Operationen mit den definierten Fuzzy-Mengen durchführen. Die wichtigsten mengentheoretischen Operationen werden im Folgenden kurz vorgestellt.

#### 6.4.2.2 Operationen mit Fuzzy-Sets

Die Operationen mit unscharfen Mengen werden über deren Zugehörigkeitsfunktionen definiert.<sup>842</sup>

Im Folgenden werden die wichtigsten Operationen (Vereinigung, Durchschnitt, Komplement) vorgestellt. Diese liefern eine gute Basis für das Arbeiten mit Fuzzy-Mengen, stellen jedoch nur einen kleinen Ausblick auf die Erweiterung der klassischen Mengenlehre dar.<sup>843</sup>



**Abbildung 52: Durchschnitt zweier Fuzzy-Mengen**

Quelle: eigene Darstellung, in Anlehnung an TU Graz, Fuzzy-Logik, [Abfrage: 10-10-2008]

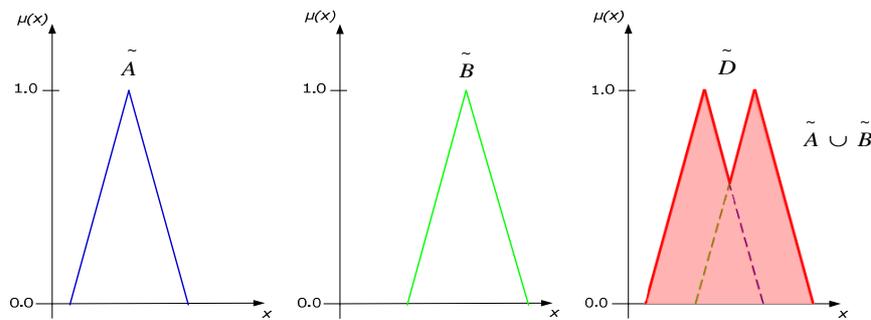
Die Zugehörigkeitsfunktion  $\mu_{\tilde{C}}(x)$  der *Schnittmenge* (oder Durchschnittsmenge, UND-

<sup>842</sup> Vgl. Zimmermann (1991), S. 16.

<sup>843</sup> Es existieren sowohl alternative als auch zusätzliche Definitionen, siehe: Zimmermann (1991), S. 17.

Operator)  $\tilde{C} = \tilde{A} \cap \tilde{B}$  ist definiert durch:  $\mu_{\tilde{C}}(x) = \min \{ \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x) \}, \quad x \in X$ .

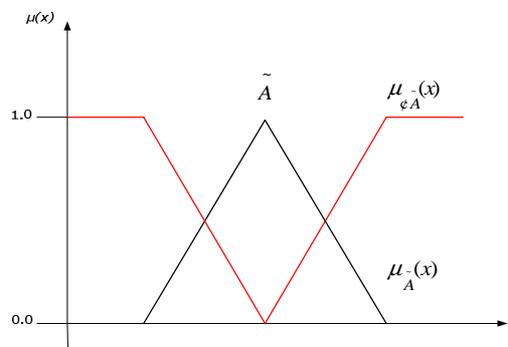
Die Zugehörigkeitsfunktion  $\mu_{\tilde{D}}(x)$  der Vereinigung (ODER-Operator)  $\tilde{D} = \tilde{A} \cup \tilde{B}$  ist definiert durch:  $\mu_{\tilde{D}}(x) = \max \{ \mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x) \}, \quad x \in X$ .



**Abbildung 53: Vereinigung zweier Fuzzy-Mengen**

Quelle: eigene Darstellung, in Anlehnung an TU Graz, Fuzzy-Logik, [Abfrage: 10-10-2008]

Die Zugehörigkeitsfunktion des Komplements (NICHT-Operator) einer unscharfen Menge  $\tilde{A}$ ,  $\mu_{\tilde{A}^c}(x)$  ist definiert durch:  $\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x), \quad x \in X$ .



**Abbildung 54: Komplement einer Fuzzy-Menge**

Quelle: eigene Darstellung, in Anlehnung an TU Graz, Fuzzy-Logik, [Abfrage: 10-09-2008]

Diese grundlegenden mengentheoretischen Operationen sind namentlich bereits aus der klassischen Mengenlehre bekannt. Allerdings gibt es auch Operationen, die lediglich in der Fuzzy-Set-Theorie durchgeführt werden können, da sie von Zugehörigkeitsgraden zwischen 0 und 1 abhängig sind. Das sind z. B. die Operationen Konzentration, Dilatation und Aggregation. Konzentration (Verstärkung, Verdichtung) und Dilatation (Aufweichung, Ausdehnung) modifizieren die Form einer unscharfen

Menge und verändern damit ihre Eigenschaft bzw. Bedeutung. Die Aggregation<sup>844</sup> stellt eine Verbindung von Mengen über Operatoren dar.<sup>845</sup>

### 6.4.3 Fuzzy-Logik als Erweiterung der klassischen Logik

In diesem Kapitel werden zunächst die Grundlagen klassischer Logik kurz eingeführt. Darauf aufbauend wird die unscharfe Logik erklärt, um im nachfolgenden Kapitel auf die unscharfe Schlussfolgerung einzugehen.

Die klassische Logik bzw. klassische logische Systeme müssen zwei semantische Bedingungen erfüllen<sup>846</sup>:

1. Das Prinzip der Zweiwertigkeit: Jede Aussage besitzt genau einen von zwei Wahrheitswerten „ja“ und „nein“, „1“ und „0“ oder „wahr“ und „falsch“.<sup>847</sup>
2. Das Prinzip der Extensionalität: Der Wahrheitswert einer Aussage ist eindeutig durch die Wahrheitswerte ihrer Teilaussagen und die Art, wie diese zusammengesetzt sind, bestimmt.

Als wichtige Teilbereiche der klassischen Logik gelten die Aussagenlogik, die Prädikatenlogik der ersten Stufe und die Prädikatenlogik höherer Stufe.<sup>848</sup> Die Aussagenlogik untersucht, ob eine Aussage in weiteren Aussagen zerlegbar ist. Lässt sich eine Aussage nicht in weitere, durch Junktoren verbundene Teilaussagen zerlegen, ist sie „atomar“ bzw. „elementar“. In der Prädikatenlogik lässt sich die innere Struktur von atomaren Aussagen darstellen. Dabei wird die innere Struktur einer Aussage durch Prädikate (Aussagefunktion) und deren Argumente (Subjekt) beschrieben. Das Prädikat drückt dabei z. B. eine Eigenschaft aus, die auf das

---

<sup>844</sup> Die Aggregation ermittelt den Wahrheitswert der gesamten Bedingung einer Fuzzy-Regel. Falls eine Bedingung aus einer Kombination (UND-, ODER-Verknüpfung) mehrerer Inputs besteht wie in der Form: „Wenn Regel A1 und/oder Regel A2 dann ...“, so werden die Wahrheitswerte der Teilbedingungen bei der Aggregation so kombiniert bzw. verknüpft, dass man genau einen Wahrheitswert aus dem Intervall [0,1] für die gesamte Bedingung erhält, vgl. Zimmermann (1995), S. 27 und Schulte (1993), S. 25f und 76.

<sup>845</sup> Vgl. Smithson (2006), S. 12.

<sup>846</sup> Vgl. Wikipedia, Begriff: Logik, [Abruf: 15-09-2008].

<sup>847</sup> Das Prinzip der Zweiwertigkeit ist vom Satz vom ausgeschlossenen Dritten zu unterscheiden: In der modernen formalen Logik wird der Satz vom ausgeschlossenen Dritten auf eine Aussage und deren Satzverneinung bezogen. Er besagt, dass sich innerhalb eines logischen Systems für eine beliebige Aussage P die Aussage  $P \vee \neg P$  (P oder nicht P) herleiten lässt.

Z. B.: „Es regnet, oder es ist nicht der Fall, dass es regnet“ stellt einen Satz der klassischen Aussagenlogik dar, kann also syntaktisch aus den Regeln und Axiomen des logischen Systems hergeleitet werden, ohne dass der Wahrheitsbegriff explizit eine Rolle spielt. Demgegenüber ist das Prinzip der Zweiwertigkeit die Eigenschaft einer Logik, dass jede Aussage P entweder „wahr“ oder „falsch“ ist. Dies bedeutet, dass semantisch jeder Aussage P genau einer von diesen zwei Wahrheitswerten zugewiesen wird, vgl. Wikipedia, Begriff: Klassische Logik, [Abruf: 15-09-2008].

<sup>848</sup> Vgl. Wikipedia, Begriff: Logik, [Abruf: 15-09-2008].

Argument zutrifft oder eine Relation, die zwischen zwei Argumenten besteht.

Eine einfache Aussage kann generell wie folgt ausgedrückt werden: „x ist P“, wobei das Symbol x für ein Subjekt aus dem Grunduniversum  $X$  ist, und P steht für das Prädikat und charakterisiert eine Eigenschaft. P kann als Funktion definiert über  $X$  aufgefasst werden, die für jeden Wert x eine Aussage bildet. Diese Funktion wird dann als  $P(x)$  bezeichnet und „Prädikat“ genannt. Wenn ein bestimmtes x aus dem Grunduniversum  $X$  in die Funktion eingesetzt wird, entsteht eine Aussage, die entweder „wahr“ oder „falsch“ ist.<sup>849</sup>

In der Logik unterscheidet man zwischen drei Elementen:<sup>850</sup> Wahrheitswerte, Vokabular (logische Operatoren) und Argumentationsverfahren (Tautologien<sup>851</sup>, Syllogismen<sup>852</sup>).

Die Wahrheitswerte können in der klassischen Logik die Werte 1 und 0 annehmen.

Für die Verknüpfung von Aussagen werden logische Operatoren verwendet. Diese sind den mengentheoretischen Operatoren ähnlich, wie z. B. Konjunktion ( $\wedge$ , entspricht dem mengentheoretischen Durchschnitt), Disjunktion ( $\vee$ , entspricht der mengentheoretischen Vereinigung), Negation ( $\neg$ , entspricht dem mengentheoretischen Komplement).

Das dritte Element der Logik ist das Schlussfolgern, welches meist auf Tautologien wie Modus Ponens, Modus Tollens, Syllogismus und Kontraposition beruht.<sup>853</sup> Diese Tautologien werden für die Ableitung deduktiver Schlussfolgerungen (Inferenzen) benutzt. Die verschiedenen Formen von Tautologien werden auch als „Inferenzregeln“ bezeichnet<sup>854</sup>:

- Modus Ponens:  $(A \wedge (A \Rightarrow B)) \Rightarrow B$
- Modus Tollens:  $((A \Rightarrow B) \wedge \neg B) \Rightarrow \neg A$
- Syllogismus:  $((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$

---

<sup>849</sup> Vg. Klir/Folger (1988), S. 26.

<sup>850</sup> Vgl. Zimmermann (1991), S. 139.

<sup>851</sup> Eine *Tautologie* ist eine Aussageform, die unabhängig von den Wahrheitswerten ihrer Bestandteile stets wahr ist.

<sup>852</sup> Logische Schlüsse. Ein *Syllogismus* besteht aus zwei Prämissen (Voraussetzungen), die Obersatz und Untersatz genannt werden und gemeinsam zu einer Konklusion (Schlussfolgerung) führen.

<sup>853</sup> Vgl. Zimmermann (1991), S. 140.

<sup>854</sup> Vgl. Klir/Folger (1988), S. 25.

- Kontraposition:  $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$

Der Modus Ponens besagt („Wenn A wahr ist und (  $\wedge$  ) (Wenn die Aussage: „Wenn A wahr ist, dann ist B wahr“) auch wahr ist) dann ist B wahr“. In der Logik wird also mittels aufgestellter Regeln eine Schlussfolgerung (Inferenz) gewonnen.<sup>855</sup>

Die klassische Mengenlehre und die Logik sind zueinander isomorph<sup>856</sup>. Das bedeutet, dass die beiden mathematischen Systeme an und für sich die gleichen Strukturen aufweisen, abgesehen von der Darstellung der Elemente und der Benennung von Relationen und Operationen. Dieser Isomorphismus erlaubt mit der Darstellung lediglich einer der beiden Theorien gleichzeitig, die beiden zu erfassen, und erstreckt sich auf die unscharfe Mengenlehre und damit auch auf die unscharfe Logik.<sup>857</sup> Daher wird in der Fuzzy-Theorie die erneute Darstellung logischer Operationen unterlassen.

Die unscharfe Logik kann als mehrwertige Logik angesehen werden, da diese bereits mehr als zwei Wahrheitswerte zulässt.<sup>858</sup> Die mehrwertige Logik, die in 1930 von Lukasiewicz eingeführt wurde, entwickelte sich aus der klassischen Logik durch die Aufhebung des Prinzips der Zweiwertigkeit. Die Standard Lukasiewicz Logik<sup>859</sup>  $L_1$  ist isomorph zur Fuzzy-Set-Theorie, die auf den Operatoren  $\min(a, b)$ ,  $\max(a, b)$  und  $1-a$  für fuzzy Durchschnitt, Vereinigung und Komplement basiert. Die Fuzzy-Logik ist eine Erweiterung der mehrwertigen Logik, da sie unendlich viele Wahrheitswerte aus dem Intervall  $[0,1]$  zulässt.<sup>860</sup>

Eine Aussage kann in der Fuzzy-Logik zu einem bestimmten Grad wahr bzw. falsch sein. Ein einfaches Prädikat, wie „x ist P“, wird in der Fuzzy-Logik durch eine Fuzzy-Menge repräsentiert. Angenommen, x steht für das Subjekt „Person Y“ und P hat die Bedeutung „jung“. Weiters werden als Grunduniversum die ganzen Zahlen von 0 bis

<sup>855</sup> Die Gültigkeit von Aussagen und Schlussfolgerungen muss sowohl semantisch als auch syntaktisch gewährleistet sein.

<sup>856</sup> Ebenso ist die Boolesche Algebra isomorph zur klassischen Mengenlehre und Logik. *Isomorphismus* ist eine Abbildung zwischen zwei mathematischen Strukturen, durch die Teile einer Struktur auf „bedeutungsgleiche“ Teile einer anderen Struktur umkehrbar eindeutig (bijektiv) abgebildet werden. *Isomorphismus* liegt also dann vor, wenn sich die Elemente zweier Bereiche (Strukturen), in denen je eine Verknüpfung erklärt ist, umkehrbar eindeutig einander so zuordnen lassen, dass die Verknüpfung zweier Elemente im ersten Bereich die entsprechende Verknüpfung der zugeordneten Elemente im zweiten Bereich zur Folge hat. So ist z. B. der Bereich der positiven reellen Zahlen  $\mathbb{R}^+$  mit der Multiplikation als Verknüpfung isomorph zu den reellen Zahlen  $\mathbb{R}$  mit der Addition, da die (bijektive) Logarithmusfunktion  $f(x) = \log x$  von  $\mathbb{R}^+$  nach  $\mathbb{R}$  einen Isomorphismus darstellt: es gilt nämlich  $\log(x \cdot y) = \log x + \log y$ , vgl. Wissen.de und Wikipedia, Begriff: Isomorphismus, [Abruf: 15-09-2008].

<sup>857</sup> Vgl. Klir/Folger (1988), S. 25.

<sup>858</sup> Vgl. Kruse/Gebhardt/Klawonn (1993), S. 58.

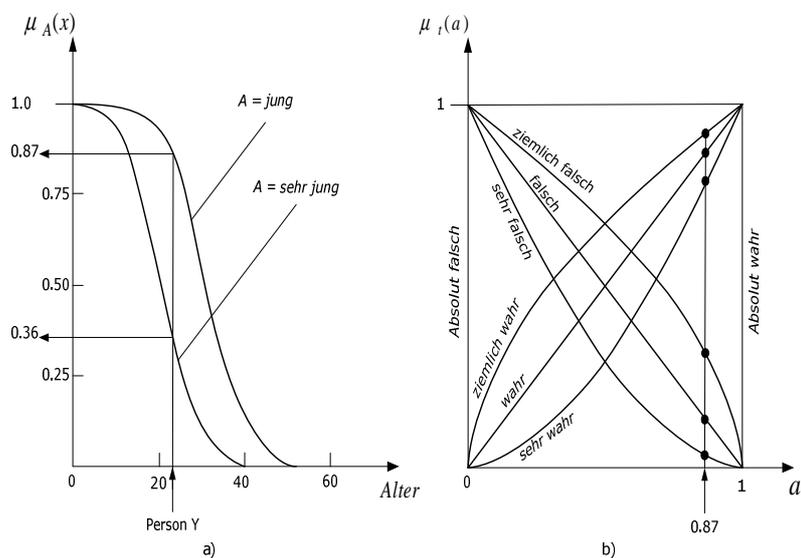
<sup>859</sup> Die Standard Lukasiewicz Logik  $L_1$  (auch *infinite-valued logic* genannt) deutet auf die Logik hin, deren Wahrheitswerte durch allen reellen Zahlen im Intervall  $[0,1]$  repräsentiert werden, vgl. Klir/Folger (1988), S. 29.

<sup>860</sup> Vgl. Zimmermann (1991), S. 143.

60 genommen, die das Alter repräsentieren. Einerseits hängt der Wahrheitswert der Aussage „Person Y ist jung“ vom Zugehörigkeitsgrad des Alters dieser Person in der unscharfen Menge „junge Menschen“ ab (siehe Abb. 55 a), andererseits spielt auch die Stärke der Wahrheit bzw. Falschheit eine Rolle (siehe Abb. 55 b):<sup>861</sup>

- Person Y ist jung ist wahr
- Person Y ist jung ist falsch
- Person Y ist jung ist ziemlich wahr
- Person Y ist jung ist sehr falsch.

Das kann man mit folgender Graphik veranschaulichen:



**Abbildung 55: Wahrheitswerte einer Fuzzy-Aussage**

Quelle: in Anlehnung an Klir/Folger (1988), S. 31.

Laut Zimmermann ist die Fuzzy-Logik „[...] an extension of set theoretic multivalued logic in which the truth values are linguistic variables (or terms of the linguistic variable truth).“<sup>862</sup>

Die Fuzzy-Logik basiert auf der unscharfen Mengenlehre und zielt auf das Approximate Reasoning also die unscharfe Schlussfolgerung mit unpräzisen Aussagen ab.<sup>863</sup>

<sup>861</sup> Vgl. Klir/Folger (1988), S. 31.

<sup>862</sup> Zimmermann (1991), S. 141.

<sup>863</sup> Vgl. Klir/Folger (1988), S. 30.

### 6.4.3.1 Unscharfe Schlussfolgerung (Approximatives Schließen)

Die Regeln können auf Basis von Erfahrung und Wissen in Form natürlicher Sprache formuliert werden. Dafür sind keine Systemmodellierungen oder komplexe mathematische Gleichungssysteme notwendig. Es ist meist möglich, ein System mit relativ wenigen Fuzzy-Regeln zu definieren.<sup>864</sup>

Die allgemeine Form einer (Fuzzy-)Regel sieht folgendermaßen aus:<sup>865</sup>

*Wenn (Bedingung) dann (Schlussfolgerung).*

Z. B.:

*Wenn (es kalt ist) dann (drehe das Ventil der Heizung auf).*

Die Argumente sind als Zugehörigkeitswerte zu unscharfen Mengen definiert, deshalb kann die Regel in die folgende abstrakte Form umgewandelt werden:

*Wenn ( $x \in A$ ) dann ( $y \in B$ ).*

Darin ist  $x$  der Momentanwert einer Basisvariablen (z. B. Temperatur) und  $y$  die dazugehörige Stellgröße.  $A$  und  $B$  sind unscharfe Mengen und  $x \in A$  bzw.  $y \in B$  beschreibt über die entsprechende Funktion die Zugehörigkeit der Basisvariablen  $x$  bzw.  $y$  zur unscharfen Menge  $A$  bzw.  $B$ .

Außerdem stellen  $x$  und  $y$  sogenannte linguistische Variablen dar,  $A$  und  $B$  sind linguistische Terme, welche mittels Fuzzy-Mengen dargestellt werden können.<sup>866</sup> Der erste Teil der Regel „*Wenn  $x$  ist  $A$* “ ist eine Prämisse (Bedingung), der zweite Teil „*dann  $y$  ist  $B$* “ stellt die Konklusion (Schlussfolgerung) dar.<sup>867</sup> Die Bedingung kann mit Hilfe von logischen Operatoren wie z. B. „und“, „oder“ mit anderen Bedingungen verbunden werden<sup>868</sup>: „Wenn Gesamtumsatz (getätigte Käufe eines Kunden) ist hoch oder Kaufhäufigkeit (Kaufrequenz dieses Kunden) ist sehr hoch und Anzahl der angekauften Produkte ist befriedigend, dann ist dieser Kunde ein guter Kunde (also als gut zu klassifizieren).“

Solche Regeln können durch Expertenschätzungen aufgestellt werden. Diese Fuzzy-

---

<sup>864</sup> Vgl. Kosko (1993), S. 194.

<sup>865</sup> Vgl. Schulte (1993), S. 74.

<sup>866</sup> Vgl. Zimmermann (1991), S. 146.

<sup>867</sup> Vgl. Zimmermann (1991), S. 146 und Schulte (1993), S. 74.

<sup>868</sup> Vgl. Schulte (1993), S. 77.

(Wenn-dann)Regeln können mittels logischen Operatoren viele verschiedene unscharfe Bedingungen (Aussagen) miteinander verknüpfen. Dadurch stellen sie den semantischen Aspekt des Schlussfolgerns dar. Die mathematische Verarbeitung von Fuzzy-Regeln wird mittels Inferenzmechanismus durchgeführt. Unter Inferenz versteht man die Auswertung der Regeln aus der Regelbasis und die anschließende Zusammenfassung der daraus abzuleitenden Konklusion, basierend auf einer kontextabhängigen Entscheidungsstrategie.<sup>869</sup>

Approximatives Schließen entspricht der Art, wie Menschen ihre Schlüsse aus unpräziser Information ziehen.<sup>870</sup> Die Semantik ist bei der Repräsentation natürlicher Sprache mittels Fuzzy-Logik besonders wichtig, vor allem, wenn sie zur Repräsentation menschlichen Wissens herangezogen wird. Die Bedeutung wird durch die Manipulation (Modifikation) von Fuzzy-Mengen operationalisiert.<sup>871</sup> Dazu werden linguistische Terme (fuzzy Prädikate wie z. B. teuer, alt, groß, klein, hoch,...), fuzzy Quantifikatoren (viele, wenige, fast alle,...), fuzzy Modifikatoren (sehr, ziemlich, wahrscheinlich, unwahrscheinlich, fast unmöglich,...) und fuzzy Wahrheitswerte (ziemlich wahr, wahr, sehr wahr, ziemlich falsch,...) herangezogen.<sup>872</sup>

#### **6.4.3.2 Linguistische Variablen und Terme**

Die linguistische Variable ist das Grundwerkzeug für Fuzzy-Logik und Approximate Reasoning.<sup>873</sup>

„In retrating from precision in the face of overpowering complexity, it is natural to explore the use of what might be called linguistic variables, that is, variables whose values are not numbers but words or sentences in a natural or artificial language. The motivation for the use of words or sentences rather than numbers is that linguistic characterizations are, in general, less specific than numerical ones.“<sup>874</sup>

Die linguistischen Variablen stellen also eine spezielle Form von unscharfen Mengen dar. Die Werte bzw. Ausprägungen sind im Gegensatz zu numerischen Variablen nicht Zahlen, sondern Wörter und Ausdrücke. Sie dienen dazu, sprachlich ausgedrücktes Wissen und dessen Bedeutung mit seinen Unschärfen zu modellieren, so dass dabei möglichst wenig vom Reichtum menschlicher Sprache verlorengeht.

---

<sup>869</sup> Vgl. Schildt/Kastner (1998), S. 203. Zu wichtigsten Strategien zählen z. B. die MAX-MIN-Inferenz und die MAX-PROD-Inferenz, siehe dazu: Schulte (1997), S. 76- 81 und Schildt/Kastner (1998), S. 203-207.

<sup>870</sup> Vgl. Klir/Folger (1988), S. 30.

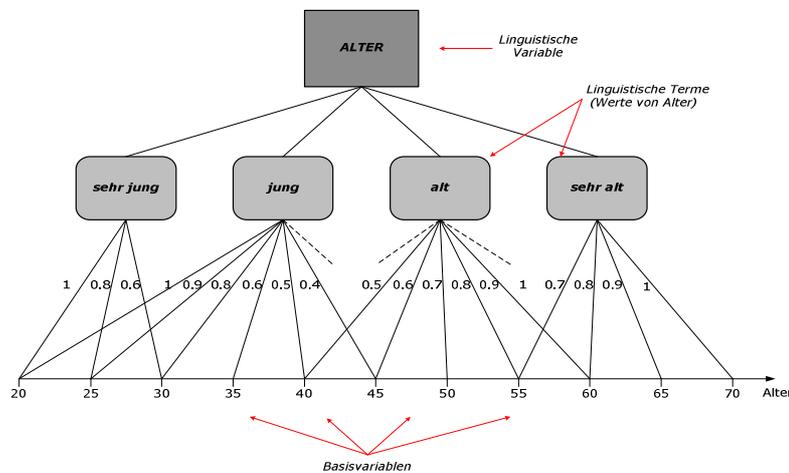
<sup>871</sup> Vgl. Klir/Folger (1998), S. 32.

<sup>872</sup> Vgl. Klir/Folger (1998), S. 30ff.

<sup>873</sup> Vgl. Zimmermann (1991), S. 131.

<sup>874</sup> Zadeh (1973), S. 3, zit. in: Zimmermann (1991), S. 131.

Das Konzept der linguistischen Variablen erlaubt es, komplexe Systeme zu beschreiben, die mathematisch nur schwer darstellbar sind. Dabei wird es zwischen fuzzy Variablen und fuzzy Termen unterschieden.



**Abbildung 56: Die linguistische Variable „Alter“**

Quelle: eigene Darstellung, in Anlehnung an Zimmermann (1991), S. 133.

Eine linguistische Variable  $X$  wird durch ein Quintupel der Form  $(x, T(x), U, G, \tilde{M})$  charakterisiert<sup>875</sup>: Sei  $X$  eine linguistische Variable mit dem Namen (oder Label)  $x$ , z. B. „Alter“, die verschiedene Ausprägungen bzw. linguistische Terme  $T(x)$  annehmen kann, wie z. B. „sehr jung“, „alt“, „sehr alt“. Die Grundgesamtheit  $U = [0,100]$  mit der Basisvariablen  $u$  stellt die möglichen Werte dar, die  $x$  annehmen kann, also die Lebensjahre von 0 bis 100.  $G$  ist eine syntaktische Regel zur Erzeugung des Namens  $X$  aus den Werten  $x$ .  $\tilde{M}(X)$  ist eine semantische Regel, die den linguistischen Termen eine Bedeutung, d. h. eine Fuzzy-Menge, zuordnet.

$$\tilde{M} = \{ (u, \mu_{alt}(u)) \mid u \in [0,100] \},$$

für die gilt,

$$\mu_{alt}(u) = \begin{cases} 0 & u \in [0, 50] \\ \left( 1 + \left( \frac{u-50}{5} \right)^{-2} \right)^{-1} & u \in [50, 100] \end{cases}$$

<sup>875</sup> Vgl. Zimmermann (1991), S. 132.

$T(x)$  definiert die Menge der Terme für die Variable  $x$  folgendermaßen:  $T(\text{Alter}) = \{\text{sehr jung, jung, alt, sehr alt}\}$ .

Somit charakterisieren die linguistischen Terme die linguistische Variable „Alter“. Meist werden pro linguistischer Variablen fünf bis zehn linguistische Terme zur Differenzierung herangezogen (siehe auch Abb. 56).

## 6.5 FUZZY-SYSTEME

Bei Fuzzy-Systemen werden unscharfe Mengen zur inhaltsdefinierten, formalen Abbildung menschlichen Wissens benutzt.<sup>876</sup> Damit wird es möglich, menschliches Erfahrungswissen auf elektronischen Datenverarbeitungsanlagen zu verarbeiten. Hierzu gehören die folgenden Funktionen:<sup>877</sup>

- *Wissensakquisition*: Das Wissen von Experten wird aus Lehrbüchern oder maschinell aus Trainingsbeispielen akquiriert.
- *Wissensrepräsentation*: Das geschieht mittels Regeln in der Wissensbasis.
- *Wissensverarbeitung*: Das geschieht mit einer Inferenzmaschine, die in der Lage ist, linguistisches Wissen inhaltserhaltend zu verarbeiten.
- *Übersetzung*: Auf der Input-Seite wird eine mögliche Übersetzung numerischer Information in linguistische Information (Fuzzifizierung) vorgenommen. Auf der Output-Seite werden bestimmte Zugehörigkeitsfunktionen entweder in Zahlen (Defuzzifizierung) oder in linguistische Ausdrücke (linguistische Approximation) übersetzt.

Die Fuzzy-Systeme basieren auf der Fuzzy-Logik, die versucht, die Art nachzuahmen, wie Menschen an Regelprobleme herangehen. Dabei wird grundsätzlich folgendermaßen vorgegangen:<sup>878</sup>

- 1) Es wird mit unscharfen Begriffen gearbeitet, analog zum Menschen, der zumeist keine Messgeräte mit exakten Messwertangaben zur Verfügung hat, um bestimmte Situationen einzuschätzen.
- 2) Die Fuzzy-Regeln sollen die ungenauen Schätzungen formulieren, die im nächsten Schritt

---

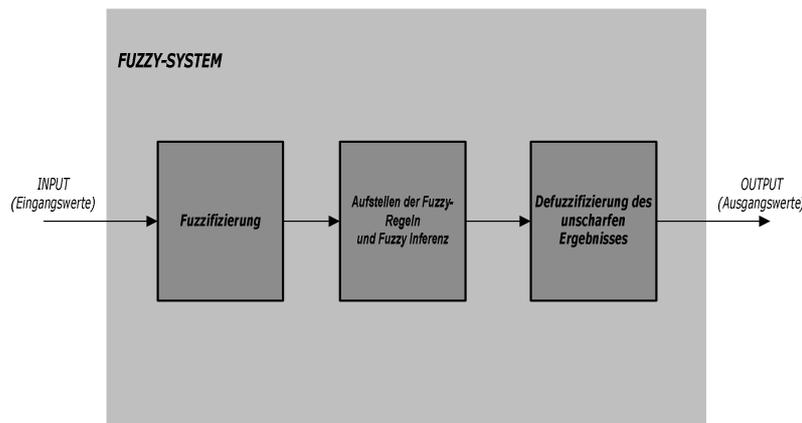
<sup>876</sup> Vgl. Zimmermann (1995), S. 24.

<sup>877</sup> Vgl. Zimmermann (1995), S. 24f.

<sup>878</sup> Vgl. Drösser (1994), S. 10.

- 3) zu einem unscharfen Ergebnis verknüpft werden. Dies nennt sich „Fuzzy-Inferenz“ (oder Schlussfolgerung) und entspricht dem Vorgang, wie Menschen bei der Lösung von alltäglichen Problemen vorgehen.
- 4) Im letzten Schritt wird meist ein scharfes Ergebnis erwartet, so dass präzise Anweisungen bzw. Handlungen vollzogen werden können. In der Fuzzy-Theorie nennt man diesen Vorgang „Defuzzifizierung“.

Die Anwendungen von Fuzzy-Logik folgen grundsätzlich einem derartigen Fuzzy-System, das folgendermaßen veranschaulicht werden kann:



**Abbildung 57: Darstellung eines Fuzzy-Systems**

Quelle: eigene Darstellung, in Anlehnung an Drösser (1994), S. 10.

Die Fuzzifizierung der Daten wurde bereits im Kapitel 6.4.2.1 Festlegung der Zugehörigkeitsfunktion erklärt. Die Erläuterungen über Approximatives Schließen und linguistische Variablen bzw. Terme können zum Fuzzy-Regelsystem herangezogen werden. Bei der Defuzzifizierung wird aus dem unscharfen Wert, der sich aus der Inferenz ergibt, ein scharfer Wert ermittelt. Dazu existieren verschiedene Methoden, die kontextspezifisch festzulegen sind. Gewöhnliche Methoden sind z. B. „maximale Höhe“, „Mittelwert der Maximalwerte“ und die „Schwerpunktmethode“.<sup>879</sup>

Ein Fuzzy-Logik-basiertes System kann jedes System modellieren und näherungsweise erfassen. Das soll durch das „Fuzzy Approximation Theorem“ (FAT) ermöglicht werden. Das FAT beruht auf der Idee, dass menschliches Wissen und Denken immer bestimmten Regeln folgt: „Wenn dies gilt, dann gilt das“. Ein Fuzzy-Logik-basierte System ist eine große Sammlung von fuzzy-logischen Wenn-Dann-Regeln.<sup>880</sup>

<sup>879</sup> Siehe dafür: Zimmermann (1995), S. 31ff., Schulte (1993), S. 84-88, Schildt/Kastner (1998), S. 207.

<sup>880</sup> Vgl. Kosko (1993), S. 192.

## 6.6 DATENANALYSE MITTELS FUZZY-LOGIK

Das Gewinnen von interessanten, d. h. validen, neuen, nützlichen und verständlichen Informationen aus großen Datenmengen ist Gegenstand und Ziel des Data Mining.

Im engeren Sinn gehört zum Data Mining nur die eigentliche Analyse der Daten, d. h. das Ermitteln von Mustern und Regeln. Zusammen mit den vorbereitenden Schritten der Bereinigung, Integration, Auswahl und Transformation der Daten sowie den nachfolgenden Schritten der Bewertung und Präsentation der gewonnenen Informationen spricht man von *Knowledge Discovery in Databases* (dem sogenannten KDD-Prozess, siehe Kapitel 5.2) bzw. *Data Mining*. Diese beiden Begriffe werden synonym verwendet.

Beim Data Mining können unterschiedliche Techniken eingesetzt werden<sup>881</sup>: Die zu untersuchenden Daten werden z. B. vorbereitet, zusammengestellt, eingeordnet, geschätzt, gruppiert, assoziiert oder visualisiert. Im Rahmen dieser Arbeit wird die Technik der Gruppierung/Segmentierung (engl.: *Clustering*) beschrieben.

Für das Clustering sind viele verschiedene Verfahren entwickelt worden. Die unterschiedlichen dabei verwendeten Ansätze und ihre Eigenschaften werden nun kurz dargestellt, danach wird auf das *Fuzzy Clustering* näher eingegangen.

Wie die vorangegangenen Kapitel deutlich gezeigt haben, besteht die besondere Stärke der Fuzzy-Logik in der Einbeziehung von Begriffen der menschlichen Sprache in die Mathematik und die daraus resultierende Zuordnung der Elemente zu unscharfen Mengen. Dies erfolgt über die Einführung von „linguistischen Variablen“. Diese werden als Gruppen von Fuzzy-Mengen mit einander mehr oder weniger überlappenden Zugehörigkeitsfunktionen bei zumeist gleicher Basisvariablen definiert. Die unscharfen Mengen einer linguistischen Variablen stehen als beschreibende Eigenschaften verschiedener Teilaspekte des übergeordneten Begriffs der linguistischen Variablen in einem begrifflichen Zusammenhang zueinander und lassen sich als „Klassen“ der linguistischen Variablen verstehen. Somit ist es z. B. möglich, bei der Erhebung von Daten nur mit linguistischen Variablen die gewünschte Information zu erfragen. Im CRM kann bspw. die linguistische Variable „Profitabilität“ oder „Loyalität“ jedes einzelnen Kunden mit unscharfen Kategorien abgefragt werden wie z. B.: „kaum“ (profitabel bzw. loyal), „wenig“, „ausreichend“, „ziemlich“, „besonders“. Vorteilhaft bei einer solchen unscharfen, dennoch leicht verständlichen Kategorisierung von Variablen mittels Fuzzy-Mengen ist, dass man auf deren Grundlage den Kunden

---

<sup>881</sup> Diese Techniken sind zu finden in: Berry/Linoff (1999).

kundenspezifische Marketingmaßnahmen einfacher und schneller zuordnen kann.

Daraus lässt sich leicht schlussfolgern, dass die klassische Mengenlehre und ihre scharfe Grenzziehung nicht adäquat zur Einteilung von Kunden in Segmente für CRM-Zwecke sind. Denn ein erfolgreiches Modell zur Kundensegmentierung in CRM sollte in der Lage sein, für jeden einzelnen Kunden im Kundenstamm angeben zu können, zu welchem Grade dieser zu einer Menge bzw. Klasse gehört. Aus marketingstrategischen Gründen ist es weitaus sinnvoller, die scharfen Grenzen der Kundenklassen in weiche bzw. gleitende Übergänge umzuwandeln, so dass in den Randbereichen geringfügige Schwankungen der Merkmalsausprägungen von Kunden sich als unterschiedliche Grade ihrer Zugehörigkeit zu einer Klasse widerspiegeln können.

### **6.6.1 Clustering**

Die Clusteranalyse ist eines der Datenanalyseverfahren, die im Data Mining eingesetzt werden.

Sie ist z. B. im Marketing von Interesse, um homogene Kundengruppen zu identifizieren und gezielt ansprechen zu können. Bei größeren Unternehmen ist hierfür der Einsatz von Clusteringverfahren unvermeidbar. Die Kunden sind nicht mehr persönlich bekannt, sondern sie sind für das Unternehmen nur durch die Informationen beschrieben, die es im Rahmen seiner Geschäftsprozesse gesammelt hat. Basierend auf diesen Informationen werden die Kundendaten durch Clusteringverfahren segmentiert. Die Kunden eines Segmentes sind hinsichtlich der betrachteten Informationen homogen und stellen daher eine einheitliche Zielgruppe dar, die entsprechend angesprochen werden kann.

Das Ziel der Clusteranalyse ist, eine Menge von Objekten in homogene Gruppen bzw. Klassen oder Cluster zu unterteilen. Dabei versteht man unter einer Einteilung in Cluster, dass

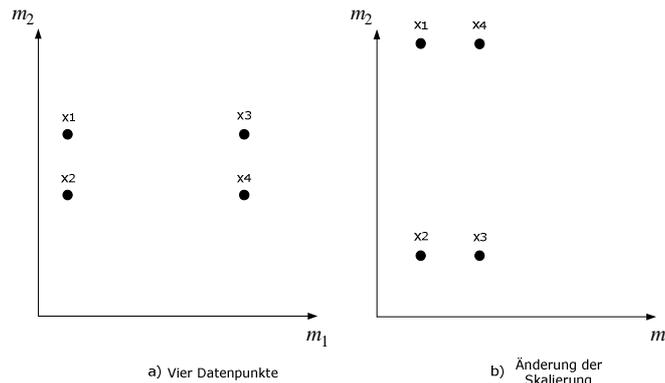
- die Objekte einer Gruppe untereinander möglichst ähnlich sind. Es wird *Homogenität innerhalb eines Clusters* gefordert.
- die Objekte verschiedener Cluster möglichst unterschiedlich sind. Es wird *Heterogenität zwischen den Clustern* gefordert.<sup>882</sup>

Die Ähnlichkeit muss hierzu in Abhängigkeit der Daten noch weiter präzisiert werden. Betrachtet man ein Datum als einen Vektor, kann als Ähnlichkeitsmaß beispielsweise

---

<sup>882</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 6.

der euklidische Abstand gewählt werden. Wichtig ist in diesem Zusammenhang, dass eine vernünftige Gewichtung<sup>883</sup> und Skalierung der Variablen (Komponenten des Vektors) verwendet wird.<sup>884</sup> Werden in einem Vektor z. B. die Höhe und die Breite als Variablen eines Objektes angegeben, wäre es normalerweise unsinnig die Höhe in Millimeteinheiten und die Breite in Metereinheiten anzugeben. Dieses Problem ist sofort zu erkennen, wenn man die Darstellung derselben Daten in unterschiedlichen Metriken vergleicht (siehe Abb. 58).



**Abbildung 58: Unterschiedliche Metriken**

Quelle: Höppner/Klawonn/Kruse (1997), S. 7.

Das Ziel der Clusteranalyse ist, die Cluster automatisch zu erkennen. Clusterverfahren lassen sich nach ihrer Vorgehensweise grundsätzlich in *hierarchische* und *partitionierende* Verfahren unterteilen.<sup>885</sup>

Partitionierende Clusterverfahren berechnen, nur eine einzige Partitionierung der zu analysierenden Daten. Sie zerteilen die Datenobjekte in eine vorgegebene Anzahl<sup>886</sup> von Gruppen. Hierarchische Verfahren erzeugen dagegen eine Hierarchie der Objekte in Form einer Baumstruktur. Sie berechnen eine Folge von Partitionierungen der Daten, indem sie sukzessiv Cluster zusammenfassen (agglomerativ) oder aufspalten (divisiv).<sup>887</sup>

Bei agglomerativen (bottom-up) Verfahren repräsentiert jedes Datenobjekt zunächst eine eigene Klasse, also in jedem Cluster befindet sich ein Objekt. Dann werden sukzessiv Klassen miteinander verschmolzen d. h. zusammengefügt, bis die gewünschten Klassen vorliegen. Divisive (top-down) Verfahren gehen umgekehrt vor: Am Anfang liegt eine einzige Klasse mit allen Objekten vor. Beginnend mit dieser

<sup>883</sup> Die einzelnen Variablen können unterschiedlich wichtig sein.

<sup>884</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 7.

<sup>885</sup> Vgl. Zimmermann (1995), S. 38f.

<sup>886</sup> d. h. die Anzahl der Cluster, in die die Daten partitioniert werden sollen.

<sup>887</sup> Vgl. Zimmermann (1995), S. 39.

Klasse werden sukzessiv neue (Teil-) Klassen erzeugt, indem bestehende Klassen in neue Klassen aufgeteilt werden, bis man eine sinnvollere Klassenbildung erhält. Nach welchen Kriterien die Cluster ausgewählt werden, die zerteilt bzw. zusammengefügt werden, unterscheiden die verschiedenen hierarchischen Clusterverfahren. Zwei der bekanntesten Verfahren heißen Single Linkage (einfache Verbindung) und Complete Linkage (Vollständige Verbindung).<sup>888</sup> Als großer Nachteil hierarchischer Clusterverfahren hat sich erwiesen, dass einmal getroffene Entscheidungen für Zerlegungen nicht mehr rückgängig gemacht werden können.<sup>889</sup>

Partitionierende Verfahren gehen von einer zufälligen Anfangspartitionierung (-verteilung) aus und ordnen die Datenobjekt zwischen den Clustern in jeder Iteration so um, dass die Güte der Gruppierung sich immer weiter verbessert. Ein Vertreter dieser Verfahrensklasse ist der Fuzzy-C-Means-Algorithmus, welcher im Laufe dieser Arbeit detailliert dargestellt wird.

Die verschiedenen Cluster-Analyseverfahren lassen sich weiter in Abhängigkeit von der Art der Zuordnung der Objekte bzw. Daten zu den Klassen in *disjunkte* und *nicht-disjunkte* Cluster-Analyseverfahren unterteilen.<sup>890</sup> Bei den disjunkten Cluster-Analyseverfahren wird jedes Datum genau einem Cluster zugeordnet. Bei den nichtdisjunkten Verfahren können Daten auch mehreren Clustern zugeordnet werden. Die Zuordnung der Daten zu den Clustern kann *deterministisch* oder *probabilistisch* erfolgen.<sup>891</sup>

Bei den deterministischen (engl.: hard clustering) Verfahren werden die Daten mit einer Wahrscheinlichkeit von „1“ einem (nicht-überlappende Zuordnung) oder mehreren (überlappende Zuordnung) Clustern zugeordnet. Bei probabilistischen (engl.: fuzzy clustering) Cluster-Analyseverfahren werden die Daten mit einer zwischen „0“ und „1“ liegenden Wahrscheinlichkeit den Clustern zugeordnet. So kann probabilistische Cluster-Analyseverfahren als eine Verallgemeinerung der deterministischen Verfahren angesehen werden. Die Verfahren der Fuzzy-Clusteranalyse können als probabilistische Cluster-Analyseverfahren bezeichnet werden, wenn die Wahrscheinlichkeiten von „0“ bis „1“ als Zugehörigkeitsgrade aufgefasst sind.<sup>892</sup> Die Zugehörigkeitsgrade repräsentieren jedoch nicht die

---

<sup>888</sup> Vgl. Zimmermann (1993), S. 74.

<sup>889</sup> Ungünstige Partitionierungen können dadurch später nicht mehr verändert oder verbessert werden.

<sup>890</sup> Vgl. Heiko (2002), S. 4

<sup>891</sup> Vgl. Heiko (2002), S. und Höppner/Klawonn/Kruse (1997), S. 8.

<sup>892</sup> Bei einer Fuzzy-Clusteranalyse werden die Objekte den Clustern mit einem Zugehörigkeitsgrad zwischen „1“ und „0“ zugeordnet, muss jedoch im Gegensatz zu allgemeinen probabilistischen Clusteranalyseverfahren die Summe der Zugehörigkeitsgrade der Objekte zu den Clustern nicht 1 ergeben.

Wahrscheinlichkeiten im wirklichen Sinne. Ein Zugehörigkeitsgrad von 0.8 besagt nicht, dass das Datum dem betreffenden Cluster mit einer Wahrscheinlichkeit von 80% zugeordnet wird. Es ist wichtig, dass die Zugehörigkeitsgrade im Sinne der Fuzzy-Logik interpretiert werden.

Wenn man die Nebenbedingung der probabilistischen Cluster-Analyseverfahren, dass die Summe der Zugehörigkeitsgrade (Wahrscheinlichkeiten) eines Datums zu den Clustern eins ergeben muss, fallen lässt, wird das probabilistische Cluster-Analyseverfahren als *possibilistisch* bezeichnet.<sup>893</sup>

Vielen Clusteranalyseverfahren liegt eine Ziel- oder Bewertungsfunktion zugrunde, die die Güte einer Partitionierung bewertet. Die Clusteranalyseverfahren versuchen, diese Funktion zu maximieren oder zu minimieren. In diesem Sinne wird die Gruppierung durch die Fuzzy-Clusteranalyse zu einem Optimierungsproblem.<sup>894</sup>

Alle möglichen Partitionen vollständig zu berechnen und anschließend die am besten bewertete Lösung auszuwählen, sind schon bei einer kleinen Anzahl von Datenobjekten aus Zeitgründen nicht durchführbar. So gibt es bspw. schon für die Zerlegung von 25 Datenobjekten in 10 Cluster ungefähr  $10^{18}$  verschiedene Möglichkeiten.<sup>895</sup> Deshalb gehen die Clusteranalyseverfahren heuristisch vor und suchen nicht nach der besten, sondern nach einer brauchbaren Lösung bei vertretbarem Aufwand.

## 6.6.2 Grundlagen der Fuzzy-Clusteranalyse

Das Ziel der Clusteranalyse ist, Daten eines nicht klassifizierten Datensatzes in Klassen bzw. Cluster einzuteilen. Dabei sollen Daten, die zu dem gleichen Cluster gehören, möglichst ähnlich und Daten, die verschiedenen Clustern zugeordnet sind, möglichst verschieden sein.

Eine eindeutige Zuordnung der Daten zu den Clustern ist jedoch bei vielen Anwendungen nicht sinnvoll. Falls die Cluster sich z. B. überlappen, wird eine eindeutige Zuordnung zu den Clustern der Struktur der Daten nicht gerecht und kann zu einem Informationsverlust führen.<sup>896</sup> Das kann anhand des sogenannten Schmetterlingbeispiels<sup>897</sup> verdeutlicht werden:

---

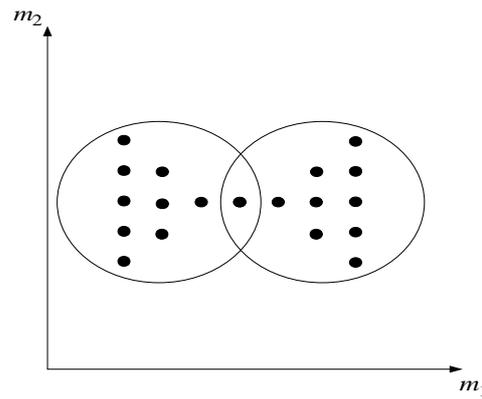
<sup>893</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 8.

<sup>894</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 9. und Zimmermann (1991), S. 232.

<sup>895</sup> Vgl. Zimmermann (1991), S. 232.

<sup>896</sup> Vgl. Heiko (2002), S. 11.

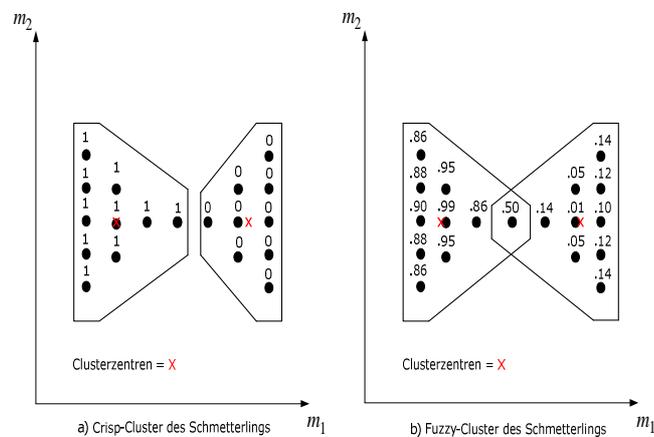
<sup>897</sup> Vgl. Zimmermann (1991), S. 225ff.



**Abbildung 59: Schmetterlingsproblem**

Quelle: eigene Darstellung

Die Abbildung 59 zeigt einen Datensatz mit sämtlichen Daten, die in zwei verschiedene Cluster eingeteilt werden sollen. Das Datum in der Mitte lässt sich aber weder dem linken noch dem rechten Cluster eindeutig zuordnen. Bei einer *crisp* Clustereinteilung (Abb. 60 a) kann man das mittlere Datum nur einem der beiden Cluster eindeutig zuordnen, obwohl der andere Cluster ebenfalls dieses in sich beinhaltet. Wenn man dennoch dieses Datum einem Cluster eindeutig zuordnet, geht die Information, dass die beiden Cluster symmetrisch sind, verloren.<sup>898</sup>



**Abbildung 60: Crisp- und Fuzzy-Cluster des Schmetterlings**

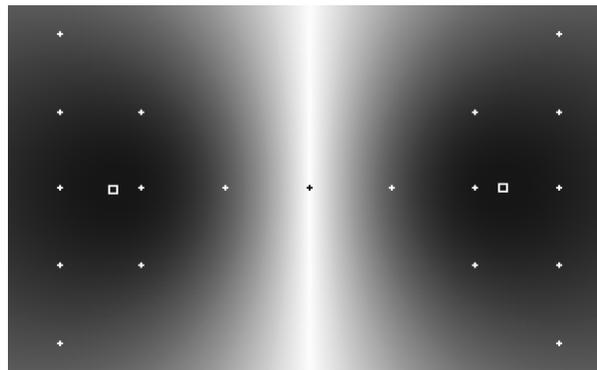
Quelle: eigene Darstellung in Anlehnung an Zimmermann (1991), S. 226f.

Ein anderer Nachteil einer eindeutigen Zuordnung ist, dass man den Daten nach einer Klassifikation nicht mehr ansehen kann, wie typisch sie für den betreffenden Cluster sind. Dagegen könnte dieses Datum bei einer *fuzzy* Clustereinteilung (Abb. 60 b) mit einem bestimmten Grad beiden Clustern angehören. Da das Datum in der Mitte beiden Clustern im gleichen Maße zugeordnet wird, ist der Zugehörigkeitsgrad des Datums zu den Clustern gleich groß.

<sup>898</sup> Vgl. Zimmermann (1991), S. 227.

Um die fließenden Übergänge zwischen den Clustern modellieren und bei der Clusteranalyse berücksichtigen zu können, werden graduelle Zugehörigkeiten verwendet. Jedem Datum  $\vec{x}_j$  wird für jedes Cluster  $\vec{\beta}_i$  ein Zugehörigkeitsgrad  $\mu_{i,j} \in [0, 1]$  zugeordnet. Ein Zugehörigkeitsgrad von 1 zeigt an, dass das Datum dem betreffenden Cluster sicher zuzuordnen ist. Ein Zugehörigkeitsgrad von 0 zeigt dagegen an, dass das Datum dem betreffenden Cluster auf keinen Fall angehört.

Die Fuzzy-Clusteranalyse basiert auf diesem Konzept der graduellen Zugehörigkeit. Die Abbildung 61 zeigt die Klassifikation des in Abbildung 59 dargestellten Datensatzes mit einem der bekanntesten Fuzzy-Clusteringverfahren, dem Fuzzy-C-Means-Verfahren. Die Quadrate symbolisieren die Zentren der Cluster und die grauen Zonen geben die Zugehörigkeit an, in Abhängigkeit von der jeweiligen Entfernung der Daten zu den Clusterzentren. Das Verfahren unterteilt den Datensatz in zwei symmetrische Cluster. Das Datum in der Mitte weist zu beiden Clustern den gleichen Zugehörigkeitsgrad auf. Die auf diese Weise erzeugte Klassifikation entspricht genau der Klassifizierung eines Menschen.



**Abbildung 61: Klassifikation mit dem Fuzzy-C-Means Algorithmus**

Quelle: Heiko (2002), S. 13.

Dieses Beispiel zeigt, dass Fuzzy-Clusteringverfahren sowohl hinsichtlich der Beschreibung der Cluster als auch bezüglich der Zugehörigkeit der Daten zu den Clustern zu Ergebnissen führen können, die genau der menschlichen Intuition entsprechen. Die Daten, die typisch für die Cluster sind, können nach der Clusteranalyse leicht identifiziert werden. Damit ist die Fuzzy-Clusteranalyse für die Datenanalyse von großem Interesse.

### 6.6.3 Probabilistische und Possibilistische Clustereinteilung

Wie im Kapitel 6.7.1 bereits erwähnt, existieren in der Fuzzy-Clusteranalyse zwei verschiedene Ansätze zur Clustereinteilung: die *probabilistische* und die *possibilistische* Clustereinteilung.

Bei der ersten Einteilung wird für jedes Datum eine Wahrscheinlichkeitsverteilung über die Cluster bestimmt, die angibt, mit welcher Wahrscheinlichkeit ein Datum einem Cluster zugeordnet wird. Dagegen wird bei der possibilistischen Einteilung jedem Datum ein Zugehörigkeits- oder Möglichkeitsgrad zugeordnet, inwieweit das Datum zum betreffenden Cluster gehört.

Wenn  $K$  die Menge der Cluster und  $X$  die Menge der zu klassifizierenden Daten darstellt, müssen bei einem probabilistischen Verfahren folgende Bedingungen erfüllt sein:<sup>899</sup>

$$1) \quad \forall x \in X : \sum_{k \in K} f(x)(k) = 1 \quad \text{und} \quad (i)$$

$$2) \quad \forall k \in K : \sum_{x \in X} f(x)(k) > 0 \quad (ii)$$

Wir interpretieren  $f(x)(k) \rightarrow [0, 1]$  als den Grad der Zugehörigkeit des Datums  $x \in X$  zum Cluster  $k \in K$  im Verhältnis zu allen anderen Clustern.

Die erste Bedingung bedeutet, dass jedes Datum in der Summe seiner Zugehörigkeiten 1 ergibt, was einer Normierung der Zugehörigkeiten je Datum entspricht. Gleichzeitig bedeutet das aber auch, dass im Vergleich mit allen Daten jedes einzelne Datum gleiches Gewicht bekommt, d. h. alle Daten werden gleichermaßen in die Clustereinteilung einbezogen. Die zweite Bedingung sagt aus, dass kein Cluster  $k$  leer ausgeht, d. h. die Zugehörigkeit  $f(x)(k)$  nicht für alle  $x$  gleich Null sein darf.

Bei einer possibilistischen Clustereinteilung muss nur die zweite Bedingung erfüllt sein.<sup>900</sup> So ist es zu erkennen, dass jede probabilistische Clustereinteilung auch eine possibilistische ist.<sup>901</sup> Bei einer probabilistischen Clustereinteilung wird das mittlere Datum am Schmetterlingbeispiel maximal mit einem Wert von 0,5 zu beiden Clustern gehören. Bei einer possibilistischen Clustereinteilung ist aber durchaus möglich, dass der Wert der Zugehörigkeit zu beiden Clustern höher liegt.

Bei der Fuzzy-Clusteranalyse werden meistens probabilistische Clusteringverfahren verwendet.<sup>902</sup> Trotz der Beliebtheit und häufigen Verwendung dieser Verfahren

<sup>899</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 17f.

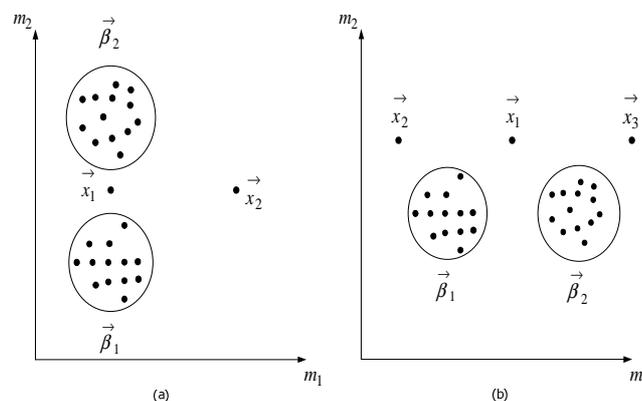
<sup>900</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 19.

<sup>901</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 19.

<sup>902</sup> Vgl. Heiko (2002), S. 25.

besitzen sie ein paar Makel, die zwar den überaus positiven Gesamteindruck dieser Verfahren nicht trüben können, aber trotzdem hier diskutiert werden müssen. Der Nachteil probabilistischer Clusteringverfahren kann anhand der in der Abbildung 62 dargestellten Beispiele (a) und (b) erklärt werden:

Die Abbildung 62 (a) zeigt einen Datensatz mit zwei Clustern. Die Daten  $\vec{x}_1$  und  $\vec{x}_2$  sind von beiden Clustern gleich entfernt. Das Datum  $\vec{x}_1$  könnte zwar wegen seiner Nähe zu  $\vec{\beta}_1$  und  $\vec{\beta}_2$  als typischen Vertreter der beiden Cluster anerkannt werden, was man aber von  $\vec{x}_2$  wegen seiner Entfernung zu den Clustern nicht behaupten kann. Anschaulich ist die Wahrscheinlichkeit für die Zugehörigkeit des Datums  $\vec{x}_2$  zu einem der Cluster sehr gering.



**Abbildung 62: Klassifikationsproblem**

Quelle: eigene Darst., in Anlehnung an Höppner/Klawonn/Kruse (1997), S. 18f.

Da aber bei der probabilistischen Clustereinteilung durch die Normierung die Summe der Zugehörigkeiten 1 ergeben, müssen die einzelnen Zugehörigkeiten der beiden Daten  $\vec{x}_1$  und  $\vec{x}_2$  zu den beiden Clustern  $\vec{\beta}_1$  und  $\vec{\beta}_2$  jeweils 0.5 betragen. Macht man sich aber diese Normierung in der Anschauung nicht bewusst, so könnte der recht hohe Wert für die Zugehörigkeit (ohne die für Klarheit sorgende Abbildung) den Eindruck erwecken, dass das Datum  $\vec{x}_2$  eben doch recht typisch für beide Cluster ist. Es wird also nicht unterschieden, dass das Datum  $\vec{x}_1$  eher beiden Clustern angehört, während das Datum  $\vec{x}_2$  eigentlich ein Stördatum ist, das keinem der beiden Cluster zugeordnet werden soll. Eine Interpretation der Zugehörigkeitsgrade  $\mu_{i,j}$  kann daher leicht zu ungenauen bzw. fehlerbehafteten Aussagen führen, wenn nicht weitere Informationen, z. B. über die Lage der Cluster, mit berücksichtigt werden.

Einen ähnlichen Fall zeigt die Abbildung 62 (b). Dieses Mal sind alle Daten sehr weit entfernt von den beiden Clustern, aber die Entfernungen der Daten links und rechts von den beiden Clustern sind unterschiedlich. Das mittlere Datum  $\vec{x}_1$  bekommt wieder die Zugehörigkeit 0.5, während die Daten  $\vec{x}_2$  und  $\vec{x}_3$  links und rechts durch die Zugehörigkeiten deutlich über 0.5 eindeutig dem jeweils dichteren Kreis zugeordnet werden. Alle Daten sind (fast) gleichermaßen untypisch für die Cluster, aber die Zugehörigkeiten verhalten sich sehr unterschiedlich.

Probabilistische Clusteringverfahren können also zu schwer interpretierbaren Zugehörigkeitsgraden und Clusteringergebnissen führen.

Eine Alternative zu probabilistischen Clusteringverfahren stellt die Anwendung von possibilistischen Clusteringverfahren dar. Im Gegensatz zu probabilistischen Verfahren ist hier für die Zugehörigkeit eines Datums  $\vec{x}_j$  zu einem Cluster  $\vec{\beta}_i$  nur der Abstand zu dem betreffenden Cluster maßgebend<sup>903</sup> und der Abstand zu den anderen Clustern wird nicht berücksichtigt<sup>904</sup>. Erst nachdem die Bedingung, dass jedes Datum in der Summe seiner Zugehörigkeiten 1 ergibt, wegfällt, können hohe Zugehörigkeiten eines Datums  $\vec{x}_j$  zu einem Cluster  $\vec{\beta}_i$  tatsächlich darauf hindeuten, dass  $\vec{x}_j$  ein typischer Vertreter von  $\vec{\beta}_i$  ist. Dann bekommen die weit entfernten Daten in der Abbildung 62 (a) und (b) geringere Zugehörigkeiten.<sup>905</sup>

Als Schlussfolgerung lässt sich nun Folgendes sagen:

Bei der possibilistischen Fuzzy-Clusteranalyse ist der Zugehörigkeitsgrad ein Maßstab, wie typisch ein Datum für einen Cluster ist, während bei der probabilistischen Fuzzy-Clusteranalyse der Zugehörigkeitsgrad eher die relative Zuordnung bzw. Lage eines Datums zu einem Cluster angibt.

#### **6.6.4 Funktionen und Bestandteile einer probabilistischen Clusteranalyse**

Das Ergebnis einer probabilistischen Clusteranalyse wird in der Regel in einer Zugehörigkeitsmatrix  $U$  angegeben. In ihr werden die Zugehörigkeiten verschiedener Daten zu einem bestimmten Cluster abgelegt.

---

<sup>903</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 38.

<sup>904</sup> Vgl. ebd., S. 38.

<sup>905</sup> Vgl. ebd., S. 19.

$$U_{i,j} = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1c} \\ u_{21} & u_{22} & \dots & u_{2c} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ u_{n1} & u_{n2} & \dots & u_{nc} \end{bmatrix} \quad \text{mit} \quad \begin{cases} i=1,\dots,c & c = \text{Clusteranzahl}, \\ j=1,\dots,n & n = \text{Datenpunkte} \end{cases}$$

$[n \times c]$

Das Element  $u_{i,j}$  dieser Matrix kennzeichnet die Zuordnung des Datenpunktes  $x_j$  zum Cluster  $k_i$ . Die Zugehörigkeits- bzw. Partitionsmatrix hat genauso viele Zeilen, wie Datenpunkte in der zu Datenmatrix  $D$  enthalten sind, und genauso viele Spalten wie Cluster in der Cluster- bzw. Clusterzentrenmatrix  $V$  vorhanden sind. Für jeden Datenpunkt hat die über alle Cluster gebildete Summe der Zuordnungen den Wert 1.

Oft liegt einer probabilistischen Clustereinteilung eine zu optimierende Funktion zugrunde<sup>906</sup>:

$$b(f) = \sum_{x \in X} \sum_{k \in K} f^m(x)(k) \cdot d^2(x,k) \quad \text{mit} \quad f : X \rightarrow F(K)$$

Diese Funktion bezeichnet man als *Bewertungs-* oder *Zielfunktion*. Die allgemeine Bewertungsfunktion der probabilistischen Clustereinteilung stellt eine Verallgemeinerung der Methode der kleinsten Quadrate dar.<sup>907</sup>

Dabei ist  $f(x)(k)$  die *Zugehörigkeitsfunktion*, welche die Zugehörigkeit des Datums  $x$  zum Cluster  $k$  angibt.  $d(x,k)$  ist eine *Distanzfunktion* und stellt ein Maß für den Abstand zwischen dem Datum  $x$  und dem des Clusters  $k$  dar. Für die Abstandsberechnung können verschiedene Distanzmaße verwendet werden. Alle Distanzen gehen quadratisch ein und somit wird die Bewertungsfunktion niemals negativ (denn die Zugehörigkeiten liegen nur zwischen 0 und 1). Die Bewertungsfunktion muss für eine optimale Lösung möglichst minimiert werden. Das heißt, je kleiner der Funktionswert ist, desto besser wird die Clustereinteilung.

Der Exponent  $m \in \mathfrak{R}_{\geq 1}$  ist der sogenannte Fuzzifier, der eine Gewichtung der Zugehörigkeit realisiert. Je größer  $m$  ist, desto schneller wird der Ausdruck  $f^m(x)(k)$  Null. Die Ausprägungen von lokalen Minima und Maxima werden damit weniger

<sup>906</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 20f.

<sup>907</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 20.

deutlich ausfallen oder sogar ganz wegfallen.<sup>908</sup> So erhält man unschärfere Ergebnisse für die Clustereinteilung. Die Wahl von  $m$  ist davon abhängig, wie gut sich die Daten damit in Cluster einteilen lassen.<sup>909</sup> Wählt man z. B.  $m=1$ , dann ist die Clustereinteilung hart, d. h. eine scharfe Trennung zwischen Clustern wird möglich (strebt  $m$  gegen eins, so konvergieren die Zugehörigkeiten gegen Null oder Eins). In diesem Fall werden die Daten, die zu dem entsprechen Cluster den geringsten Abstand haben, mit der maximalen Zugehörigkeit  $f(x)(k)=1$  diesem Cluster zugeordnet. Wählt man dagegen  $m=\infty$ , werden die Cluster voneinander nicht unterscheidbar, da die Zugehörigkeit der Daten zu allen Clustern gleich groß ist (lässt man  $m$  gegen unendlich laufen, so nähern sich die Zugehörigkeiten  $f(x)(k)$  dem Wert  $\frac{1}{c}$ ,  $c$  repräsentiert dabei die Clusteranzahl).

Es gibt keine theoretischen Erkenntnisse zur Wahl des Fuzzifiers. In der Praxis haben sich aber für  $m$  Werte zwischen 1 und 2,5 als geeignet herausgestellt.<sup>910</sup> Häufig wird als Fuzzifier  $m=2$  gewählt.<sup>911</sup>

Eine gute Datenanalyse fordert ein (globales) Minimum der Bewertungsfunktion  $b(f)$ , unter der Nebenbedingung einer probabilistischen Clustereinteilung. Da die Anzahl der möglichen Zugehörigkeiten unendlich groß ist, ist dieses Minimum jedoch nicht leicht zu finden.

Des Weiteren ist die Bewertungsfunktion mehrdimensional, so dass die bekannten eindimensionalen Verfahren zur Suche der Minima nicht angewendet werden können. Man muss sich also von vornherein mit lokalen Minima zufriedengeben.<sup>912</sup>

Für einen endlichen Datensatz  $X = \{x_1, x_2, \dots, x_n\}$  und eine endliche Menge von Clustern  $K = \{k_1, k_2, \dots, k_c\}$  lässt sich das Analyseergebnis  $f: X \rightarrow F(K)$  als  $c \times n$ -Matrix  $U$  darstellen, wobei für die einzelnen Einträge  $u_{i,j}$  der Matrix  $U$  für die Zugehörigkeit gilt:  $u_{i,j} := f(x_j)(k_i)$ . Gesucht ist nun ein Analyseergebnis, das die Bewertungsfunktion unter Berücksichtigung von (i) bis (iv) minimiert. Diese Minimierung geschieht iterativ, wobei in jedem Iterationsschritt nacheinander die

---

<sup>908</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 21.

<sup>909</sup> Vgl. ebd., S. 22.

<sup>910</sup> Vgl. Wikipedia, Begriff: Fuzzy C-Means, Abruf [02-10-2008].

<sup>911</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 22.

<sup>912</sup> Vgl. Höppner/Klawonn/Kruse., S. 22.

Matrix  $U$  und die Clustermenge  $K$  möglichst optimal aufeinander abgestimmt werden. Dieses Verfahren wird auch als *Alternation Optimization* (AO) bezeichnet.<sup>913</sup>

Sei  $f: X \rightarrow F(K)$  eine probabilistische Clustereinteilung,  $x \in X$ , setze  $u := f(x)$ ,  $u_k := f(x)(k)$  für  $k \in K$ . Die Bewertungsfunktion  $b$  sei für alle probabilistischen Clustereinteilungen in  $f$  minimal. Zudem gelten die im vorigen Kapitel eingeführten Bedingungen (i) und (ii) für  $f$ . Man kann die Bewertungsfunktion

$$b(f) = \sum_{x \in X} \sum_{k \in K} f^m(x)(k) \cdot d^2(x, k) \text{ minimieren, indem für alle } x \in X \text{ die Summe} \\ \sum_{k \in K} u_k^m \cdot d^2(x, k) \text{ minimiert wird.}$$

Nach der Anwendung des Lagrangeverfahrens auf diese Summenfunktion und der partiellen Ableitungen nach  $\lambda$  und  $u_k$  ergibt sich folgende Zugehörigkeitsfunktion:

$$f(x)(k) = u_k = \frac{1}{\sum_{j \in K} \left(\frac{d^2(x, k)}{d^2(x, j)}\right)^{\frac{1}{m-1}}} = \frac{1}{\sum_{j \in K} \left(\frac{d(x, k)}{d(x, j)}\right)^{\frac{2}{m-1}}} \quad (iii)$$

Es ist zu beachten, dass diese Formel (iii) im Fall  $d(x, j) = 0$  nicht anwendbar ist. Diese Situation liegt vor, wenn das Datum  $x$  genau im Clusterzentrum  $j$  liegt, d. h. wenn der Abstand des Datums zu dem betreffenden Cluster 0 beträgt und damit selbst diesen Cluster repräsentiert bzw. vertritt. In diesem Fall ist die Zugehörigkeit des Datums zu dem Cluster genau 1.

Die Minimierung der Bewertungsfunktion durch partielle Ableitungen wird in dem Buch *Fuzzy-Clusteranalyse* von Höppner, Klawonn und Kruse detailliert beschrieben.<sup>914</sup> Dort wird auch eine Alternative für eine Bewertungsfunktion vorgeschlagen, mit der man Algorithmen zur possibilistischen Clustereinteilung realisieren kann.

Außer den partiellen Ableitungen nach  $\lambda$  und  $u_k$  erfordert die Minimierungsaufgabe auch die Bildung der Richtungsableitung der Bewertungsfunktion  $b$  nach  $k_i \in K, i \in \mathbb{N}_{\leq c}$ . So wird die Bewertungsfunktion bezüglich allen probabilistischen Clustereinteilungen  $X \rightarrow F(K)$  mit  $K = \{k_1, k_2, k_3, \dots, k_c\}$  bei gegebenen

<sup>913</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 28

<sup>914</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 23f.

Zugehörigkeiten  $f(x_j)(k_i) = u_{i,j}$  durch  $f : X \rightarrow F(K)$  minimiert, indem man alle Richtungsableitungen von  $b$  nach  $k_i$  bildet. So gilt für alle  $\xi \in \mathbf{R}^p$  mit  $t \in \mathbf{R}$  folgendes:

$$\begin{aligned}
 0 &= \frac{\partial}{\partial k_i} \sum_{j=1}^n \sum_{i=1}^c u_{i,j}^m |x_j - k_i|^2 \\
 &= \sum_{j=1}^n u_{i,j}^m \frac{\partial}{\partial k_i} |x_j - k_i|^2 \\
 &= \sum_{j=1}^n u_{i,j}^m \lim_{t \rightarrow 0} \frac{||x_j - (k_i + t\xi)|^2 - |x_j - k_i|^2}{t} \\
 &= \sum_{j=1}^n u_{i,j}^m \lim_{t \rightarrow 0} \frac{1}{t} \left( ((x_j - k_i) - t\xi)^T ((x_j - k_i) - t\xi) - (x_j - k_i)^T (x_j - k_i) \right) \\
 &= \sum_{j=1}^n u_{i,j}^m \lim_{t \rightarrow 0} \frac{-2t(x_j - k_i)^T \xi + t^2 \xi^T \xi}{t} \\
 &= -2 \sum_{j=1}^n u_{i,j}^m (x_j - k_i)^T \xi,
 \end{aligned}$$

es folgt

$$\begin{aligned}
 0 &= \frac{\partial}{\partial k_i} b \\
 \Leftrightarrow \sum_{j=1}^n u_{i,j}^m \cdot (x_j - k_i) &= 0 \\
 \Leftrightarrow k_i &= \frac{\sum_{j=1}^n u_{i,j}^m \cdot x_j}{\sum_{j=1}^n u_{i,j}^m} \quad (iv)
 \end{aligned}$$

Mit dieser Gleichung werden die Clustermittelpunkte (bzw. -zentren)  $k_i$  berechnet. Dies geschieht iterativ in Form einer verallgemeinerten Mittelwert-Bildung. Das bei jedem Iterationsschritt resultierende Ergebnis für Clusterzentren wird in einer Matrix  $V$  gespeichert. Der Ausdruck  $||x_j - k_i|^2$  drückt eine Norm für Ähnlichkeits- bzw. Distanzmaße (z. B. euklidische Distanz) aus.

### 6.6.5 Probabilistischer Clustering-Algorithmus

Im vorherigen Kapitel wurde ein Überblick darüber gegeben, wie eine mögliche Berechnung der probabilistischen Clustereinteilung realisiert werden kann. Nun wird

der allgemeine Algorithmus<sup>915</sup> zur probabilistischen Clustereinteilung vorgestellt. Der folgende Algorithmus liefert eine probabilistische Clustereinteilung der Eingabedaten. Dabei nähert er sich iterativ einer Lösung an. Es muss eine Abbruchschwelle  $\varepsilon$  als Genauigkeits- bzw. Abbruchkriterium festgelegt werden, bei der die Berechnung abgebrochen und ein Clusteringergebnis zurückgeliefert wird.

---

*Algorithmus zur probabilistischen Clustereinteilung*

---

*Gegeben sei eine Datenmenge  $X = \{x_1, x_2, \dots, x_n\}$ . Jedes Cluster sei durch sein Clusterzentrum eindeutig charakterisierbar.*

*Wähle die Anzahl  $c$  der Cluster,  $2 \leq c < n$ ;*

*Wähle ein  $m \in \mathfrak{R}_{\geq 1}$ ;*

*Wähle eine Abbruchgenauigkeit  $\varepsilon$  und die maximale Anzahl der Iterationsschritte  $I$ ;*

*Initialisiere eine Start – Zugehörigkeitsmatrix  $U^{(0)}$ ;*

*Setze den Iterationsschritt auf Null  $\rightarrow I := 0$ ;*

*REPEAT*

*Erhöhe den Iterationsschritt  $I$  um 1;*

*Berechne die Clusterzentren  $V^{(I)}$  mittels Zugehörigkeitsmatrix  $U^{(I-1)}$ ;*

*Berechne die neue Zugehörigkeitsmatrix  $U^{(I)}$  mittels  $V^{(I)}$ ;*

*UNTIL Änderung der Clusterzentren bzw. Änderung der Zugehörigkeiten kleinergleich als die Abbruchgenauigkeit, d.h.  $|U^{(I-1)} - U^{(I)}| \leq \varepsilon$  oder Überschreitung der maximalen Anzahl der Iterationsschritte.*

---

Durch die Wahl des Distanzmaßes und des Fuzzifiers kann das Resultat des Algorithmus wesentlich beeinflusst werden.

Der Nachteil des Algorithmus ist, dass die Anzahl der Cluster vorher bekannt sein muss. Beim Vorliegen einer Bewertungsfunktion, durch die sämtliche Einteilungen mit unterschiedlicher Clusteranzahl miteinander verglichen werden können, ist es schon denkbar, den Algorithmus für mehrere Clustergrößen (2-Cluster, 3-Cluster, ..., N-Cluster) ablaufen zu lassen und dann für die eigentliche Verwendung jene Clustereinteilung auszuwählen, die im Vergleich zu anderen das beste Ergebnis liefert. Die Bewertungsfunktion  $b$  ist jedoch für einen solchen Vergleich eher ungeeignet, da diese Funktion bei wachsender Clusteranzahl  $c$  monoton fallend ist.<sup>916</sup> Die Ermittlung der optimalen Clusteranzahl kann also nicht über ein Minimum oder Maximum dieser

---

<sup>915</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 28.

<sup>916</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 31 und 157.

Bewertungsfunktion angegeben werden. Im Kapitel 6.6.8 werden andere Funktionen (auch Gütefunktionen genannt) vorgestellt, die diese Aufgabe übernehmen können.

### 6.6.6 Fuzzy-C-Means-Algorithmus

Die Fuzzy-Clusteranalyse gehört zu den zielfunktionsbasierten (engl.: objective-function) unüberwachten Klassifikations- bzw. Clusteringverfahren, bei denen ein Datensatz klassifiziert wird, ohne die Klassen vorher zu kennen.<sup>917</sup> Das Klassifikationsproblem wird durch eine Bewertungs- bzw. Zielfunktion beschrieben, die unter Berücksichtigung von Restriktionen zu optimieren ist. Die Clusteranalyse zielt darauf ab, Daten in homogene Gruppen bzw. Klassen oder Cluster zu unterteilen. Jedes Cluster wird dabei durch ein repräsentatives bzw. typisches Datum eindeutig beschrieben. Demnach ist die Forderung, dass Daten, die zu einem Cluster gehören, homogen sein sollen, so zu interpretieren, dass diese Daten dem typischen Datum möglichst ähnlich sein sollen.

Um dem Homogenitäts- bzw. Ähnlichkeitskriterium gerecht zu werden, kann man den euklidischen Abstand verwenden. Das Homogenitätskriterium wird erst dann erfüllt, wenn der Abstand zwischen den Daten eines Clusters und dem typischen Datum möglichst klein ist. Das ist die Grundidee der bei den Fuzzy-Clusteringverfahren verwendeten Bewertungsfunktion: *„Die Daten sind den Clustern so zuzuordnen, dass die Summe der Abstände zwischen den Clustern und den ihnen zugeordneten Daten minimal wird.“*<sup>918</sup>

Der Fuzzy-C-Means-Algorithmus (FCM) ist der Basisalgorithmus aller Fuzzy-Clusteranalyseverfahren nach der probabilistischen Clustereinteilung. Er ist eng mit dem Hard C-Means-Algorithmus von Duda und Hart verwandt und kann als seine Fuzzy-Variante verstanden werden. Der Algorithmus versucht, einen Datensatz in C-Cluster einzuteilen, die durch ihren Mittelwert beschrieben werden. Das „C“ in seinem Namen hat der Algorithmus dadurch, dass die Anzahl der Cluster a priori angegeben werden muss.<sup>919</sup> Der Fuzzy-C-Means-Algorithmus ist also für eine feste Anzahl von Clustern vorgesehen, das heißt, er bestimmt diese nicht selbstständig.<sup>920</sup> Das Wort „Means“ kommt von der Abschätzung der Clusterzentren durch die Mittelwertbildung.

Der Unterschied zwischen Hard-C-Means und Fuzzy-C-Means-Verfahren ist die Zuordnung der Daten zu den Clustern. Während bei dem Hard-C-Means-Verfahren die

---

<sup>917</sup> Vgl. Zimmermann (1991), S. 220-225.

<sup>918</sup> Heiko (2002), S. 14.

<sup>919</sup> Es handelt sich z. B. bei vier Clustern, also die Anzahl der Cluster beträgt 4, um den Fuzzy-4-Means.

<sup>920</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 35.

Daten den Clustern eindeutig zugeordnet werden, verwendet man bei dem Fuzzy-C-Means-Verfahren graduelle Zugehörigkeiten zwischen 0 und 1. Das ermöglicht es, die Übergänge zwischen den verschiedenen Clustern geeignet zu modellieren. Der Fuzzy-C-Means-Algorithmus erkennt kugelförmige Punktwolken im p-dimensionalen Raum.<sup>921</sup> Die Cluster werden als gleich groß angenommen. Jeder Cluster wird durch seinen Mittelpunkt dargestellt. Diese Repräsentation jedes einzelnen Clusters wird auch „Prototyp“ genannt, da sie als Stellvertreter aller dem betreffenden Cluster zugeordneten Daten angesehen wird.<sup>922</sup>

Gegeben sei eine Datenmenge  $X = \{x_1, x_2, x_3\}$ . Dann sind mehrere Fuzzy-2-Means-Partitionen möglich, wie z. B.:

$$\tilde{\mu}_1 = \begin{bmatrix} x_1 & x_2 & x_3 \\ 1 & 0.5 & 0 \\ 0 & 0.5 & 1 \end{bmatrix}, \tilde{\mu}_2 = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.8 & 0.5 & 0.2 \\ 0.2 & 0.5 & 0.8 \end{bmatrix}, \tilde{\mu}_3 = \begin{bmatrix} x_1 & x_2 & x_3 \\ 0.8 & 1 & 0.9 \\ 0.2 & 0 & 0.1 \end{bmatrix}, \dots \text{ usw.}$$

Die Lage eines Clusters wird durch sein Clusterzentrum  $k_i$  angegeben. Ausgehend von der Lage des Clusterzentrums wird die Entfernung der Daten zum betreffenden Cluster berechnet. Generell verwendet der Algorithmus als Distanzmaß den euklidischen Abstand zwischen den Daten  $x_k$  und den Clusterzentren  $k_i = (k_{i1}, \dots, k_{ip}) \in \mathbf{R}^p$ ,  $i = 1, \dots, c$ .<sup>923</sup>

Dabei ist  $c$  die Clusteranzahl bzw. Partition, und  $\mathbf{R}^p$  stellt einen p-dimensionalen Vektorraum dar. Man kann  $k = (k_1, \dots, k_c) \in \mathbf{R}^{cp}$  als den Vektor aller Clusterzentren verstehen. Die Clusterzentren bzw. Cluster werden in einer Matrix  $K$  (statt  $K$  wird sie häufig mit dem Buchstaben  $V$  gekennzeichnet) gespeichert, die folgendermaßen aussieht:

$$K_{i,j} = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1p} \\ k_{21} & k_{22} & \dots & k_{2p} \\ \dots & \ddots & & \dots \\ \dots & \ddots & & \dots \\ \dots & \ddots & & \dots \\ k_{c1} & k_{c2} & \dots & k_{cp} \end{bmatrix} \quad \text{mit} \quad \begin{cases} i = 1, \dots, c & c = \text{Clusteranzahl,} \\ j = 1, \dots, p & p = \text{Merkmale} \end{cases}$$

[c×p]

<sup>921</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 35.

<sup>922</sup> Vgl. ebd., S. 35.

<sup>923</sup> Vgl. Zimmermann (1991), S. 231.

Bei dieser Matrix werden die Cluster  $c$  durch die Zeilen und die Dimensionen bzw. Merkmale jedes einzelnen Clusters  $p$  durch die Spalten charakterisiert.

$$D_{i,j} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad \text{mit} \quad \begin{cases} i = 1, \dots, n & n = \text{Datenpunkte,} \\ j = 1, \dots, p & p = \text{Merkmale} \end{cases}$$

$[n \times p]$

Die zu klassifizierenden Daten können ebenfalls mit einer Matrix  $D$  dargestellt werden.

### 6.6.7 Ähnlichkeits- bzw. Distanzmaße

Clusteringverfahren gehen davon aus, dass durch die Merkmalswerte, die die zu klassifizierenden Objekte bzw. Daten beschreiben, gemessen werden kann, wann beispielsweise zwei Objekte  $x$  und  $k$  einen großen Abstand (Unähnlichkeit) voneinander oder eine große Ähnlichkeit zueinander haben.<sup>924</sup> Der Wahl des geeigneten Ähnlichkeitsmaßes kommt daher eine wichtige Rolle beim Clusteringverfahren zu. Schließlich unterteilt die Clusteranalyse eine Menge von ungeordneten Objekten bzw. Daten in Cluster, deren Elemente zueinander ähnlich sind.

Das Ähnlichkeitsmaß gibt den Abstand einzelner Objekte voneinander als Zahlenwert an. Dieser Wert beträgt zwischen den Objekten eines Clusters null, da alle Mitglieder dieses Clusters ähnlich sind. Hingegen ist der Abstand zwischen den Objekten unterschiedlicher Cluster größer als null.

Im zweidimensionalen Raum wird der Abstand zwischen zwei Datenpunkten  $k$  und  $x$  mit den Datensätzen  $x = (x_1, x_2)$  und  $k = (k_1, k_2)$  durch die euklidische Metrik bzw. Distanz bestimmt<sup>925</sup>:

$$d(x, k) = \left( \sum_{i=1}^2 |x_i - k_i|^2 \right)^{\frac{1}{2}} = \sqrt{(x_1 - k_1)^2 + (x_2 - k_2)^2}.$$

<sup>924</sup> Vgl. Zimmermann (1995), S. 39.

<sup>925</sup> Vgl. Mertens/Wieczorrek (2000), S. 216.

Für metrischskalierte  $p$ -dimensionale Vektorräume lässt sich das Ähnlichkeitsmaß für zwei Punkte  $k$  und  $x$  allgemein als Minkowski-Metrik-Distanz

$$d(x, k) = \left( \sum_{i=1}^p |x_i - k_i|^r \right)^{\frac{1}{r}}$$

angeben, mit den Datensätzen bzw. -vektor  $x = (x_1, x_2, x_3, \dots, x_p)$  und  $k = (k_1, k_2, k_3, \dots, k_p)$ ,  $p > 0$  und dem ganzzahligen Wert  $r > 0$ .

Bei  $r=1$  liegt eine Manhattan-Metrik vor, im Fall  $r=2$  die sogenannte Euklidische Metrik bzw. Distanz<sup>926</sup>.

Je größer der Wert für  $r$  ist, desto stärker werden stark abweichende Werte der einzelnen Variablen von Datensätzen berücksichtigt und desto sensibler ist die Minkowski-Metrik-Distanz für Ausreißer einzelner Werte.<sup>927</sup> Die Manhattan-Metrik hat also gegenüber der Euklidischen-Metrik den Nachteil, da die geometrische Distanz überschätzt wird, und andererseits auch den Vorteil, da die Ausreißerwerte nicht so stark berücksichtigt werden, weil nur die Differenzen und nicht die quadrierten Differenzen in die Abstandsfunktion eingehen. Kleine wie große Differenzen  $|x_i - k_i|$  werden bei Manhattan-Metrik gleich bewertet, bei Euklidischen-Metrik fallen große Differenzen stärker ins Gewicht.

In der Literatur wird zur Clusteranalyse von metrisch skalierten Daten häufig eine Minkowski-Metrik oder die quadrierte Mahalanobis-Distanz als Distanz- bzw. Ähnlichkeitsmaß verwendet.

Für die quadrierte Mahalanobis-Distanz gilt:

$$d^2(x, k) = (x - k)^T A^{-1} (x - k), \quad \text{mit dem Datensatz } x = (x_1, x_2, x_3, \dots, x_p), \text{ dem Clusterzentrum } k = (k_1, k_2, k_3, \dots, k_p).$$

Da das Datum  $x$  und das Clusterzentrum  $k$  aus mehreren Variablen ( $p$ -dimensionaler Raum) bestehen, kann man diese als Vektorform darstellen. Somit ist der

<sup>926</sup> Diese Metrik, mit der man die Unähnlichkeit zwischen Datenpunkten in einem Cluster und dessen Clusterzentrum bestimmt, wird auch als Varianz-Kriterium der Partition bezeichnet, vgl. Zimmermann (1991), S. 231.

<sup>927</sup> Vgl. Mertens/Wieczorrek (2000), S. 217.

Mahalanobis-Abstand zwischen einem Datum  $\vec{x}_j$  und einem Cluster  $C_i$  mit dem Clusterzentrum  $\vec{k}_i$  folgendermaßen definiert:

$$d^2(\vec{x}_j, \vec{k}_i) = \left( \vec{x}_j - \vec{k}_i \right)^T A_i^{-1} \left( \vec{x}_j - \vec{k}_i \right) \quad \begin{cases} j = 1, \dots, n ; \quad i = 1, \dots, c \text{ mit} \\ n = \text{Datenobjekte und } c = \text{Clusteranzahl} \end{cases}$$

$A_i$  ist die Kovarianzmatrix<sup>928</sup>, welche die Korrelationen zwischen allen Komponenten des Merkmals- bzw. Dimensionsvektors  $\vec{p} = p = (m_1, m_2, \dots, m_z)$  im i-ten Cluster wiedergibt. Über diese symmetrische und positiv definite Gewichtungsmatrix  $A_i$  wird das innere Produkt der Vektorkomponenten definiert und damit die Norm bzw. die Metrik im Merkmalsraum festgelegt.

Die Bezeichnung  $A_i^{-1}$  steht für die Inverse der  $(z \times z)$ -Kovarianzmatrix.

$$A_i = \begin{bmatrix} \text{cov}_{1,1} & \text{cov}_{1,2} & \dots & \text{cov}_{1,p} \\ \text{cov}_{2,1} & \text{cov}_{2,2} & \dots & \text{cov}_{2,p} \\ \dots & \dots & \ddots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \text{cov}_{p,1} & \text{cov}_{p,2} & \dots & \text{cov}_{p,p} \end{bmatrix} \quad \begin{cases} \text{cov}_{j,k} = \text{cov}(m_j, m_k) = \text{Kovarianz der entsprechenden} \\ \text{Merkmalpaare } (m_j, m_k) \text{ im } i\text{-ten Cluster} \\ p = \text{Anzahl der Merkmale bzw. Dimensionen} \\ j, k \in \{1, \dots, p\} \end{cases}$$

$[p \times p]$

Die Kovarianzen der Merkmalpaare  $\text{cov}(m_j, m_k)$  im i-ten Cluster werden wie folgt berechnet:

$$\text{cov}_{i,j,k} = \frac{1}{n_i} \sum_{i=1}^{n_i} (m_{j_i} - \bar{m}_j)(m_{k_i} - \bar{m}_k) \quad \begin{cases} \text{für } j, k \in \{1, \dots, p\} \text{ und} \\ n = \text{Datenobjekte im } i\text{-ten Cluster} \end{cases}$$

Dabei sind  $\bar{m}_j$  und  $\bar{m}_k$  die Mittelwerte der betreffenden Merkmale  $j$  und  $k$  im i-ten Cluster.

$$\bar{m}_j = \frac{1}{n_i} \sum_{i=1}^{n_i} m_{j_i}, \quad \bar{m}_k = \frac{1}{n_i} \sum_{i=1}^{n_i} m_{k_i}$$

<sup>928</sup> Wenn man als Kovarianzmatrix die Einheitsmatrix verwendet, erhält man dann den euklidischen Abstand.

Mit der Kovarianz  $\text{cov}_{i,j,k}$  der Merkmale  $m_j$  und  $m_k$  im  $i$ -ten Cluster erhält man folgenden Mahalanobis-Abstand zwischen einem Datum  $x_j$  und einem Cluster  $C_i$  mit dem Clusterzentrum  $k_i$ :

$$d^2(x_j, k_i) = \sum_{j=1}^p \sum_{k=1}^p a_{m_j, m_k} (m_{j_i} - m_{j_j})(m_{k_i} - m_{k_j}) \left\{ \begin{array}{l} p = \text{Merkmals- bzw. Dimensionanzahl} \\ a_{m_j, m_k} = \text{das entsprechende Element der} \\ \text{Kovarianzmatrix des } i\text{-ten Clusters} \end{array} \right.$$

oder

$$d(x_j, k_i) = \sqrt{\sum_{j=1}^p \sum_{k=1}^p a_{m_j, m_k} (m_{j_i} - m_{j_j})(m_{k_i} - m_{k_j})}$$

Als Schlussfolgerung könnte man sagen, dass die Mahalanobis-Distanz eine generalisierte Form der euklidischen Distanz darstellt: Durch die Multiplikation mit dem Inversen der Kovarianzmatrix (und somit einer Gewichtung mittels Kovarianzen) wird aus der euklidischen Distanz die Mahalanobis-Distanz. Der Vorteil der Mahalanobis-Distanz gegenüber der euklidischen Distanz liegt in der Möglichkeit zur Berücksichtigung von Korrelationen zwischen den Merkmalen.<sup>929</sup>

Die Korrelation zwischen den verschiedenen Merkmalen führt i. d. R. dazu, dass korrelierte Merkmale die Distanzbestimmung stärker beeinflussen als unkorrelierte: Die Einbeziehung von stark korrelierenden Merkmalen in die Clusterbildung führt bei der euklidischen Distanz zu einer Überbetonung der hinter den korrelierenden Daten stehenden Dimensionen. Dies bedeutet wiederum, dass diese korrelierten Merkmale eine höhere Gewichtung bei der Clusterbildung bekommen.<sup>930</sup> In diesem Fall entscheidet man zwischen zwei Möglichkeiten: entweder eliminiert man korrelierte Merkmale aus dem Merkmalskomplex, indem diese nicht mehr berücksichtigt werden, oder man verwendet als Gütekriterium der gefundenen Partition die Mahalanobis-Distanz.

Die Verwendung der Mahalanobis-Distanz als Abstandsmaß erfordert jedoch einen höheren Rechenaufwand gegenüber den Minkowski-Metriken, darunter auch die Euklidische Metrik, da nach jeder Änderung der Clusterzentren im Iterationsschritt die

<sup>929</sup> Vgl. Zimmermann (1995), S. 40.

<sup>930</sup> Merkmale werden umso schwächer gewichtet, je höher ihre Korrelation zu anderen Merkmalen ist. Merkmalen, die nicht korreliert sind, kommt eine höhere Bedeutung in der Clusteranalyse zu, vgl. Zimmermann (1995), S. 40.

inverse Kovarianzmatrix zwischen den Datenmerkmalen in den veränderten Clustern neu berechnet werden muss.

### **6.6.8 Bewertung der Clustereinteilung - Bestimmung der Clusteranzahl**

Die Fuzzy-Clusteranalyse ist ein Verfahren zur unüberwachten Klassifikation. Die Anzahl der Cluster muss dabei vorgegeben werden. Da aber die mögliche optimale Clustereinteilung oft a priori nicht bekannt ist, muss das Ergebnis einer Fuzzy-Clusteranalyse auf die Richtigkeit hin überprüft sowie bewertet werden. Für die Bewertung werden sogenannte *Cluster-Gütemaße* verwendet.

Es gibt viele Gründe dafür, warum eine bestehende Clustereinteilung bewertet werden sollte. Bei einer Clustereinteilung treten zahlreiche Probleme und Fragen auf, die eine Bewertung notwendig machen<sup>931</sup>:

- Wenn man alle existierenden Clustering-Algorithmen auf eine vorhandene Datenmenge anwendet, so bekommt man eine Vielzahl unterschiedlicher Einteilungen. Welcher Algorithmus liefert aber die richtige Zuordnung?
- Wenn die Anzahl der Cluster nicht a priori bekannt ist, so lässt sich nur für jede mögliche Clusterzahl eine Einteilung ermitteln. Welche Einteilung bzw. Clusterzahl entspricht am besten der Klassifikationsaufgabe?
- Die Clustering-Algorithmen nehmen i. d. R. eine Einteilung vor, ohne Rücksicht darauf, ob diese herausgefundene Einteilungsstruktur in den Daten tatsächlich vorhanden ist.

Wie bereits erwähnt, muss für die Durchführung eines probabilistischen Algorithmus zur Clustereinteilung die Anzahl der Cluster vorher angegeben werden. Diese Anzahl ist aber oft a priori noch nicht bekannt. Mit Hilfe einer Gütefunktion, die alle möglichen Clustereinteilungen miteinander vergleichbar macht, kann man leicht die beste bzw. optimale Clusterzahl bestimmen und diese für die eigentliche Klassifikationsaufgabe verwenden.

Solange sich die Daten im zwei- oder dreidimensionalen Raum darstellen lassen, fällt es einem menschlichen Betrachter leicht, Clustereinteilungen, die von verschiedenen Algorithmen gemacht wurden, zu bewerten. Bei umfangreichen Daten mit einer höheren Dimensionalität ist diese Aufgabe für den Menschen jedoch schwer zu lösen.

---

<sup>931</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 153-157.

In solchen Fällen ist eine Bewertungsfunktion unmittelbar hilfreich für die Bestimmung der richtigen Clustereinteilung bzw. -anzahl. Es ist keinesfalls eine Seltenheit, dass jene Daten für eine Klassifikationssaufgabe vorliegen, in denen überhaupt keine Struktur vorhanden ist. Die meisten Algorithmen nehmen auf die Struktur der Daten keine Rücksicht und nehmen trotzdem eine Clustereinteilung vor. Wenn die Güte für die gefundene Einteilung, welche als optimal bzw. richtig eingeschätzt wurde, immer noch sehr gering ist, liegt die Ursache vermutlich in den Daten.

Es gibt sehr unterschiedliche Arten von Gütemaßen bzw. -funktionen, die stark auf den jeweiligen Verwendungszweck ausgerichtet sind.<sup>932</sup> Daher ist es notwendig, für den entsprechenden Verwendungszweck das richtige Gütemaß auszuwählen. Man unterscheidet zwischen globalen Gütemaßen, die eine Klassifikation samt mit allen Clustern als Ganzes bewerten, und lokalen Gütemaßen, bei denen jeder Cluster separat bzw. einzeln bewertet wird.<sup>933</sup>

Um eine Clusteranalyse bei einer unbekanntem Clusteranzahl durchzuführen, gibt es unterschiedliche Ansatzmöglichkeiten<sup>934</sup>:

- Die erste Möglichkeit ist, den Fuzzy-Clusteringalgorithmus mit einer unterschiedlichen Zahl von Clustern durchzuführen. Das Ergebnis für jeweils verschiedene Clusterzahl wird nach jedem Durchlauf mit Hilfe einer Gütefunktion bewertet. Diese Gütefunktion liefert für jede mit einer bestimmten Zahl von Clustern durchgeführten Clustereinteilung einen Wert, mit dem man die Analyseergebnisse indirekt miteinander vergleichen kann. Anschließend wird die Clusterzahl, die das beste Ergebnis liefert, für die eigentliche Klassifizierungsaufgabe ausgewählt.
- Eine andere Vorgehensweise ist, mit einer eher größeren Anzahl von Clustern den Clusteringalgorithmus durchzuführen und die Clusteringergebnisse mittels einer Gütefunktion miteinander zu vergleichen, um ähnliche Cluster zu identifizieren. Diese ähnlichen Cluster (die zueinander kompatibel sind) werden dann zusammengefasst. Bei jedem erneuten Durchlauf des Clusteringalgorithmus werden neue ähnliche Cluster gefunden und diese zu einer einzigen Cluster vereinigt. Dadurch wird die Clusteranzahl schrittweise verringert, so dass schließlich bei einer Clustereinteilung nur noch Cluster übrig bleiben, die zueinander inkompatibel, d. h. ganz unterschiedlich sind. Die

---

<sup>932</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 156.

<sup>933</sup> Vgl. ebd., S. 157 und 171.

<sup>934</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 30ff.

auf diese Weise ermittelte Clustereinteilung enthält dann laut Gütefunktion nur gute Cluster. Ein bekannter Fuzzy-Clusteringalgorithmus, der auf dieser Idee beruht, ist z. B. der „Compatible-Cluster-Merging-Algorithmus“ (CCM).

- Statt kompatible Cluster zu identifizieren ist es auch möglich, nach der Durchführung des Clusteringalgorithmus nach schlechten Clustern zu suchen und diese zu entfernen.<sup>935</sup>
- Die letzte Möglichkeit wäre z. B. den Fuzzy-Algorithmus für mehrere Clusterzahl durchzuführen, dabei jeweils den Grad der Unschärfe zu ermitteln und anschließend diejenige Einteilung als optimal einzustufen, die einer harten Einteilung am nächsten kommt.<sup>936</sup>

Im Folgenden werden die wichtigsten globalen Gütemaße vorgestellt.

Globale Gütemaße bewerten eine komplette Clustereinteilung. Sie geben die Güte einer gesamten Clustereinteilung durch einen einzigen reellen Wert an. Das einfachste globale Gütemaß ist die zu minimierende Zielfunktion bzw. Bewertungsfunktion  $b$ .

Da die zu minimierende Zielfunktion die Klassifikationsaufgabe beschreibt, kann man aus einem niedrigeren Wert dieser Funktion auf ein besseres Klassifikationsergebnis schließen. Jedoch ist die Bewertungsfunktion  $b$  als Gütefunktion nur bedingt geeignet, denn bei einer Erhöhung der Clusterzahl wird der Abstand der Daten zu den dichtesten Prototypen (Clusterzentren) zwangsweise immer geringer, d. h. die Funktion  $b$  ist für steigendes  $c$  monoton fallend.<sup>937</sup> Die optimale Clusterzahl kann also nicht über ein Minimum oder Maximum dieser Funktion ermittelt werden. Es ist zwar möglich, bei derartigen monotonen Gütefunktionen die optimale Einteilung alternativ durch einen Scheitelpunkt der Funktion zu bestimmen (auch als Ellenbogen-Kriterium genannt), jedoch werden dazu andere Funktionen verwendet. Einige dieser Gütefunktionen bewerten nur die Zuordnung der Daten zu den Clustern; andere hingegen bewerten die Form der Cluster oder wie gut die Cluster separiert sind.<sup>938</sup> Da die Gütefunktionen unterschiedliche Aspekte einer guten Klassifikation betrachten, ist es problematisch, von dem besten Gütemaß zu sprechen. Stattdessen sollten bei der Bewertung einer Clustereinteilung mehrere Gütemaße betrachtet werden.

---

<sup>935</sup> Auch bei diesem Ansatz wird die Clusteranalyse mit einer größeren Anzahl von Clustern durchgeführt, vgl. Höppner/Klawonn/Kruse (1997), S. 31.

<sup>936</sup> Vgl. ebd., S. 155.

<sup>937</sup> Vgl. ebd., S. 157. Je mehr Cluster erlaubt sind, desto kleiner wird die Bewertungsfunktion, bis schließlich jedem Datum ein eigener Cluster zugeordnet wird und  $b$  der Wert null annimmt, vgl. ebd., S. 31.

<sup>938</sup> Eine detaillierte Beschreibung verschiedener Gütefunktionen ist in: Höppner/Klawonn/Kruse (1997), S. 157-210 zu finden.

Im Folgenden werden die globalen Gütemaßen vorgestellt, die die Zuordnung der Daten zu den Clustern berücksichtigen:

#### *Partitionskoeffizient (Pk)*

Wenn unter einer guten Klassifikation, eine möglichst harte Clustereinteilung verstanden wird, ist der Partitionskoeffizient (Pk) ein geeignetes Maß.<sup>939</sup> Er beruht auf der Idee, dass bei einer guten Klassifikation die Daten eindeutig den Clustern zugeordnet werden können. Die Zugehörigkeitsgrade sollten möglichst nahe bei 1 bzw. nahe bei 0 sein. So deutet eine unscharfe Zuordnung auf eine schlechte Klassifikation hin.

Seien  $X = \{x_1, x_2, \dots, x_n\}$  die  $n$  Datenobjekte, die in  $c$  Cluster eingeteilt wurden und  $u_{i,j}$  die Zugehörigkeit von Objekt  $j$  zu Cluster  $i$  ( $j = 1, \dots, n; i = 1, \dots, c$ ), so gilt:

$$Pk = \sum_{j=1}^n \sum_{i=1}^c \frac{(u_{i,j})^2}{n} = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c (u_{i,j})^2 .$$

Dabei gilt für jede probabilistische Clustereinteilung folgende Beziehung:

$$\frac{1}{c} \leq Pk \leq 1 \quad (v)$$

Enthält eine Clustereinteilung keine Information, ist also jedes Datum jedem Cluster mit gleichen Teilen zugeordnet. Daraus lässt sich der Minimalwert ableiten.

Da  $u_{i,j} \in [0, 1]$  und  $\sum_{i=1}^c u_{i,j} = 1$  gilt, folgt  $\sum_{i=1}^c (u_{i,j})^2 \leq 1$ , und somit auch  $Pk \leq 1$ . Der

Maximalwert  $Pk = 1$  gilt im Falle einer harten Einteilung. Also je größer  $Pk$  ist, desto unschärfer ist die Clustereinteilung. Stehen mehrere Einteilungen zur Wahl (mit gleichen Clusterzahlen), dann ist diejenige Einteilung mit dem maximalen Partitionskoeffizienten die Clustereinteilung mit der eindeutigsten Zuordnung der Daten zu den Clustern.<sup>940</sup>

#### *Partitionsentropie (Pe)*

Ebenso wie bei dem Partitionskoeffizienten wird auch bei dem Gütemaß der

<sup>939</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 158.

<sup>940</sup> Vgl. ebd., S. 158.

Partitionsentropie das Ergebnis des Fuzzy-Clusteringverfahrens nur unter Verwendung der Zugehörigkeitsgrade der Daten zu den Clustern beurteilt.

Die Partitionsentropie basiert auf der aus der Informationstheorie bekannten Shannon-Entropie.<sup>941</sup> Der Begriff Entropie bezeichnet den mittleren Informationsgehalt einer Nachrichtenquelle.<sup>942</sup> Die gesuchte Information ist die Zuordnung eines Datums  $j$  zu einem Cluster  $i$ . Bei einer probabilistischen Clustereinteilung kann die Zugehörigkeit  $u_{i,j}$  als die Wahrscheinlichkeit für die Zugehörigkeit eines Datums  $j$  zu einem Cluster  $i$  aufgefasst werden: Wenn man nun von irgendeiner Quelle die Information erhält, dass das Datum  $j$  dem Cluster  $i$  zuzuordnen ist, so hat diese Information umso mehr Informationsgehalt, desto geringer die Wahrscheinlichkeit  $u_{i,j}$  für diese Klassifizierung ist.<sup>943</sup>

Der Informationsgehalt für die Zugehörigkeit wird durch  $I_{i,j} = -\ln(u_{i,j})$  angegeben.

Dann ist  $H = \sum_{i=1}^c u_{i,j} I_{i,j}$  der mittlere Informationsgehalt der Quelle als Entropie. Die

Partitionsentropie (Pe) ist demnach wie folgt definiert:

$$Pe = \frac{\sum_{j=1}^n \sum_{i=1}^c u_{i,j} \cdot I_{i,j}}{n} = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{i,j} \cdot \ln(u_{i,j})$$

Analog zum Partitionskoeffizienten lässt sich auch hier eine Beziehung bestimmen. So gilt:  $0 \leq Pe \leq \ln c$  (vi).

Wenn die Daten eindeutig in Cluster eingeteilt sind, dann ist die Information über die Zuordnung der Daten zu den Clustern vollständig vorhanden und der Informationsgehalt ist Null. Bei einer Gleichverteilung der Zugehörigkeiten ist die Entropie hingegen maximal.<sup>944</sup> Wie bei dem Partitionskoeffizient wird auch bei der Partitionsentropie eine möglichst eindeutige Zuordnung der Daten zu den Clustern angestrebt.<sup>945</sup> Hierfür ist aber im Gegensatz zum Partitionskoeffizienten die Partitionsentropie zu minimieren.

<sup>941</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 158.

<sup>942</sup> Vgl. ebd., S. 159.

<sup>943</sup> Vgl. ebd., S. 159.

<sup>944</sup> Vgl. ebd., S. 159.

<sup>945</sup> Vgl. ebd., S. 159.

*Proportionsexponent (Pex)*

Der Proportionsexponent kann als Logarithmus des Maßes für die Anzahl der Einteilungen interpretiert werden, bei denen alle Daten besser als bei der vorliegenden Einteilung klassifiziert werden.<sup>946</sup> Für jede Spalte  $i$  der Zugehörigkeitsmatrix  $U$  ist er definiert als

$$Pex = -\ln \left\{ \prod_{j=1}^n \left( \sum_{i=1}^{m_j} (-1)^{i+1} \binom{c}{i} (1 - i \cdot u_j)^{c-1} \right) \right\}$$

mit folgender Beziehung:

$$0 \leq Pex \leq \infty \quad (vii)$$

Dabei ist  $m_j$  die größte ganze Zahl kleiner oder gleich  $\frac{1}{u_j}$  mit  $u_j = \max \{u_{i,j}\}_{i=1,\dots,c}$ .

$u_{i,j}$  stellt die Zugehörigkeit eines Datenobjekts  $j$  zum Cluster  $i = 1, \dots, c$  dar.

Die Wertebereiche für den Partitionskoeffizienten (Pk) und für die Partitionsentropie (Pe) sind stark abhängig von der Clusterzahl  $c$  (siehe (v) und (vi)). Bei wachsendem  $c$  werden die Wertebereiche jeweils größer. Aus diesem Grund eignen sich diese Gütemaße nicht für den Vergleich zweier Partitionen mit unterschiedlichen Clusterzahlen.<sup>947</sup> Für diesen Vergleich kann man den Partitionsexponenten heranziehen, dessen Wertebereich unabhängig von der Clusterzahl  $c$  ist. Es ist jedoch zu berücksichtigen, dass bei diesem Gütemaß kein Datum einen Zugehörigkeitsgrad von 1 zu einem Cluster haben darf, d. h. dass kein einziges Datum eindeutig einem Cluster zugeordnet werden darf.<sup>948</sup> Falls dies dennoch passiert, dann wird  $u_j = 1$  sowie  $(u_j)^{-1} = 1$  und der Summenterm für dieses Datum  $j$  Null ergeben. Damit wird auch das Produkt innerhalb des Logarithmus Null und der Proportionsexponent wird undefiniert. Der Proportionsexponent (Pex) lässt sich als ein Maß für die Anzahl der Einteilungen, die alle Daten besser als bei der vorliegenden Einteilung klassifiziert, interpretieren. Es ist aber zu beachten, dass durch die Anwendung dieses Gütemaßes selbstverständlich nicht möglich ist, ein bereits eindeutig zugeordnetes Datum immer

<sup>946</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 160.

<sup>947</sup> Vgl. Zimmermann (1995), S. 44

<sup>948</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 160.

noch besser zu klassifizieren. Die eindeutigen (harten) Zuordnungen von Daten in der Partition stellen eine Problematik für den Partitionsexponent dar.<sup>949</sup> Dieses Problem kann gelöst werden, indem man einen höheren Fuzzifier beim Fuzzy-C-Means-Algorithmus benutzt. So können harte Zuordnungen von Daten in der Partition beinahe ausgeschlossen werden.

Der Proportionsexponent liefert für vollständig unscharfe Partitionen den minimalen Wert 0 und wächst mit zunehmender Schärfe der Partition ins Unendliche. Bei der schärfsten Trennung von Daten bekommt er seinen maximalen Wert.

Um die Bestimmung der besten Clusteranzahl zu erleichtern, kann eine Gütegraphik benutzt werden, die verschiedene Gütemaße visuell darstellen lässt.

Der Partitionskoeffizient ( $P_k$ ) und die Partitionsentropie ( $P_e$ ) tendieren zu monotonem Verhalten in Abhängigkeit von der Clusterzahl  $c$ . Gemäß dieser Monotonie ist bei der Suche nach der optimalen bzw. richtigen Clusterzahl  $c$  in der Gütegraphik diejenige zu wählen, für die beim Übergang von der Clusterzahl  $c-1$  zur Clusterzahl  $c$  der Wert für die Entropie unter den steigenden Trend fällt bzw. der Wert des Partitionskoeffizienten über den fallenden Trend steigt.<sup>950</sup>

Die Stelle der optimalen Clusterzahl ist bei Betrachtung der jeweiligen Kurve in der Graphik aus allen verbundenen Werten als „Knickpunkt“ zu erkennen. Genau an dieser Stelle erfolgt die beste Trennung der Cluster (siehe auch Abb. 65).

Es ist jedoch zu beachten, dass es keine endgültigen Regeln für die Wahl der richtigen Clusterzahl gibt. Häufig muss der Anwender aufgrund seiner fachlichen Kenntnisse die gefundene Clusterzahl interpretieren und somit selbst auf ihre Richtigkeit hin überprüfen.

### **6.6.9 Ablauf des Fuzzy-C-Means-Algorithmus**

Im Folgenden wird der Fuzzy-C-Means-Algorithmus (FCM) genauer betrachtet und seine Vor- und Nachteile erläutert.

Der iterative FCM-Algorithmus baut grundsätzlich auf dem klassischen Isodata-Verfahren (engl.: iterative self-organizing data analysis technique) von Ball und Hall.<sup>951</sup>

---

<sup>949</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 160f.

<sup>950</sup> Vgl. Zimmermann (1995), S. 154.

<sup>951</sup> Vgl. Zimmermann (1995), S. 41.

Das Isodata-Verfahren ähnelt dem Hard-C-Means-Verfahren, bei dem die Daten nur einem einzigen Cluster zugeordnet werden (Hard Partition).

Im Gegensatz zu Hard C-Means-Verfahren erlaubt Isodata-Verfahren zusätzlich die dynamische Bestimmung der Clusteranzahl (also die Clusteranzahl ist nicht a priori bestimmt und fix wie bei C-Means-Verfahren<sup>952</sup>). Dafür können die Cluster vereinigt, geteilt und entfernt werden und je nach der Operation erhält man somit in der endgültigen Partition mehrere oder wenige Cluster als die Ausgangspartition enthält.

Das FCM-Verfahren erlaubt die Zuordnung von Daten in gewissem Grade zu verschiedenen Clustern und die Daten können somit mehreren Clustern gleichzeitig angehören (Fuzzy Partition).

Der FCM-Algorithmus wurde 1973 erstmals von Joe Dunn als eine Fuzzy-Variante des Isodata-Algorithmus für einen Spezialfall (Fuzzifier  $m = 2$ ) vorgestellt und von Jim Bezdek auf die endgültige Variante (Fuzzifier  $m > 1$ ) verallgemeinert.<sup>953</sup>

Das FCM-Verfahren hat sich in der Mustererkennung besonders bewährt.

Der Ablauf des Fuzzy-C-Means-Algorithmus besteht aus folgenden Schritten<sup>954</sup>:

*Schritt 1)*

Man gibt die Clusterzahl  $c$  vor, wobei  $c$  größer gleich zwei und kleiner gleich der Datenobjektzahl  $J$  sein muss. Weiterhin müssen die Zugehörigkeiten  $u_{i,j}$  des  $j$ -ten Objekts  $X_j$  zum  $i$ -ten Cluster zur Initialisierung des Verfahrens angegeben werden. Der Merkmalsvektor zu Objekt  $X_j$  sei mit  $x_j$  bezeichnet. Außerdem muss der Parameter  $m$  (Fuzzifier) für die Gleichungen a) und b) bestimmt werden.

*Schritt 2)*

Mit der Gleichung a) werden aus den vorgegebenen Zugehörigkeiten  $u_{i,j}$  die Schwerpunkte  $v_i$  der Cluster bestimmt.

---

<sup>952</sup> Wenn man die Clusteranzahl  $c$  für die Ausgangspartition angibt, dann liefert Hard-C-Means Algorithmus eine Endpartition mit dieser anfänglichen Clusteranzahl. Hingegen bestimmt Isodata-Algorithmus selbst die Clusteranzahl der Endpartition. Somit ist Isodata-Algorithmus flexibler als Hard-C-Means-Algorithmus.

<sup>953</sup> Vgl. Höppner/Klawonn/Kruse (1997), S. 35.

<sup>954</sup> Vgl. Zimmermann (1995), S. 41ff.

$$a) v_i = \frac{\sum_{j=1}^J (u_{i,j})^m \cdot x_j}{\sum_{j=1}^J (u_{i,j})^m}, \quad \forall i = 1, \dots, c$$

Schritt 3)

Aus den neu berechneten Clusterzentren  $v_i$  werden mit b) die neuen Zugehörigkeiten ermittelt.

$$b) u_{i,j} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{i,j}}{d_{k,j}}\right)^{\frac{2}{m-1}}}, \quad \begin{array}{l} \forall i = 1, \dots, c \\ \forall j = 1, \dots, J \end{array}$$

Dabei ist  $d_{i,j}$  der Abstand des Merkmalsvektors  $x_j$  zum Zentrum des Clusters  $i$ .

Schritt 4)

Die neue Zugehörigkeitsmatrix  $U(t+1)$ , deren Elemente die neuen Zugehörigkeitswerte  $u_{i,j}$  sind, wird mit der alten Matrix  $U(t)$  verglichen und dadurch der Abbruch des Verfahrens gesteuert. Ist für eine geeignet gewählte Matrixnorm der Abstand dieser Matrizen kleiner als ein vorgegebener Wert  $e$ , so stoppt der Algorithmus. Im anderen Fall wird, beginnend bei Schritt 2, ein neuer Iterationszyklus gestartet.

$$c) \|U^{(t+1)} - U^{(t)}\| \leq e$$

Um die Konvergenzeigenschaft des Algorithmus sicherzustellen, müssen zwei Nebenbedingungen erfüllt sein, d. h.

$$d) \sum_{i=1}^c u_{i,j} = 1, \quad \forall j = 1, \dots, J$$

und für die Zugehörigkeitswerte muss gelten:

$$u_{i,j} \in [0,1] \text{ für alle } i=1, \dots, c \text{ und } j=1, \dots, J.$$

Der FCM-Algorithmus liefert neben der Lage der Clustermittelpunkte mit Hilfe von b)

Zugehörigkeitswerte der einzelnen Datenobjekte zu den verschiedenen Clustern. Diese Resultate können dann als Grundlage einer Klassifikation neuer, unbekannter Objekte dienen.

Das Ergebnis des Algorithmus kann mit zwei Parametern beeinflusst werden: Das Abbruchkriterium bestimmt die Zahl der Iterationen und damit verbunden die Genauigkeit des Ergebnisses. Die Zahl der Iterationen und damit die Rechenzeit steigen mit der Verkleinerung von  $\epsilon$  spürbar an. Der Parameter  $m$  in den Exponenten der Gleichungen a) und b) bestimmt den Grad der Unschärfe der Resultate. Für  $\epsilon \rightarrow 1$  nähert man sich dem scharfen Clusterergebnis, das auch der klassische Isodata liefert. Für  $m \rightarrow \infty$  streben die Zugehörigkeitswerte  $u_{i,j}$  aller Objekte gegen den reziproken Wert  $1/c$  der Clusterzahl.

Ein Nachteil des Verfahrens ist, dass bei Verwendung der euklidischen Distanz nur bei einer kreis- bzw. kugelförmigen Struktur der Cluster sinnvolle Ergebnisse erzielt werden. Langgestreckte Klassen werden sonst durch diesen Algorithmus nicht erkannt.<sup>955</sup> Liegen solche Clusterformen vor, empfiehlt sich die Verwendung verallgemeinerter Distanzmaße oder Cluster-Algorithmen (z. B. Gustafson-Kessel-Algorithmus mit Mahalanobis Distanz<sup>956</sup>).

### **6.6.10 Was kann mit Hilfe der Fuzzy-Datenanalyse im CRM erreicht werden?**

Der Ansatz der Fuzzy-Logik und seine unscharfe Mengentheorie finden ihren Einsatz in vielen Bereichen der Wissenschaft und können erfolgreich verwendet werden.

Eine Möglichkeit, die auch für diese Arbeit besonders relevant ist, stellt aber die Verwendung der Fuzzy-Logik bei der Analyse von umfangreichen Kundeninformationen, die meistens in Form von Daten in einer Datenbank gespeichert sind und für diverse Zwecke zur Verfügung stehen.

Aufgrund des großen Konkurrenzdrucks am Markt stellt sich bei vielen Unternehmen die Notwendigkeit einer gezielten Kundenansprache und besseren Kundenorientierung. Die Problemstellung einer unscharfen Kundensegmentierung ist daher für viele

---

<sup>955</sup> Vgl. Zimmermann (1995), S. 43.

<sup>956</sup> Der FCM-Algorithmus geht davon aus, dass alle Cluster ungefähr gleich groß und (hyper-)kugelförmig sind. Durch die Änderung der verwendeten Metrik (Distanzmaß) beim Fuzzy C-Means können statt kugelförmigen Clustern auch (hyper-)ellipsoide Cluster erkannt werden. Durch die Verwendung der Mahalanobis-Distanz anstelle der euklidischen Distanz und dem GK-Algorithmus anstelle des FCM-Algorithmus ist es möglich, die Form der Cluster zu beschreiben und damit nicht nur etwa gleich große sondern auch unterschiedlich große (Hyper-)Ellipsoiden zu erkennen, vgl. Höppner/Klawonn/Kruse (1997), S. 41-48. Auf diese Weise wird die Form der Cluster in der Clusterbildung mit berücksichtigt.

Unternehmen von großer Relevanz. So ist z. B. bei großen Versandhäusern und Retail-Unternehmen die Identifizierung von Zielgruppen ein wichtiges Erfolgskriterium. Versandhäuser besitzen große Datenbanken mit umfangreichen Informationen über ihre Kunden. Eine unscharfe Kundensegmentierung in dieser Branche sowie auch bei Retail-Unternehmen könnte für ein gezieltes und am Kundenpotential orientiertes Database-Marketing und Customer Relationship Management genutzt werden. Das Database-Marketing und Customer Relationship Management wurden bereits in vorigen Kapitel detailliert beschrieben und erläutert, wie diese dem Erfolg eines Unternehmens beitragen, indem sie dem Unternehmen die Möglichkeit bieten, seine Kunden gezielt anzusprechen. Dazu soll zunächst eine Kundenanalyse durchgeführt werden, um die Zielgruppen zu identifizieren und zu segmentieren. Der Vorteil einer unscharfen Kundensegmentierung liegt darin, dass die Kunden nicht scharf einem Segment zugeordnet werden. Hingegen werden die Kunden so klassifiziert, dass sie nicht eindeutig einem Segment, sondern mit verschiedenen Zugehörigkeitswerten zu unterschiedlichen Segmenten gehören. Dementsprechend können alle Kunden im Rahmen des CRM durch ein individuelles Produkt bzw. Dienstleistungsangebot adäquat angesprochen werden. Die unscharfe Segmentierung von Kunden weist ein hohes Potential in der Kundenanalyse auf. Denn die scharfe Zuordnung nur einem Segment hat den Nachteil, dass ein Teil der Informationen verloren geht, da eine eindeutige Zuordnung oder Nicht-Zuordnung nicht immer gegeben ist.

Nachdem bereits vorhandene Kunden klassifiziert und Zielsegmente gefunden werden, können die Klassifikations-Regeln für jeweils diese Segmente erzeugt werden. Mit Hilfe dieser aufgestellten Regeln kann man dann leicht feststellen, welchem Segment ein neu gewonnener Kunde zuzuordnen und dementsprechend wie er anzusprechen ist.

Weiterhin bietet sich die Kundensegmentierung dazu an, Tendenzen in der Bewegung der Kunden zwischen den Segmenten aufzudecken, indem in regelmäßigen Abständen die Klassifikation der Kunden wiederholt wird. Dadurch lässt sich früh erkennen, wann die Kunden-Bearbeitungsstrategie geändert werden sollte.

# KAPITEL 7

---

## **FUZZY-LOGIK BASIERTE SEGMENTIERUNG VON KUNDEN FÜR CRM**

### **7.1 EINFÜHRUNG**

Für eine effiziente Durchführung von Marketing-Aktionen im Rahmen des erfolgreichen Customer Relationship Management ist es von großer Bedeutung, gezielt nur die Kunden anzusprechen, für die das betreffende Produkt von Relevanz ist. Je spezifischer die Kunden ausgewählt werden, desto höher werden die Responsraten und desto größer auch der Erfolg der Marketing-Aktion. Die grundlegende Voraussetzung für eine zielgruppengerechte Kundenansprache ist eine differenzierte Betrachtung des vorhandenen Kundenstamms. Dies verlangt eine Segmentierung der Kunden in Gruppen, die sich hinsichtlich ihrer Struktur und ihres Verhaltens möglichst deutlich voneinander unterscheiden.

### **7.2 METHODOLOGIE**

In diesem praktischen Teil meiner Diplomarbeit wird eine Anwendung von Fuzzy-Clustering-Verfahren an einem Fallbeispiel aus dem betriebswirtschaftlichen Bereich dargestellt. Das Ziel dieses Fallbeispiels ist es, einen Überblick über das mögliche Anwendungsgebiet des Fuzzy-Clustering zu geben und dabei einen Eindruck von der Arbeitsweise des Fuzzy-C-Means-Algorithmus zu vermitteln.

Bei konventionellen Clusterverfahren ist es häufig problematisch, Objekte (z. B. Kunden) eindeutig einem der Cluster zuzuordnen.<sup>957</sup> Eine unscharfe Clusteranalyse dagegen ordnet die Kunden nicht scharf einem Cluster zu, sondern ermittelt den Zugehörigkeitsgrad zu jeder der gebildeten Cluster (Segmente/Klassen).

---

<sup>957</sup> Vgl. Zimmermann (1995), S. 143.

Dabei stellen die Segmentierungskriterien einen großen Problembereich dar, da man im Vorfeld der durchgeführten Analyse nicht weiß, welche Merkmale zur Bildung trennscharfer Segmente besonders beitragen und somit zur Abgrenzung gut geeignet sind. Dieses Wissen entsteht erst im Verlauf der Kundensegmentierung, wenn die Merkmalsausprägungen jedes Kunden im multidimensionalen Raum mit denen anderer Kunden verglichen werden.

Nach einer Einführung in die Problemstellung der Kundensegmentierung bei Großhändlern und der Vorstellung des verwendeten Datenmaterials erfolgt die Beschreibung der Vorverarbeitung dieser Daten. Danach wird die Vorgehensweise zur Clusterbildung (Bildung der Kundensegmente) mit dem Fuzzy-C-Means-Algorithmus anhand des in Java selbst implementierten Programms dargestellt. Abschließend werden die gefundenen Kundentypen als Ergebnis der Kundenanalyse fallspezifisch interpretiert sowie ihre Eigenschaften beschrieben.

### **7.3 PROBLEMSTELLUNG**

Im wettbewerbsspolitischen Umfeld der Großunternehmen gab es in den letzten Jahren grundlegende Veränderungen. Verantwortlich hierfür war in erster Linie die Globalisierung und Transparenz der Verkaufsmärkte. Aufgrund des sich wandelnden Marktumfeldes sowie der Transparenz und großen Nachfrage für Produkte und Dienstleistungen wurde die Konkurrenz zwischen Unternehmen größer. Die Produkte und Leistungen von Unternehmen wurden austauschbar, wodurch Unternehmen ihre Wettbewerbskraft und damit ihre signifikante Position im Markt verloren.

Außerdem können noch mehr Unternehmen als zuvor im Markt ihre Produkte und Dienstleistungen online oder offline zu unterschiedlichen Preisen und Konditionen anbieten. Dies führt selbstverständlich dazu, dass Unternehmen auf der einen Seite viele Kunden erreichen können, auf der anderen Seite jedoch ihre Kunden wegen der leichten Substituierbarkeit ihrer Produkte und Dienstleistungen verlieren.

Vor dem Hintergrund dieser Entwicklungen auf dem Markt wurde die Forderung nach zielgruppengerechter Kundenorientierung bei den Unternehmen zur wichtigsten Geschäftsstrategie. Die zentrale Bedeutung der Kunden und die Pflege ihrer Geschäftsbeziehung wurden so auch im Retail-Geschäft erkannt. Dementsprechend sind die Erfüllung der Bedürfnisse jedes einzelnen Kunden durch die für ihn „passenden“ Produkte bzw. Leistungen die Voraussetzung für eine gute Geschäftsbeziehung des Retail-Unternehmens zu diesem Kunden.

Die Problemstellung, mit der sich dieses Fallbeispiel befasst, besteht darin, aus über

einem begrenzten Betrachtungszeitraum (online oder offline) fallenden Geschäftsdaten bezüglich der Kunden zunächst die Informationen herauszufiltern, die für eine Kundensegmentierung relevant sein könnten. (das heißt, die wesentlichen Parameter des zu entwerfenden Analysemodells wie z. B. Gesamtumsatzniveau und Kauffrequenz des Kunden, Deckungsbeitrag u. a.). Daran anschließend stellt sich die Aufgabe, aus diesen unterschiedlichen Kundendaten bestimmte typische Kundengruppen zu ermitteln.

Die Gruppierung/Segmentierung wird mit Hilfe des Fuzzy-C-Means-Algorithmus durchgeführt. Dabei werden aus einer großen Anzahl von Kunden möglichst homogene Kundensegmente gebildet, innerhalb derer die Kunden ähnliche Bedürfnisse aufweisen. Zudem sollte jedes Segment gegenüber den anderen Segmenten möglichst heterogen sein. Diese vom Fuzzy-C-Means-Verfahren gefundenen Kundensegmente können dann beispielsweise zielgruppenorientiert mit entsprechenden Marketingaktivitäten bearbeitet bzw. behandelt werden.

## **7.4 ANFORDERUNGEN**

### TECHNISCHE ANFORDERUNGEN

1. *Die Konfigurierbarkeit:*

Sie umfasst die Datenbank bzw. die Datenfelder, das heißt, es soll konfigurierbar sein, welche Daten wir segmentieren bzw. clustern wollen. Außerdem soll die Clusteranzahl konfiguriert werden können.

2. *Die Bestimmung und Berechnung der Cluster durch Fuzzy-C-Means Algorithmus.*

3. *Die Speicherung der Clusteringergebnisse.*

4. *Die Zuordnung der Kunden zu den entsprechenden Kundensegmenten und die Speicherung der einzelnen Ergebnisse der Clusteranalyse mit verschiedenen Konfigurationen.* Die Speicherung der Ergebnisse dient dazu, diese zu einem späteren Zeitpunkt wieder aufrufen und nochmals untersuchen zu können, falls es notwendig wird.

5. *Die Benutzeroberfläche GUI:* Über der Benutzeroberfläche soll es möglich sein, die Konfiguration (Parameter wie z. B. Dimensionanzahl, Clusteranzahl, Unschärfe, Schwellenwert u. a.) einzugeben bzw. die berechneten Segmente (Clusterzentren) abzufragen.

6. Die *graduelle Zugehörigkeit jedes einzelnen Datenobjekts (Kunden) zu verschiedenen Segmenten (Kundengruppen)* sollen in der GUI angezeigt werden.

## ANFORDERUNGEN FÜR DIE ANALYSE IM UNTERSUCHUNGSUMFELD

### ➤ FESTLEGUNG DES UNTERSUCHUNGSTYPUS

Mit der Festlegung des Untersuchungsdesigns werden die wesentlichen Entscheidungen über die anzuwendenden Forschungsmethoden getroffen, indem die Art der Datenerhebung bestimmt wird. Häufig stellt sich aber die Frage, ob es überhaupt notwendig ist, für einen bestimmten Untersuchungszweck im CRM die Daten neu zu erheben und auszuwerten. In manchen Fällen könnte es ausreichen, vorhandene Daten im Hinblick auf das aktuelle Problem neu zu analysieren.

Grundsätzlich unterscheidet man zwischen Primär- und Sekundärforschung:<sup>958</sup> Als Primärforschung bezeichnet man die Neu-Erhebung von Daten für ein anstehendes Untersuchungsproblem („Field Research“). Dagegen ist die Sekundärforschung dadurch gekennzeichnet, dass bereits erhobene und gespeicherte Daten für einen gegebenen Untersuchungszweck neu aufbereitet und analysiert werden. Da auf diese Weise die Untersuchungen am PC durchgeführt werden, findet man in diesem Zusammenhang auch den Begriff „Desk Research“.

Unter dem Gesichtspunkt der Datenbeschaffungsart entspricht die vorliegende Analyse eher der Sekundärforschung, da wir für dieses Fallbeispiel den Kundenstamm selbst aufgebaut haben und seine Struktur untersuchen möchten. Wir nehmen an, dass die Kunden bereits vorhanden und ihre Daten in einer Datenbank abgelegt sind. Die Ergebnisse der Kundensegmentierung basieren auf diesen fiktiv gebildeten Daten und könnten daher nicht den realen Gegebenheiten entsprechen. Schließlich stellen die Ergebnisse in diesem Fallbeispiel nicht das Ziel, sondern den Zweck der Analyse dar bei. Typische Fragestellungen bei dieser sekundären Kundenanalyse sind z. B.:

- Welche Segmente/Gruppen existieren in dem Kundenstamm des Unternehmens?

---

<sup>958</sup> Vgl. Neckel/Knobloch (2005), S. 91.

- Welche Zielgruppen sind besonders interessant?
- Wer sind die besten und schlechtesten Kunden?
- Wo liegen die Stärken und Schwächen der Kunden, die zu verschiedenen Segmenten gehören.

➤ DATENMATERIAL

Die selbst erschafften Kundendaten eines Retail-Unternehmens bilden die Grundlage für die Kundensegmentierung. Diese Kundendaten werden in einem relationalen Datenbanksystem (für die Herstellung des Database-Systems wird Toad for MySQL Version 2.0 verwendet) abgelegt. Die Customer Datentabelle enthält somit 200 Kundendatensätze. Diese Datensätze beinhalten alle Informationen, die laut unserem Fallbeispiel im Kundeninformationssystem des untersuchten Retail-Unternehmens über die Kunden zur Verfügung stehen, wie z. B. Name, Deckungsbeitrag, Gesamtanzahl der Produktankäufe, Gesamtumsatz, Recency (seit dem letzten Kauf vergangene Zeit) und Frequency (Kauffrequenz/Kaufhäufigkeit pro Zeiteinheit) des betreffenden Kunden.

➤ DATENVORVERARBEITUNG

Aus der Vielzahl der einen Kauffall beschreibenden möglichen Merkmale müssen im ersten Schritt die für die Kundensegmentierung relevanten Merkmale ausgewählt werden. In der Untersuchung, die dieses Fallbeispiel beschreibt, wurden folgende Merkmale (mit zugehöriger Maßeinheit) der Kunden als Segmentierungskriterien selektiert:

1. *Gesamtumsatz für bis 18 Monate* (kumulierter Wert in €)
2. *Recency*: Seit dem letzten Kauf vergangene Zeit (in Monaten zwischen 1-18). Je höher dieser Wert ist, desto schlechter ist der Kunde im wirtschaftlichen Sinne. Es herrscht also indirekte Proportionalität.
3. *Frequency*: Kauffrequenz; Kaufhäufigkeit (zwischen 1-5 mal). Je öfter ein Kunde gekauft hat, desto besser ist er zu klassifizieren. Hier herrscht direkte Proportionalität.
4. *Der kundenindividuelle Deckungsbeitrag*: Der Deckungsbeitrag (DB) ist die Differenz zwischen den erzielten Erlösen und den variablen Kosten.

Es handelt sich also um den Betrag, der zur Deckung der Fixkosten zur Verfügung steht. Der Deckungsbeitrag kann sowohl positiv als auch negativ sein. Den individuellen Deckungsbeitrag erhält man normalerweise, indem man von Gesamterlösen (= Bruttoerlöse - Nettoerlöse) des Kunden diesem Kunden zurechenbare Marketingkosten, Verkaufskosten und Service- und Transportkosten subtrahiert.

5. Gesamtanzahl der Produktankäufe (zwischen 1-10 mal)

Diese Merkmale/Dimensionen wurden bei unserem Fallbeispiel als relevant für eine Kundensegmentierung erkannt. Die ersten 3 Merkmale sind auch die Faktoren der RFMR-Methode<sup>959</sup> zur Kundenbewertung (siehe dafür auch Kapitel 4.6.2). Noch wurden individueller DB und Produktankäufe als zwei weitere wichtige Faktoren/Merkmale für die Kundensegmentierung gewählt. Die Kunden werden mit diesen 5 Merkmalen beschrieben, auf deren Grundlage die Segmentierung erfolgen soll (Es können noch weitere Merkmale für die Segmentierung hinzugefügt werden). Das folgende Bild 1 zeigt einen Ausschnitt aus der Kundendatentabelle.

Kundennummer	Deckungsbeitrag	Vorname	Nachname	Umsatz	Ankaeufe	Age	Frequency	Recency
809	-15			1300	6	{null}	4	2
810	-3			190	1	{null}	1	4
811	85			1300	7	{null}	5	8
812	6			200	1	{null}	1	4
813	265			4200	5	{null}	4	3
814	-26			1290	7	{null}	3	9
815	20			240	2	{null}	2	5
816	58			700	1	{null}	1	10
817	20			350	2	{null}	1	5
818	165			2550	7	{null}	3	10
819	75			1450	2	{null}	1	12
820	100			1700	4	{null}	2	13
821	20			300	2	{null}	1	6
822	-25			2250	6	{null}	4	14
823	170			2600	7	{null}	4	2
824	11			230	1	{null}	1	3
825	-50			1950	3	{null}	2	12
826	40			490	1	{null}	1	5
827	68			710	3	{null}	1	14
828	10			200	2	{null}	1	3
829	-55			2700	8	{null}	5	15
830	28			800	1	{null}	1	15
831	7			180	2	{null}	1	4
832	90			1450	5	{null}	5	11
833	190			2700	7	{null}	3	3
834	110			1570	2	{null}	2	17
835	-17			110	3	{null}	1	4
836	-75			1700	6	{null}	5	5

Bild 1: Datentabelle für die Kundensegmentierung

<sup>959</sup> Für die Analyse werden die Faktoren der RFMR-Methode gewählt, da seine Faktoren als Dimensionen für die Kundensegmentierung sinnvoll benützt werden können.

Wie bereits erwähnt, enthält die im Bild 1 dargestellte Datentabelle 200 Kundendatensätze. Diese bilden die Grundlage für die Clusteranalyse mit dem Fuzzy-C-Means-Algorithmus. Die zur Verfügung stehenden Informationen zu einem Kunden sind in einer Zeile dargestellt.

Um die Clusteranalyse nicht zu verzerren, muss sichergestellt werden, dass die ausgewählten Segmentierungsmerkmale (DB, Umsatz, Ankäufe, Frequency und Recency) weitgehend voneinander unabhängig sind. Mit Hilfe einer Korrelationsanalyse kann man feststellen, in wie weit diese Merkmalpaare (z. B. DB-Umsatz; DB-Ankäufe; DB-Frequency; Ankäufe-Frequency usw.) voneinander abhängig sind.

Die Datentabelle kann aus dem Database-System in MS-Excel-Format exportiert werden. Mit der mathematischen Korrelation-Funktion von MS-Excel kann dann die Anhängigkeit zwischen Merkmalpaaren leicht feststellen. Diese Funktion gibt den Korrelationskoeffizienten als Maß für die Stärke der Abhängigkeit zwischen jeweils zwei Merkmalen an. Wenn man die Funktion auf jedes Merkmalpaar anwendet, ergibt sich eine 5×5-Korrelationsmatrix. Der Korrelationskoeffizient kann Werte zwischen 1 und -1 annehmen, wobei Werte um Null auf Unkorreliertheit und Werte nahe 1 bzw. -1 auf positive bzw. negative Abhängigkeit hinweisen.

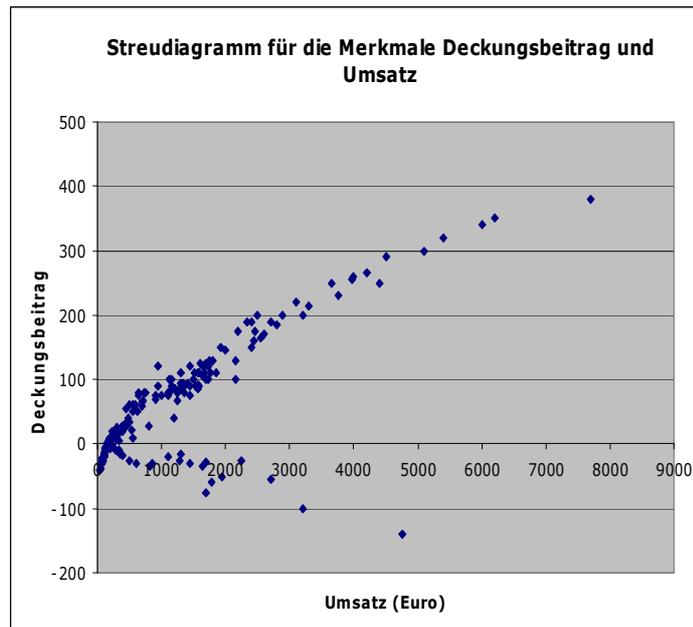
Die Korrelationsanalyse für 200 Kundendatensätze mit MS-Excel ergibt folgende Korrelationsmatrix:

**KORRELATIONSMATRIX**

	<b>DB</b>	<b>UMS</b>	<b>ANK</b>	<b>FR</b>	<b>RC</b>
<b>DB</b>	1	0,79245966	0,433069	0,36178825	-0,16612415
<b>UMS</b>	0,79245966	1	0,61021325	0,51097339	-0,15342217
<b>ANK</b>	0,433069	0,61021325	1	0,86290321	-0,18237007
<b>FR</b>	0,36178825	0,51097339	0,86290321	1	-0,18737272
<b>RC</b>	-0,16612415	-0,15342217	-0,18237007	-0,18737272	1

Wie aus der Matrix erkennbar ist, hat die höchsten Korrelationskoeffizienten die Werte 0,79245966 und 0,86290321 für die Merkmalpaare Deckungsbeitrag-Umsatz und Ankäufe-Frequency. Diese Werte deuten auf eine starke Abhängigkeit zwischen den betreffenden Merkmalpaaren. Diese Abhängigkeiten

können wie folgt interpretiert werden: Außer Ausnahmesituationen<sup>960</sup> verursachen hohe Umsätze hohe Deckungsbeiträge für Kunden. Obwohl die Deckungsbeiträge i. d. R. mit zunehmendem Umsatz steigen, soll man sich auf diese steigende Tendenz nicht verlassen. Denn Kunden mit hohen Umsätzen können sowohl hochprofitabel als auch verlustbringend sein. Solche Kunden können um ein Vielfaches mehr an Verlust erwirtschaften als Kunden mit niedrigen Umsätzen (siehe Abb. 63) Daher spielt das Merkmal Deckungsbeitrag (trotz der starken Korrelation zwischen DB und UMS) eine unterstützende Rolle in der Clusteranalyse.



**Abbildung 63: Streudiagramm DB-Umsatz für 200 Datenobjekte**

Quelle: eigene Darstellung

Die Merkmale Ankäufe und Frequency korrelieren sich stark miteinander. Also je oft man kauft, desto mehr wird selbstverständlich die Gesamtsumme seiner Ankäufe für einen Betrachtungszeitraum. Es herrscht also eine starke Korrelation (d. h. diese Merkmale drücken fast Ähnliches aus) zwischen diesen Merkmalen. Eines von beiden Merkmalen soll aus dem Merkmalkomplex eliminiert werden. Da die Kaufhäufigkeit/Frequency für die langfristige Kundenbindung wichtiger als die Gesamtankäufe ist, kann man sich dafür entscheiden, das Merkmal Ankäufe nicht in die Clusteranalyse mit einzubeziehen. Die Werte der Korrelationskoeffizienten für andere Merkmale deuten auf keine signifikanten Abhängigkeiten zwischen Merkmalpaaren hin (die absoluten Werte der Korrelationskoeffizienten sind kleiner). So bleiben nur

<sup>960</sup> siehe dazu: Neckel/Knobloch (2005), S. 205f.

noch 4 Merkmale übrig, mit denen die Clusteranalyse durchgeführt werden soll. Die Unabhängigkeit stellt eine gleiche Gewichtung der ausgewählten Segmentierungsmerkmale in der Clusteranalyse sicher.

Weiterhin sind die in der Clusteranalyse zu untersuchenden Daten dimensionsbehaftet und die Merkmalskalen (alle sind metrisch skaliert) weisen unterschiedliche, mehr oder weniger willkürlich gewählte Wertebereiche auf (für den Fallbeispiel in dieser Arbeit: Umsatz: 0-7700 €; DB: -140-380; Frequency: 1-10 Male; Recency: 1-18 Monate). Daher sind sie kaum miteinander vergleichbar.

Der Fuzzy-C-Means-Algorithmus verwendet die euklidische Distanz als Abstandsmaß, um die geometrische Entfernung als Differenz zweier metrisch

skalierter Daten zu berechnen: 
$$d(x, k) = \left( \sum_{i=1}^p |x_i - k_i|^r \right)^{\frac{1}{r}}$$

Bei der Berechnung der euklidischen Distanz werden durch die Quadratisierung die großen Differenzwerte überproportionaler gewichtet, während geringere Distanzen ein kleineres Gewicht bekommen. Um diese unterschiedliche Gewichtung auszuschalten, müssen die Werte der Merkmalsausprägungen vergleichbar gemacht werden.

Das ist aber schwierig, wenn die Verteilungen der Werte unterschiedliche Zentren und Streuungen, die zu vergleichenden Merkmalen also unterschiedliche „arithmetische Mittel“ und „Standardabweichungen“ aufweisen. Durch eine sogenannte  $z$ -Transformation können die ursprünglichen Werte der Merkmalsausprägungen so umgerechnet werden, dass alle zu untersuchenden Merkmalen ein arithmetisches Mittel  $\mu=0$  und eine Standardabweichung  $\sigma=1$  haben, also vergleichbar sind.<sup>961</sup> Man sagt daher auch, dass die Merkmalen mit Hilfe einer  $z$ -Transformation „standardisiert“ wurden.

Konkret gibt der  $z$ -transformierte Merkmalswert eines Datums an, um wie viele Standardabweichungen der ursprüngliche Merkmalswert dieses Datums unterhalb oder oberhalb des arithmetischen Mittels für dieses Merkmal liegt. Die Verteilung der  $z$ -Werte entspricht dann einer Standardnormalverteilung.

---

<sup>961</sup> Vgl. Andreß (2001), Begriff:  $z$ -Transformation.

Durch diese Standardisierung/Normierung werden die Daten auf einen einheitlichen Wertebereich transformiert.<sup>962</sup>

Für k-tes Merkmal lautet die z-Transformation (Mittelwertsubtraktion und Division durch die Standardabweichung):

$$z_{i,k} = \frac{x_{i,k} - \bar{x}_k}{s_r} \quad i = 1, \dots, n \text{ und } k = 1, \dots, p$$

mit  $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{i,k}$  und  $s_r = \left( \frac{1}{n} \sum_{i=1}^n (x_{i,k} - \bar{x}_k)^2 \right)^{\frac{1}{r}}$ . Dabei ist  $r=1$  für die

Manhattan-Metrik und  $r=2$  für die euklidische Distanz. Außerdem ist  $\bar{x}_k$  das arithmetische Mittel und  $s_r$  die Standardabweichung des Merkmals k (das normalisierte bzw. standardisierte Merkmal k hat den Mittelwert  $\mu=0$  und die Standardabweichung  $\sigma=1$ ).

Die Daten auf den gemeinsamen Wertebereich zu transformieren und dann die Clusteranalyse durchzuführen ist sinnvoll und sicherlich die richtige Vorgehensweise. Das Problem dabei ist jedoch, dass die Darstellung der Daten nach deren Transformation weniger anschaulich bzw. nachvollziehbar wird. Dementsprechend wird in der folgenden Clusteranalyse auf eine derartige Transformation von Daten verzichtet und die Clusteranalyse ohne Normierung durchgeführt, um die Clusterergebnisse einfach interpretieren zu können.

<sup>962</sup> Eine Clusteranalyse von Kunden anhand der Merkmale „Alter“ und „Einkommen“ könnte (bei Verwendung eines euklidischen Distanzmaßes) z. B. aufgrund der Unterschiede in der Standardabweichung der beiden Merkmale vollständig von den Einkommensunterschieden dominiert werden. Durch die Verwendung eines Normierungsverfahren wie z. B. Z-Transformation werden die Merkmale „Alter“ und „Einkommen“ mit Mittelwert  $\mu=0$  und Standardabweichung  $\sigma=1$  in transformierte Merkmale überführt, die nun gleichgewichtig in die Clusteranalyse eingehen können, vgl. Hippner/Wilde (2001), S. 93.

## 7.5 ANALYSE und ENTWURF

### SYSTEMARCHITEKTUR

Zur Realisierung des Projektes ist eine den gestellten Anforderungen entsprechende, softwaretechnische Umgebung erforderlich, welche im folgenden Abschnitt beschrieben wird.

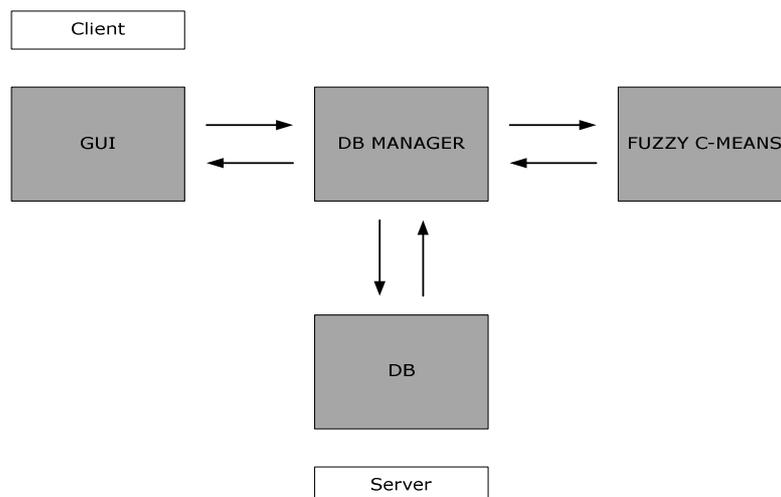
### CLIENT-SERVER-ARCHITEKTUR

Das Client-Server-System bildet eine Netzwerkstruktur, bestehend aus einem zentralen Datenbankserver als Server-Komponente und einem Benutzer-Client als Client-Komponente. Die Server-Komponente bietet Dienste für die Client-Komponente an. Die Kommunikation zwischen einem Client und dem Server erfolgt nachrichtenorientiert und wird über DB Manager geregelt. Die Nutzung der Dienste des Servers geht immer vom Client aus.

Den Client bildet im Wesentlichen das Anwendungsprogramm, über dessen Benutzerschnittstelle der Benutzer über das Netzwerk auf Ressourcen des Datenbankservers zugreift. Er liest und pflegt die Daten in der Datenbank durch „Abschicken“ von SQL-Befehlen an den Server. Das Ergebnis (Kundendaten) liefert der Server als Antwort an den Client zurück.

Fuzzy C-Means stellt das eigentliche Programm für das Clustering dar.

Die folgende Abbildung veranschaulicht die Client-Server-Architektur des für das Fallbeispiel entworfenen Prototyps:

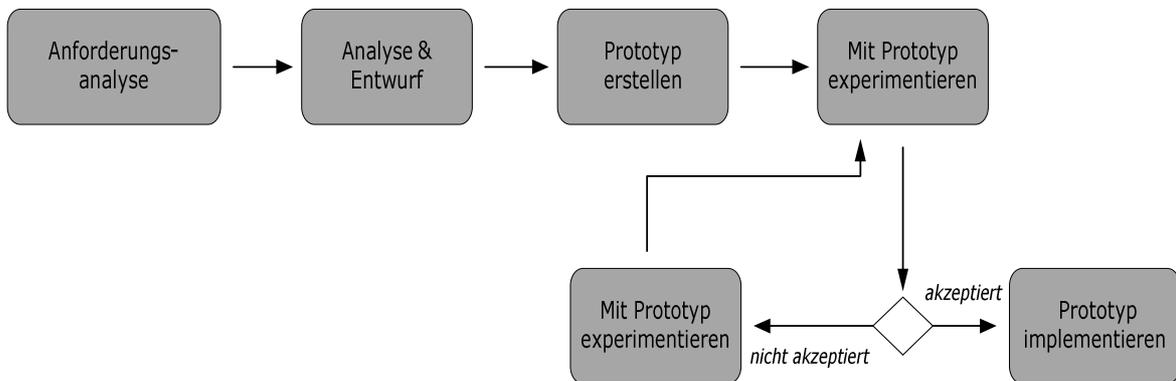


**Abbildung 64: Client-Server-Architektur**

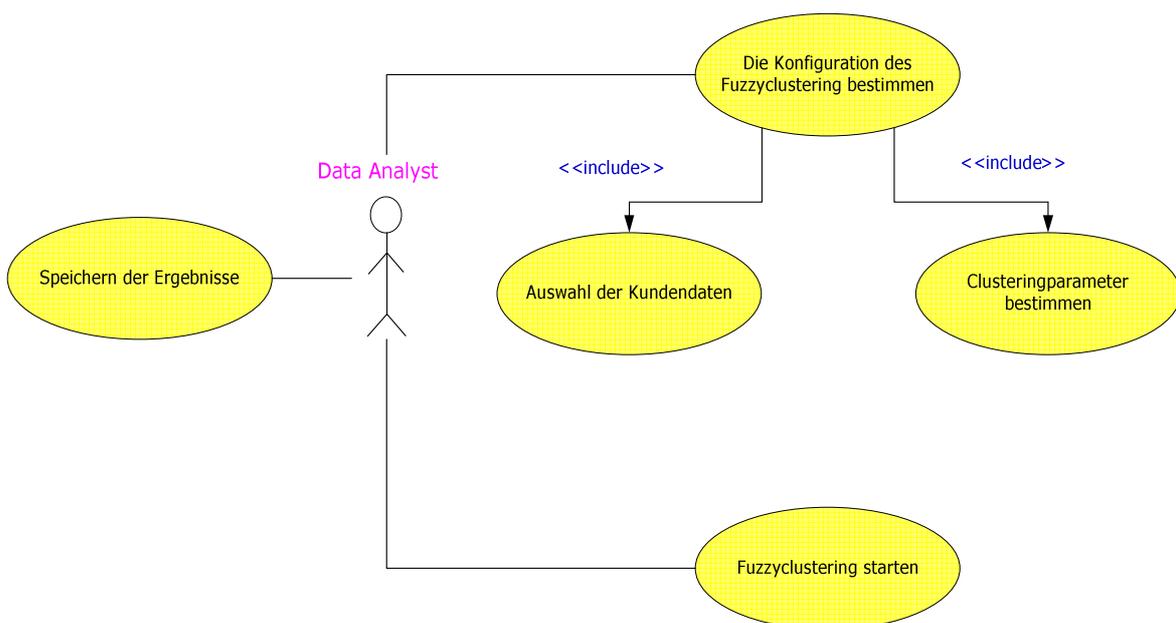
Quelle: eigene Darstellung

Mit Hilfe der Benutzeroberfläche wird über den DB-Manager die in der Datenbank gespeicherte Customer-Tabelle abgefragt. Der Benutzer wählt die notwendigen Felder der Customer-Datentabelle für die Segmentierung aus. Diese Informationen werden an den DB-Manager zugesandt. Der DB-Manager legt temporäre Tabellen in der Datenbank an (TempTables) und nimmt die Input-Daten auf. Diese werden dann dem Fuzzy C-Means Programm zur weiteren Berechnung zugeschickt. Die Input-Daten werden in Fuzzy C-Means verarbeitet und die berechneten Ergebnisse werden wieder dem DB-Manager zurückgesandt. Der DB-Manager speichert die Clusteringergebnisse in FuzzyTables in der Datenbank. Danach liefert der DB-Manager diese in FuzzyTables angelegten Clusteringergebnisse in der graphischen Benutzeroberfläche an.

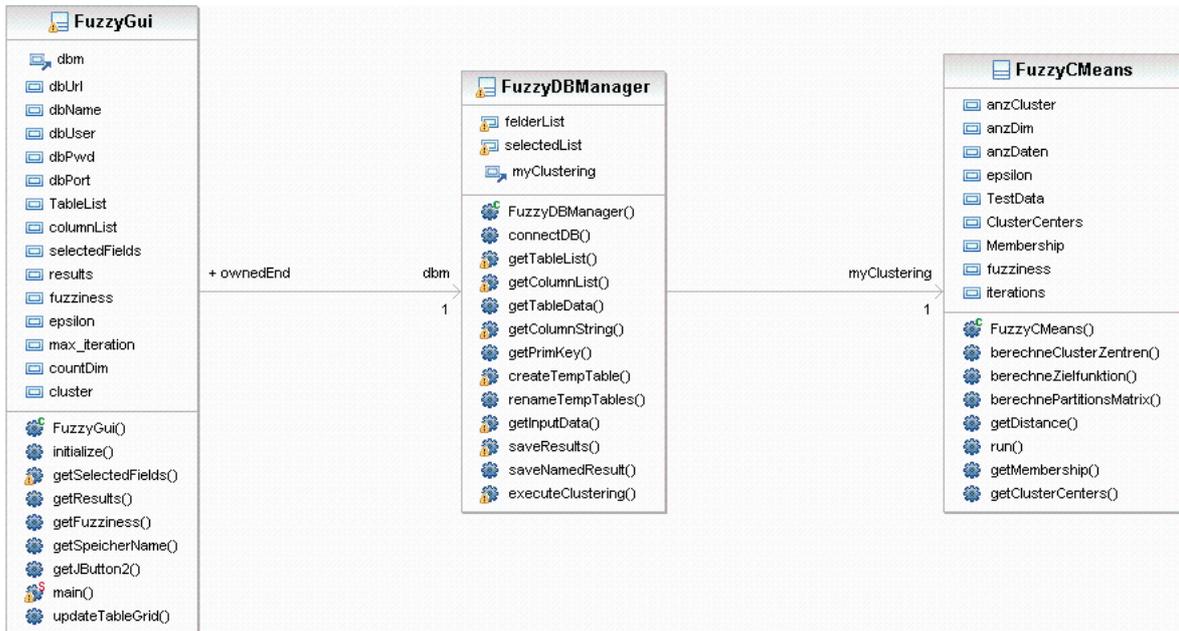
### PROTOTYP-ORIENTIERTE MODELL



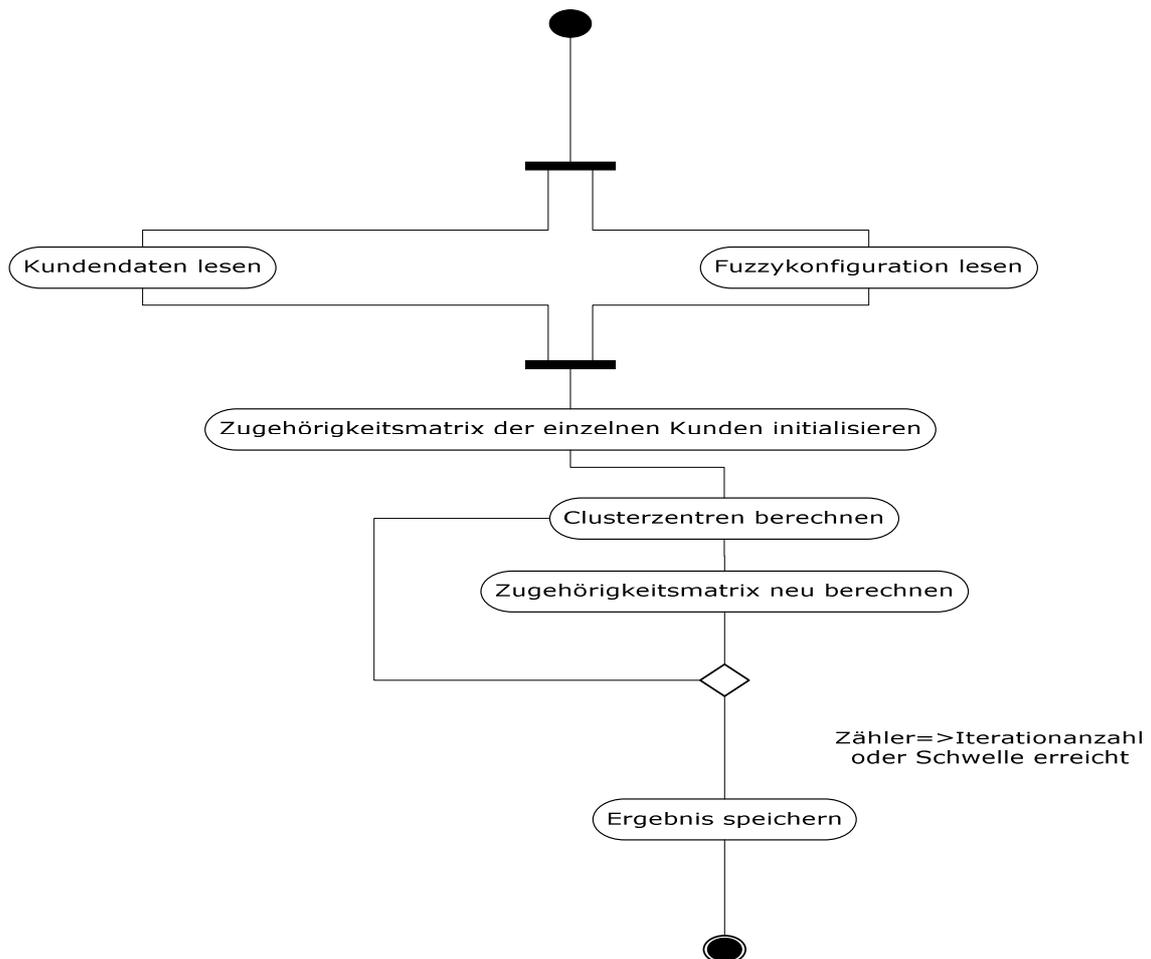
### USECASE-DIAGRAMM



## KLASSENDIAGRAMM

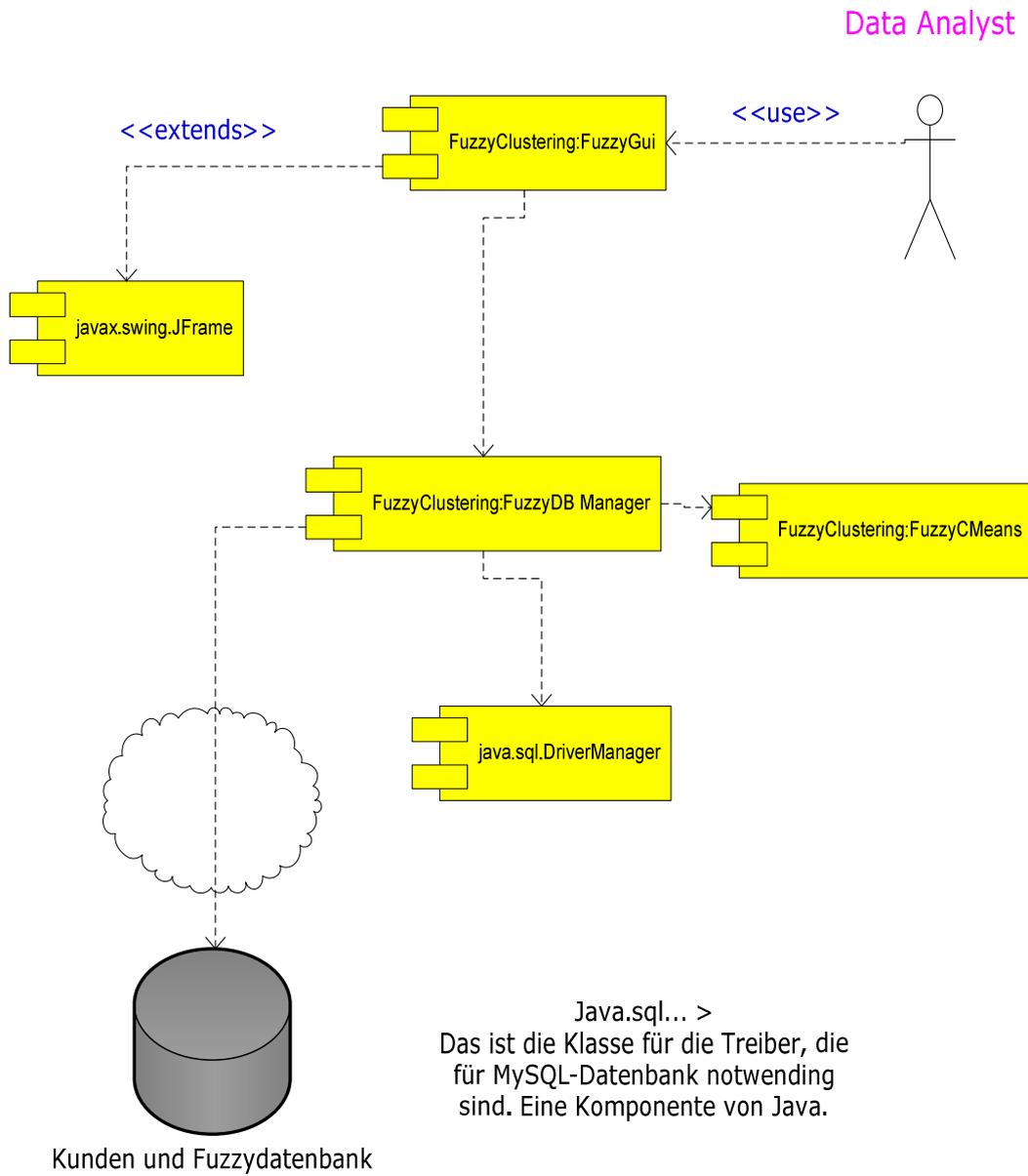


## AKTIVITÄTSDIAGRAMM



## KOMPONENTENDIAGRAMM

FuzzyGui vererbt JFrame von Swing-Komponente der Java



## 7.6 IMPLEMENTIERUNG und DEMONSTRATION

### KOMPONENTEN

➤ Benutzeroberfläche-GUI: FuzzyGui

Die Benutzeroberfläche bietet uns folgende Möglichkeiten:

- Die Auswahl der Datenbank: Parameter Url (DB-Url), Name der Datenbank (DB-Name), User, Password (Pwd).
- Die Auswahl der Tabelle (Customer Tabelle) in der Datenbank: die Datenfelder der Tabelle werden in einer Liste (Auswählbare Felder) dargestellt.
- Die Felder (Deckungsbeitrag, Frequency, Recency, Umsatz), die für das Fuzzy-Clustering zu verwenden sind, werden ausgewählt (Ausgewählte Felder).

Der Fuzzyparameter bietet uns 4 Werte manuell einzugeben:

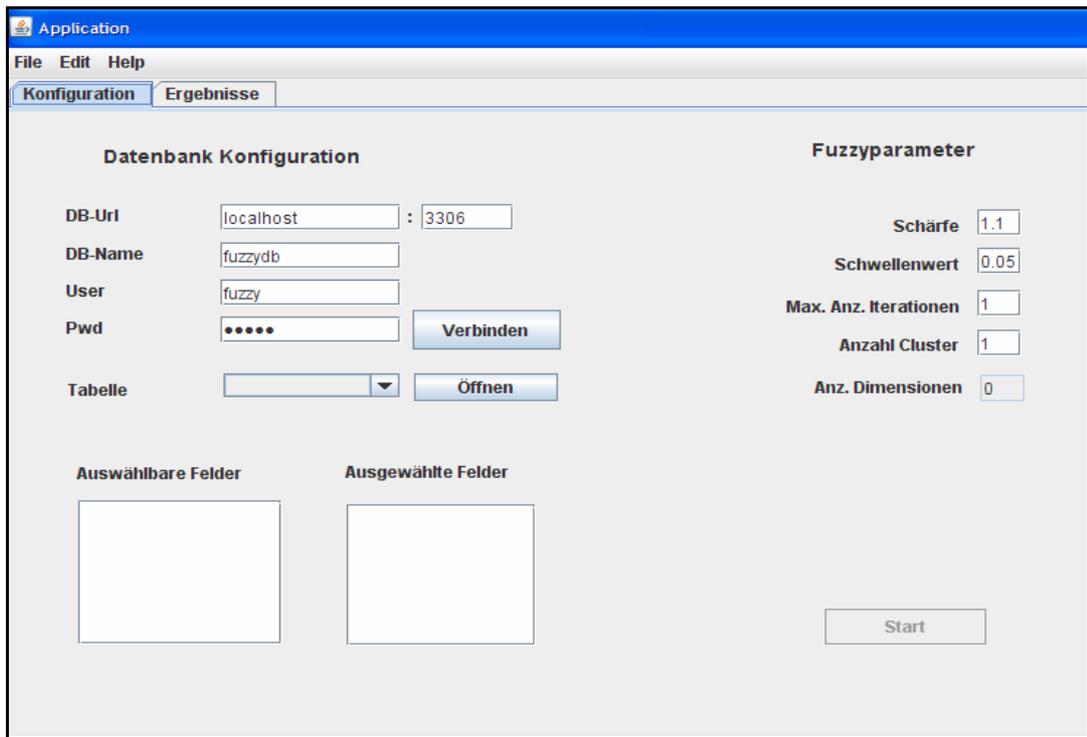
Diese Werte sind die Unschärfe, der Schwellenwert (Abbruchsbedingung) und die maximale Anzahl der Iterationen (Wieviel Mal soll max. iteriert werden, falls der von uns gegebene Schwellenwert nicht erreicht werden kann). Darüber hinaus kann der Anwender die Anzahl der Cluster beliebig auswählen und manuell in das Parameterfeld eingeben. Damit soll bestimmt werden, zu wie vielen Clustern (Kundenklassen) der Algorithmus die Kundendatensätze (Datenfelder in der Datentabelle) teilen soll. Im Gegensatz zur Clusteranzahl kann die Dimensionanzahl (Merkmale/Faktoren) dynamisch nicht angegeben werden. Sie ist fix, und die Höchstanzahl wird durch die ausgewählten Datenfelder der Kundentabelle bestimmt (bei dem untersuchten Fallbeispiel max. 5, also Ankäufe<sup>963</sup>-Deckungsbeitrag-Frequency-Recency-Umsatz).

Die Ausgabefunktion der Ergebnisse von Fuzzy-Clustering ermöglicht uns, die Zugehörigkeiten einzelner Datensätze zu entsprechenden Clustern und die Zentren von Clustern in einer Tabellenform darzustellen. Zudem bietet diese Funktion auch die Möglichkeit, dass diese hergestellten Tabellen unter einem bestimmten Namen in der Datenbank für zukünftige Untersuchungen gespeichert werden. Auf diese Weise kann der Anwender auf die bereits vorhandenen Analyseergebnisse zu einem späteren Zeitpunkt wieder

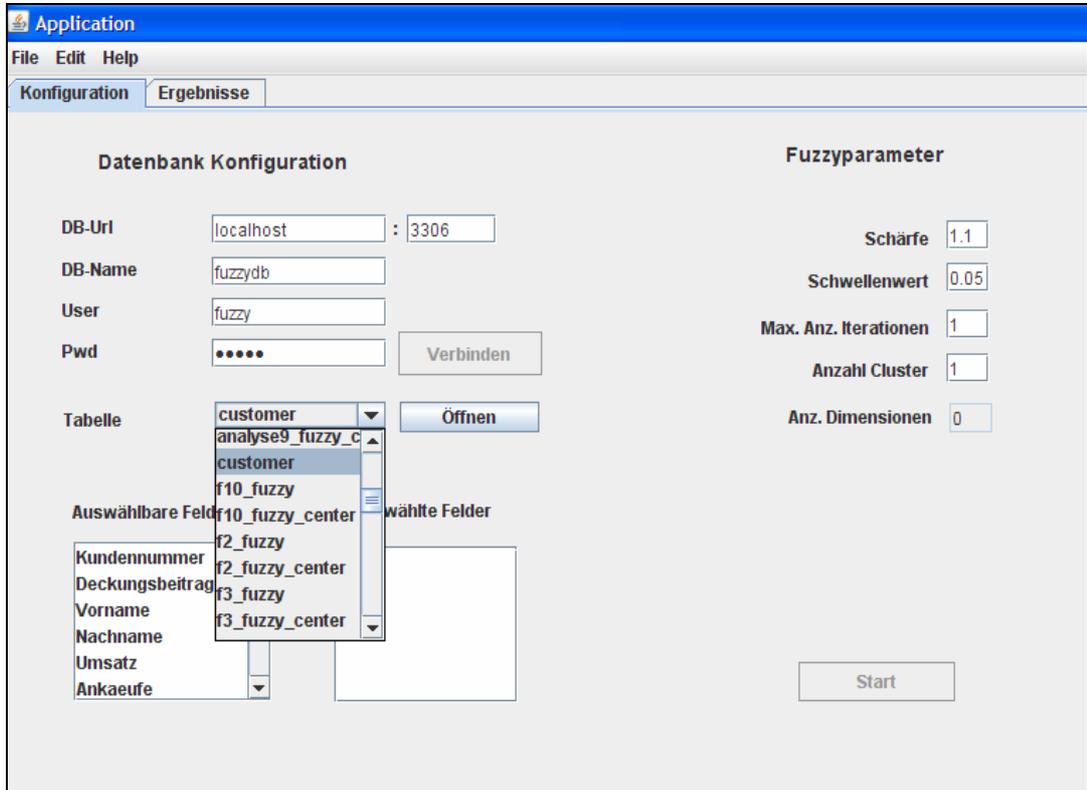
---

<sup>963</sup> Das Merkmal bzw. die Dimension „Ankäufe“ wird in die folgende Clusteranalyse nicht einbezogen.

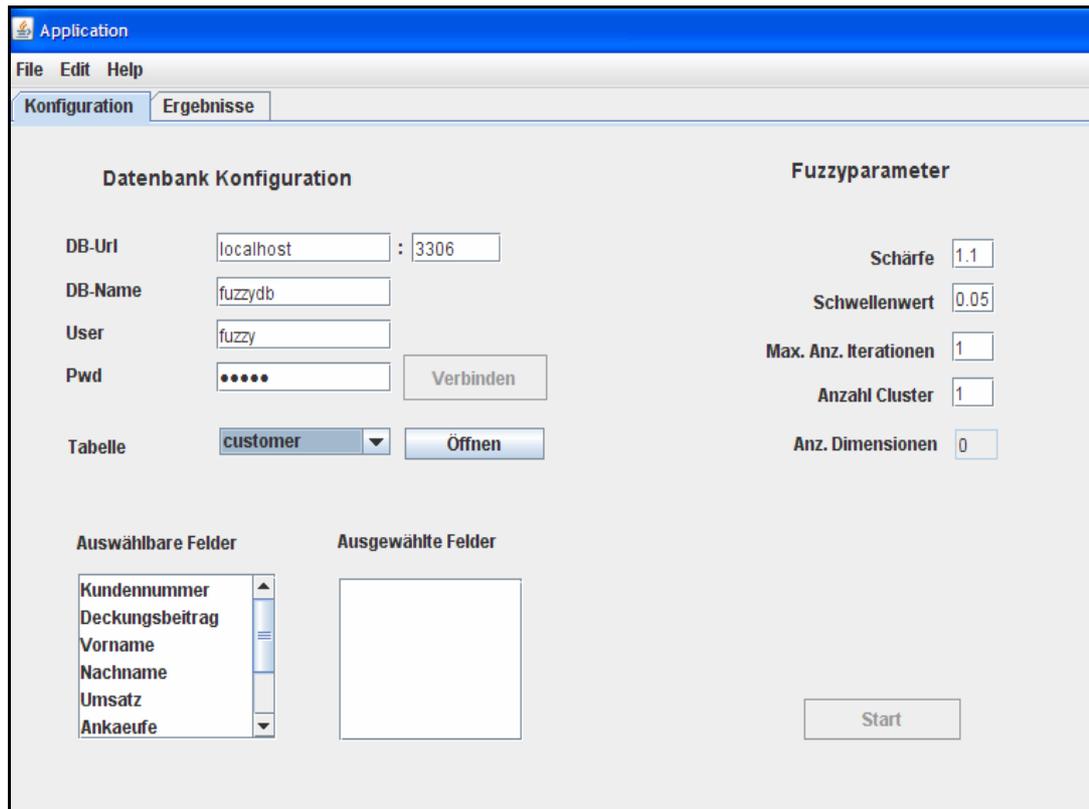
zurückgreifen.



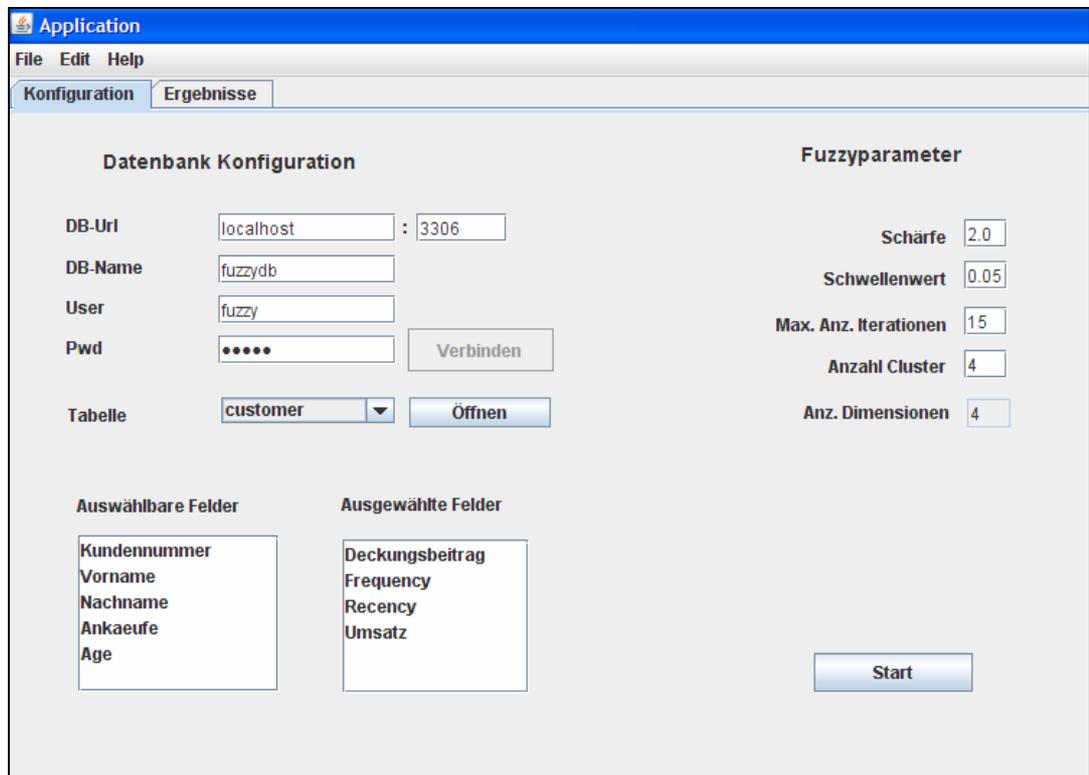
**Bild 2: Benutzeroberfläche für die Konfiguration**



**Bild 3: Die Konfiguration-GUI nach der Datenbankanbindung und Auswahl der Customer Tabelle**



**Bild 4: Die Konfiguration-GUI mit auswählbaren Feldern der Customer Tabelle**



**Bild 5: Die Konfiguration-GUI nach der Auswahl der Felder (Dimensionen) und Fuzzyparameter**

The screenshot shows a software application window titled 'Application' with a menu bar 'File Edit Help'. The 'Ergebnisse' tab is active. It contains two tables and a save button.

**Ergebnis**

Kundennu...	Deckungs...	Frequency	Recency	Umsatz	
809	-15	4	2	1300	0
810	-3	1	4	190	0
811	85	5	8	1300	0
812	6	1	4	200	0
813	265	4	3	4200	0
814	-26	3	9	1290	0
815	20	2	5	240	0
816	58	1	10	700	0
817	20	1	5	350	0
818	165	3	10	2550	0
819	75	1	12	1150	0

**Clusterzentren**

cluster	Deckungs...	Frequency	Recency	Umsatz
cluster1	9.61996	1.2248	6.76053	297.809
cluster2	77.6432	2.6018	5.60253	1430.31
cluster3	149.215	3.23508	5.25272	2643.18
cluster4	250.067	3.31055	3.91769	5077.68

Buttons: Ergebnis speichern, Speichern

Bild 6: Die Ergebnis-GUI für die Darstellung der Fuzzy-Ergebnisse der Clusterbildung (a)

The screenshot shows the same software application window as Bild 6, but with a different 'Ergebnis' table. The 'Clusterzentren' table is identical to the one in Bild 6.

**Ergebnis**

cluster1	cluster2	cluster3	cluster4	fuzzy_id
0.0244429	0.960428	0.0134157	0.00171295	1
0.989518	0.0075515	0.00193106	0.0009994...	2
0.016424	0.973241	0.00917431	0.00116027	3
0.991157	0.00625665	0.0015865	0.0009993...	4
0.034358	0.0681798	0.215591	0.681871	5
0.0294289	0.952979	0.0155808	0.00201131	6
0.995582	0.00242089	0.0009987...	0.0009987...	7
0.735634	0.226178	0.0319012	0.00628738	8
0.995591	0.00241143	0.0009986...	0.0009986...	9
0.00173919	0.00702721	0.989848	0.00138575	10
0.000076	0.007007	0.000076	0.000076	11

**Clusterzentren**

cluster	Deckungs...	Frequency	Recency	Umsatz
cluster1	9.61996	1.2248	6.76053	297.809
cluster2	77.6432	2.6018	5.60253	1430.31
cluster3	149.215	3.23508	5.25272	2643.18
cluster4	250.067	3.31055	3.91769	5077.68

Buttons: Ergebnis speichern, Speichern

Bild 7: Die Ergebnis-GUI für die Darstellung der Fuzzy-Ergebnisse der Clusterbildung (b)

## ➤ DATENBANKMANAGER: FuzzyDBManager

- übernimmt die Kommunikation zu Datenbank,
- bietet die Möglichkeit die Daten (Tabellen und Felder) aus einer Datenbank zu lesen,
- wandelt die ausgewählten Felder in für den Fuzzy-C-Means-Algorithmus nötigen Inputdatenmatrix um,
- triggert das Fuzzy-Clustering,
- speichert die Ergebnisse des Clustering in temporären Tabellen,
- bietet die Möglichkeit die temporären Ergebnisse in selbstbenannte Tabellen zu speichern.

Methoden:

*connectDB(String, String, String, String, String)*: Stellt die Verbindung zu einer angegebenen Datenbank mit angegebenem User.

*getTableList()*: Liest die Tabellen verbundener Datenbank und gibt sie in einer Liste zurück.

*getColumnList(String)*: Liest die Felder einer angegebenen Tabelle und gibt sie in einer Liste zurück.

*getTableData(String)*: Liest Felder (incl. Feldernamen) einer Tabelle und wandelt sie in eine String-Matrix um.

*getPrimaryKey(String)*: Bestimmt den primären Schlüssel einer Tabelle.

*getInputData(String, Vector)*: Liest die für Fuzzy-C-Means nötigen Inputdaten aus einer angegebenen Tabelle und wandelt sie in eine Inputdatenmatrix um.

*createTempTable(String, Vector, int)*: Erzeugt zwei temporäre Tabellen für das Speichern von Clustering-Ergebnissen (eine für die Zugehörigkeit der Datensätze zu einzelnen Klassen (Clustern), eine für die Clusterzentren, siehe Bild 6 und 7).

*saveNamedResult(String)*: Ändert temporäre Tabellen zu den angegebenen Namen („Speichert“ quasi die Ergebnisse).

*saveResults()*: Speichert die Clustering-Ergebnisse in temporäre Tabellen.

*executeClustering(String, Vector, float, double, int, int)*: Triggert Fuzzy-C-Means-Clustering.

➤ FUZZY-CLUSTERING: FuzzyCMeans

implementiert den Fuzzy-C-Means-Algorithmus.

Methoden:

*FuzzyCMeans()*: Konstruktor; initiiert die Zugehörigkeitsmatrix mit zufälligen Werten.

*berechneClusterZentren()*: Berechnet die Clusterzentren.

*berechnePartitionsMatrix()*: Die Zugehörigkeitsmatrix wird berechnet.

*berechneZielfunktion()*: Berechnet den Zielfunktionswert.

*run()*: Führt das Fuzzy-Clustering durch, indem sie die Methoden *berechneClusterZentren()* und *berechnePartitionsMatrix()* hintereinander aufruft, bis der Schwellenwert bzw. die maximale Anzahl der Iterationen erreicht wird.

*getMembership()*: Gibt die Zugehörigkeitsmatrix zurück.

*getClusterCenters()*: Gibt die Clusterzentren zurück.

*getDistance()*: Berechnet die euklidische Distanz zwischen den zwei angegebenen Vektoren.

## FUZZY-CLUSTERING MIT DEM FUZZY-C-MEANS-ALGORITHMUS

➤ BESTIMMUNG DER CLUSTERANZAHL

Zur Initialisierung der Clusterbildung muss eine Matrix angegeben werden, deren Zeilenanzahl mit der Anzahl der Objekte und deren Spaltenzahl mit der Anzahl der zu findenden Cluster übereinstimmt. Das bedeutet, dass die Clusteranzahl für den Prozess der Clusterbildung vorgegeben werden muss. Die

Matrix gibt an der Stelle  $(i, j)$  die Zugehörigkeit von Objekt  $j$  zum Cluster  $i$  an. Bei der Initialisierung wird eine Startmatrix zufällig vom Programm vorbelegt. Dies bedeutet, dass das Programm für jedes Datenobjekt gleichverteilte Clusterzugehörigkeiten vorgibt. Das Programm berechnet die Zugehörigkeiten durch die Verwendung der euklidischen Distanz als Distanzmaß zwischen den Datenobjekten und Clusterzentren.

Der Schwellenwert bestimmt die Genauigkeit des Analyseergebnisses und damit auch die Zahl der Iterationen (Das Programm kann abbrechen, bevor die max. Iteration erreicht wird). Für die vorliegende Anwendung wurde ein Wert von 0,05 verwendet und die max. Iterationsanzahl beträgt 15.

In dem Fallbeispiel werden 4 Merkmale (Dimensionen) untersucht. Diese sind Gesamtumsatz, den das Retail-Unternehmen von jedem einzelnen Kunden in 18 Monaten erzielt hat, individuellen Deckungsbeitrag des jeweiligen Kunden, die Kaufhäufigkeit und die vergangene Zeit von dem letzten Kauf (Recency).

Für die Clusterbildung muss die Anzahl der Cluster vorgegeben werden. Der Fuzzy-C-Means-Algorithmus soll aus der Menge der Kunden eine Struktur mit der vorgegebenen Clusterzahl finden. Dem Anwender ist jedoch die richtige Clusterzahl a priori unbekannt, deshalb müssen mehrere Durchläufe mit den vorgegebenen Clusterzahlen durchgeführt werden, um zu einer möglichst gut segmentierten Clusterbildung zu gelangen. Nach Ablauf des Fuzzy Clustering erfolgt eine Bewertung der Cluster-Ergebnisse, um die richtige Clusterzahl zu bestimmen. Die Güte jeder Clusterbildung kann mit Hilfe sogenannter Gütekriterien bzw. maße, die im Kapitel 6.6.8 detailliert beschrieben sind, überprüft werden. Sie messen die Qualität einer Clusterbildung durch die Zuordnung einer Zahl, wobei man von der Vorstellung ausgeht, dass die Clusterstruktur umso besser identifiziert wird, je stärker sich die Datenpunkte um ein jeweiliges Clusterzentrum konzentrieren.<sup>964</sup>

In der Untersuchung des Fallbeispiels werden Clusterzahlen von 2 bis 10 betrachtet.

Ein weiterer einzustellender Parameter ist ein Wert, der den Grad der Schärfe bzw. Unschärfe von Cluster-Ergebnissen bestimmt. Für unsere Anwendung wurde dieser Wert auf 2.0 eingestellt. Als Eingangsdaten für die Clusteranalyse werden die vorbereiteten 200 Kundendatensätze eingesetzt. Damit sind alle

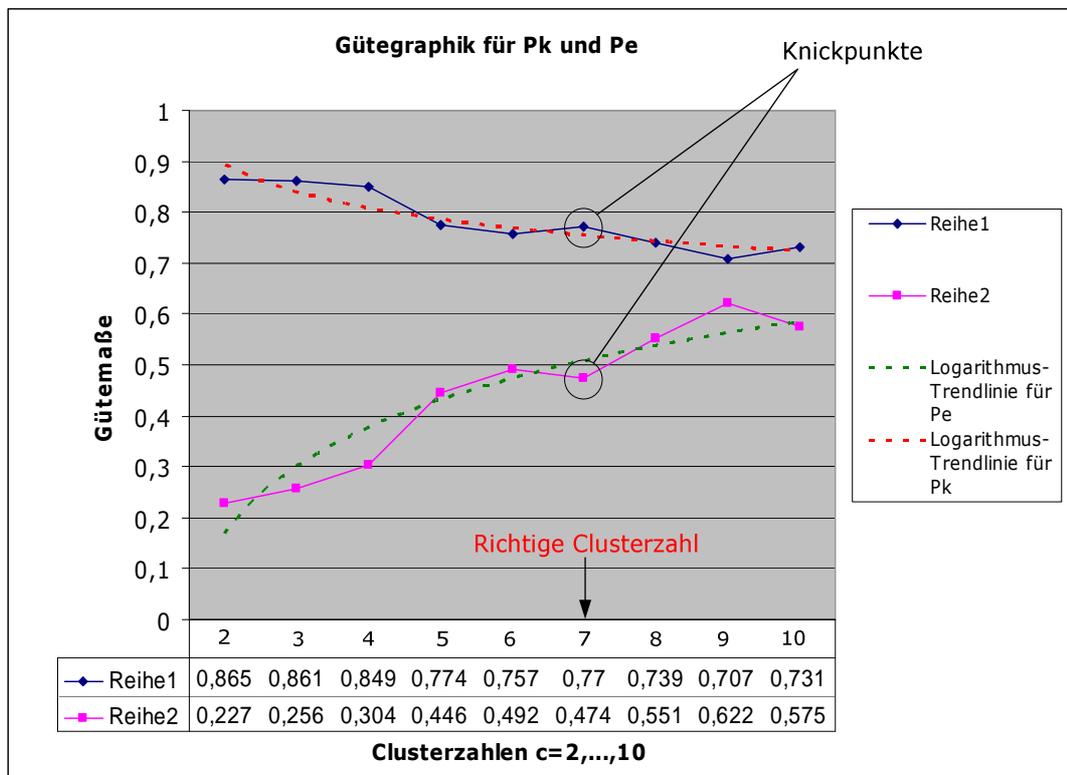
---

<sup>964</sup> d. h. je schärfer die Zugehörigkeitsmatrix der Endpartition ist.

Einstellungen getätigt und die Clusterbildung kann gestartet werden. Die folgende Liste stellt die Werte der Gütemaße bei verschiedenen Clusterzahlen.

Clusterzahl (c)	Partitionskoeffizient (Pk)	Partitionsentropie (Pe)	Proportionsexponent (Pex) - ungefähr
2	0,865287954	0,227281365	558,6588148
3	0,861208412	0,256130926	706,8295229
4	0,848743569	0,30430493	706,0734683
5	0,773713217	0,445815435	704,142824
6	0,756668349	0,491762862	687,1685372
7	0,770337426	0,47424491	702,7928545
8	0,738791476	0,550984169	700,6443907
9	0,707127412	0,621600835	693,6201145
10	0,730595008	0,574964453	686,9137351

Diese Werte können zur besseren Übersicht in einer Gütegraphik (siehe Abb. 65) visualisiert werden. Wie im Kapitel 6.6.8 bereits erwähnt ist, haben der Partitionskoeffizient (Pk) und die Partitionsentropie (Pe) ein monotones Verhalten in Abhängigkeit von der Clusterzahl  $c$ .



**Abbildung 65: Gütegraphik**

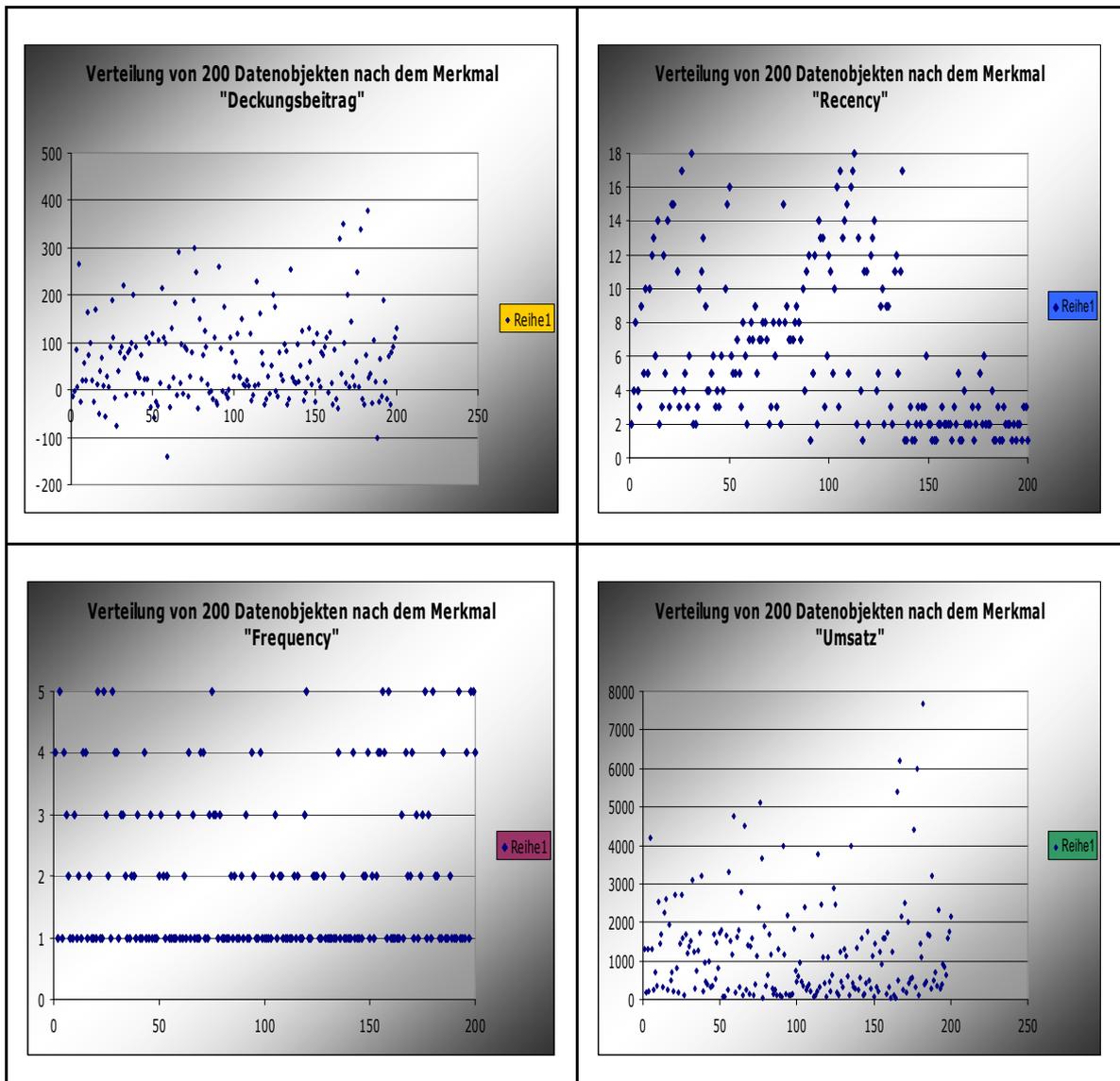
Quelle: eigene Darstellung

Gemäß dieser Monotonie ist bei der Suche nach der optimalen bzw. richtigen Clusterzahl  $c$  in der Gütegraphik diejenige zu wählen, für die beim Übergang von der Clusterzahl  $c-1$  zur Clusterzahl  $c$  der Wert für die Entropie unter den steigenden Trend fällt bzw. der Wert des Partitionskoeffizienten über den fallenden Trend steigt.<sup>965</sup> Diese Stelle ist bei Betrachtung der jeweiligen Kurve in der Graphik aus allen verbundenen Werten als „Knickpunkt“ zu erkennen (deshalb auch „Ellbogen-Kriterium“ genannt). Genau an dieser Stelle erfolgt die beste Trennung der Cluster.

---

<sup>965</sup> Vgl. Zimmermann (1995), S. 154.

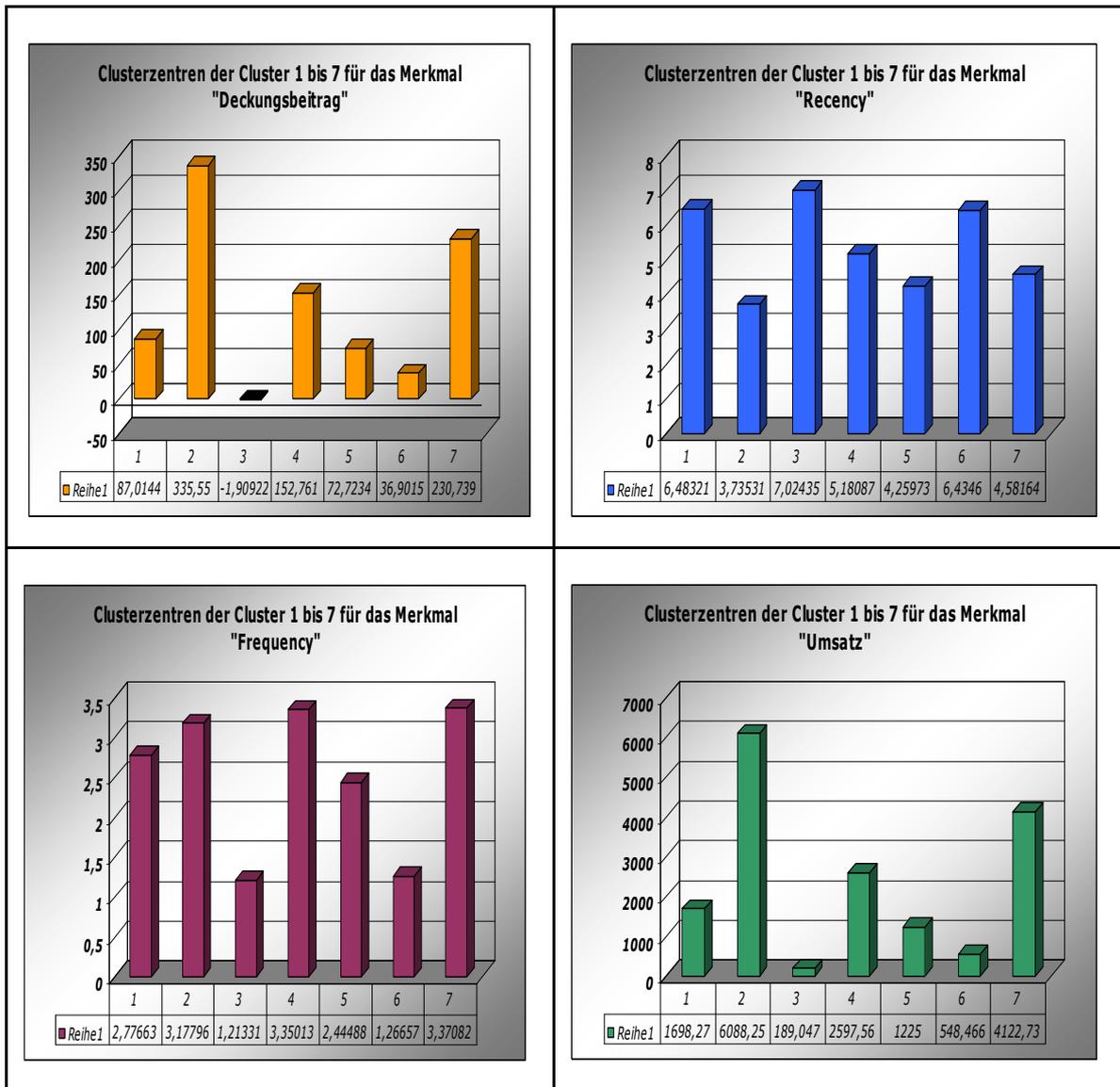
## 7.7 EVALUIERUNG DER ANALYSEERGEBNISSE



**Tabelle 9: Verteilung von 200 Datenobjekten bzw. Kundenmerkmale**

Quelle: eigene Darstellung

Fuzzyparameter	Cluster	DB	Frequency	Recency	Umsatz
<b>Schärfe:</b> 2,0 <b>Klassenanzahl:</b> 7 <b>Anzahl der Iterationen:</b> 15 <b>Schwellenwert:</b> 0,05 <b>Dimensionen:</b> 4	1	87,0144	2,77663	6,48321	1698,27
	2	335,55	3,17796	3,73531	6088,25
	3	-1,90922	1,21331	7,02435	189,047
	4	152,761	3,35013	5,18087	2597,56
	5	72,7234	2,44488	4,25973	1225
	6	36,9015	1,26657	6,4346	548,466
	7	230,739	3,37082	4,58164	4122,73



**Tabelle 10: Die gefundenen Clusterzentren für vier Merkmale (Dimensionen) nach der Analyse von 200 Datenobjekten**

Quelle: eigene Darstellung

Die dargestellten Diagramme zeigen die Clusterzentren bezüglich der vier Merkmale (Dimensionen) als Balkendiagramme. Das Programm wurde mit einem Fuzzifier von  $m=2.0$  mit unterschiedlichen Clusterzahlen  $c=2, \dots, 10$  ausgeführt und die Clusterbildung mit 7 Clustern für die folgende Interpretation ausgewählt (siehe Abb. 65).

## INTERPRETATION DES ANALYSEERGEBNISSES

- **Klasse Best:** Cluster 2 in der Tabelle 10

Diese Klasse ist eindeutig das Bestsegment. Die kumulierten Umsätze der Kunden bewegen sich um ca. 6000 €. Die Kunden, die zu dieser Klasse gehören, weisen die

besten Charakteristiken einer Superkunde, nämlich einen sehr hohen individuellen Deckungsbeitrag, einen geringen Wert für Recency (seit dem letzten Kauf vergangene Zeit) und eine hohe Kauffrequenz. Dieses Segment beschreibt das Segment der Kunden mit sehr hohem Einkommen, die ihr Geld gerne für Produkte und Dienstleistungen ausgeben. Für das untersuchte Unternehmen stellt dieses Segment das gewinnbringendste dar, da der Deckungsbeitrag hier mit einem Wert um 335.55 am höchsten ist.

- **Klasse Worst negativ:** Cluster 3 in der Tabelle 10

Dieses Segment ist geprägt durch negative Deckungsbeiträge und sehr geringe kumulierte Umsätze. Mit großer Sicherheit befinden sich in dieser Klasse die schlechtesten Kunden, die negative Deckungsbeiträge haben und in dem Betrachtungszeitraum der Analyse nur einen Kauf tätigen. Außerdem beträgt bei diesen Kunden die Zeitspanne seit ihrem letzten Kauf fast 7 Monate. Das Unternehmen soll Wege finden, die Situation dieser Kunden zu verbessern oder diese völlig aus dem Kundenstamm auszusortieren.

- **Klasse Frequenter 1th:** Cluster 7 in der Tabelle 10

Dieses Segment beschreibt die Kundengruppe mit hoher Kauffrequenz, welche auch auf eine hohe Anzahl der gekauften Produkte hindeutet. Diese Klasse überholt sogar mit der Kaufhäufigkeit für einen Moment die Besten. Da die Kunden dieser Klasse Produkte und Dienstleistungen mit großer Häufigkeit kaufen, kann das Unternehmen mit denen hohe Umsätze erwirtschaften. So ist die Klasse Frequenter die zweitbeste gewinnbringende Klasse für das Unternehmen (von dem hohen Umsatz in der Tabelle 9 erkennbar).

- **Klasse Uncertain:** Cluster 1 in der Tabelle 10

Die Klasse repräsentiert die Kundengruppe, die Käufe in Wert von ca. 1700 € tätigen. Der Wert des Deckungsbeitrags ist im Vergleich zu dessen Werten der Klassen Best und Frequenter gering aber diese Klasse verbirgt ein hohes Potential für zukünftige Käufe, denn sie weist eine hohe Kauffrequenz auf. Die Klasse Uncertain weist zudem einen hohen Wert für Recency, also fast denselben Wert dafür wie bei der Klasse Worst auf. Dies deutet darauf hin, dass die Kunden dieser Klasse dem Unternehmen nicht genug angebunden sind. Vermutlich kaufen sie Produkte und Dienstleistungen auch bei anderen Unternehmen. Um diese Kunden dem Unternehmen zu nähern, kann man auf deren Bedürfnisse ausgerichtete Produkte und Dienstleistungen entwickeln. Gezielt gesetzte Marketingstrategien und -kampagnen könnten dann die Kunden

dieser Klasse dazu veranlassen, dass sie sich in die Richtung besserer Klassen bewegen, wo die Anzahl der getätigten Kaufgeschäfte mit dem Unternehmen steigt und dafür ausgegebene Geldmenge erhöht. Diese Kunden können dann als *Spender* bezeichnet werden. Je öfter ein Kunde im Verlauf der Geschäftsbeziehung oder im Laufe des Betrachtungszeitraums (bei unserem Fallbeispiel 18 Monate) gekauft hat, desto stärker neigt er zu Wiederkaufen. Im ersten Blick könnte der Umsatz und Deckungsbeitrag als entscheidendes Kriterium für die Kundenbewertung betrachtet werden, dennoch können diese für längere noch profitablere Geschäftsbeziehungen irre führend sein, wenn man im CRM auf dieses Kriterium übermäßigen Wert legt. Schließlich kann es oft passieren, dass ein Kunde hunderte von Produkten in einem Geschäftstag kauft, somit einmalig hohe Umsätze und Deckungsbeiträge verursacht und nicht wieder mal das Unternehmen besucht. Daher ist die Kauffrequenz/-häufigkeit entscheidend für diese Klasse und damit auch für seine Bewertung im Zuge des CRM.

Die Klassen, die bisher beschrieben sind, stellen die wichtigsten Kundengruppen in dem Kundenstamm des Unternehmens dar. Sie sind die Grundklassen. Die übrigen Cluster 4, 5 und 6 können als Variationen diesen gefundenen Grundklassen zugeordnet werden. Die Kunden in dem Cluster 6 sind als eine Variante der Kunden in der Klasse Worst mit positiven Deckungsbeiträgen einzustufen und dementsprechend ist diese Kundengruppe mit dem Klassennamen *Worst positiv* zu bezeichnen. Die Kundengruppe der Cluster 4 und 5 können gemäß ihrer Kauffrequenz und ihren Umsatz jeweils *Frequenter 2th* (second) und *Frequenter 3th* (third) genannt werden.

# KAPITEL 8

---

## SCHLUSSFOLGERUNG

Die Zielstellung der Clusteranalyse, Strukturen bzw. Cluster in Daten zu finden, ist bei der Datenanalyse von großem Interesse. Häufig sind die in den Datensätzen vorliegenden Cluster nicht deutlich separiert. Viele Daten sind Mischformen verschiedener Cluster. Die Fuzzy-Clusteranalyse bietet eine Möglichkeit, diese Daten bei der Clusteranalyse entsprechend zu berücksichtigen. Hierbei werden die Daten den Clustern mit einem Zugehörigkeitsgrad zwischen 0 und 1 zugeordnet. Ein Datum kann daher gleichzeitig mehreren Clustern angehören, was für die Datenanalyse einen besonderen Vorteil darstellt. Die scharfe Zuordnung zu nur einem Cluster hat dagegen den Nachteil, dass ein Teil der Informationen verloren geht, da eine eindeutige Zuordnung bzw. Nicht-Zuordnung nicht stets gegeben sein muss.

In komplexen Systemen, in denen es schwer ist, exakte Aussagen über das Verhalten des Systems zu machen, kommt vielfach die Fuzzy-Logik zum Einsatz. Die Fuzzy-Logik imitiert die menschliche Denkweise für das Lösen komplexer Probleme. Menschen nehmen bei ihren Denkprozessen nicht nur zwei Werte „ja“ oder „nein“ wahr, sondern die Ein- und Ausgabe des menschlichen Denkprozesses können Zwischen- oder sogenannte Graustufen besitzen. Dieser Denkprozess beruht größtenteils auf unscharfen Wahrnehmungen und Schlussfolgerungen.

Fuzzy Logik-basierte Anwendungen sind Anwendungen, in denen die Verständlichkeit von großer Relevanz ist. Unscharfe Formulierungen des Fuzzy-Logik-Ansatzes liefern paradoxerweise eher verständliche Ergebnisse bei komplexen Problemen. Der Fuzzy-Logik-Ansatz zieht die Verständlichkeit der Genauigkeit vor. Das soll jedoch nicht bedeuten, dass seine Ergebnisse ungenau sind, sondern es gibt ein Trade-off zwischen

Verständlichkeit und Genauigkeit in Fuzzy-Anwendungen.

Die Fuzzy-Logik kommt auch in Anwendungen zum Einsatz, in denen man die klassische Logik ersetzen soll, wenn es dadurch möglich wird, bessere Analyseergebnisse zu erzielen. Es ist jedoch ratsam, in jedem einzelnen Fall zu prüfen, ob die Fuzzy-Logik tatsächlich zu besseren Ergebnissen führt oder ob nicht Lösungen mit klassischen Methoden und Verfahren genauso gut erreicht werden können. Denn einerseits werden die Ergebnisse des Fuzzy-Logik-Ansatzes als eine Approximation betrachtet, so dass sie nicht immer die beste Lösung liefern können. Andererseits ist die Anwendung des Fuzzy-Logik-Ansatzes mit einem erhöhten Informationsverarbeitungsaufwand verbunden. Da in der Fuzzy Set Theory im Vergleich zu klassischen Verfahren an die Stelle von Zahlen Funktionen treten (Zugehörigkeitsfunktion), erhöht die zu verarbeitende Informationsmenge zwangsmäßig. Die Vielzahl der vorhandenen Operatoren und Zugehörigkeitsfunktionen der Fuzzy-Mengen-Theorie stellt nicht nur den Vorteil dar, dass man daraus für den jeweiligen Fall die richtigen auswählen kann, sondern man ist auch gezwungen, kontextabhängig eine entsprechende Wahl zu treffen. Diese Wahl setzt unter Umständen eine größere Anwendungserfahrung oder aber auch umfangreichere empirische Untersuchungen voraus, die dann selbstverständlich auch Kosten verursachen.

Fuzzy-Logik's Arbeitsweise enthält folgende drei Schritte: Fuzzifizierung, Regelformulierung und Defuzzifizierung. Fuzzifizierung ist die Umwandlung alltäglicher scharfer Daten in unscharfe Werte, Regelformulierung ist die Berechnung der unscharfen Werte und deren Verknüpfung zu einem unscharfen Ergebnis und Defuzzifizierung ist die Umwandlung der Ergebnisse von dem unscharfen Bereich in den scharfen Bereich, so dass crisp bzw. scharfe Daten entstehen.

In dem sechsten Kapitel wurde die Theorie von Fuzzy-Mengen vorgestellt und seine mathematische Grundlage vermittelt. Anschließend wurde der Fuzzy-C-Means-Algorithmus erklärt, welcher aus der Idee des probabilistischen Clustering hervorgeht. Der Fuzzy-C-Means-Algorithmus gilt als Basisalgorithmus aller Fuzzy-Clusteringverfahren und zeichnet sich vor allem durch eine einfache Berechnungsvorschrift aus. In der Praxis liefert der Algorithmus schon nach einigen wenigen Iterationen gute Näherungen an die endgültige Clusterlösung.

Aufbauend auf der Vorstellung der wichtigsten Ideen und Konzepte in Bezug auf Fuzzy-Logik wurde im siebten und praktischen Kapitel eine Anwendung im wirtschaftlichen Bereich untersucht. Dabei wurde die Clusteranalyse anhand eines

Fallbeispiels näher betrachtet. Gemäß dieser Clusteranalyse wurden die dem Fallbeispiel zugrunde liegenden Kunden anhand ihrer Merkmale in Klassen gruppiert bzw. segmentiert. Anschließend wurden diese Klassen als Analyseergebnisse inhaltlich interpretiert. Ein wichtiger Vorteil der Segmentation ist, dass man die auf der Grundlage der bereits vorhandenen Kunden erfolgte Analyse für die Beurteilung neu gewonnener Kunden verwenden kann. Nachdem ein Klassifikator zur Einteilung von bestehenden Kunden in Cluster erzeugt wurde und diese Cluster auch inhaltlich interpretierbar sind, kann dieser Klassifikator eingesetzt werden, um neue Kunden in die bereits vorhandenen Cluster entsprechend einzuordnen.

Einige der durch Experimente mit dem Prototyp bestätigten Vorteile der Fuzzy-Logik Anwendungen sind:

- Das Imitieren der menschlichen Denkweise.
- Die Fähigkeit mit unpräzisen oder vagen Begriffen umzugehen.
- Modellierung komplexer, nicht linearer Probleme.
- Verbesserte Wissensdarstellung.

Obwohl es viele Diskussionen über den Einsatz von Fuzzy-Logik gibt, birgt sie ein großes Potential in sich. Dementsprechend findet sie ihren Einsatz immer mehr in vielen Anwendungen der Datenanalyse aus verschiedenen Bereichen.

## 9 LITERATURVERZEICHNIS

Die folgende Literatur habe ich zur Erstellung meiner Diplomarbeit direkt oder indirekt benutzt.

### BÜCHER

**[Ackerschott, Harald 2001]**

Wissensmanagement für Marketing und Vertrieb: Kompetenz steigern und Märkte erobern, 1. Aufl., Wiesbaden: Gabler 2001.

**[Amberg, Michael/Schumacher, Jörg 2002]**

CRM-Systeme und Basistechnologien, in: Meyer, Matthias (Hrsg.): CRM-Systeme mit EAI - Konzeption, Implementierung und Evaluation, 1. Aufl., Wiesbaden: Vieweg 2002, S. 21-59.

**[Bandemer, H./Gottwald, S. 1993]**

Einführung in Fuzzy-Methoden, 4. überarbeitete und erweiterte Auflage, Berlin: Akademie Verlag 1993.

**[Berry, L. L. 1983]**

Relationship Marketing, in: Berry, L. L.; Shostack, G. L.; Upah, G. D. (Hrsg.): Emerging Perspectives on Services Marketing, Chicago 1983, S. 25-28.

**[Berry, M. J. A./Linoff, G. 1997]**

Data Mining Techniques: For Marketing, Sales and Customer Support, New York: Wiley 1997.

**[Berry, M. J. A./Linoff, G. 1999]**

Mastering Data Mining: The Art and Science of Customer Relationship Management. New York : Wiley 1999

**[Biewer, Benno 1997]**

Fuzzy-Methoden. Praxisrelevante Rechenmodelle und Fuzzy Programmiersprachen, Berlin [u. a.]: Springer 1997.

**[Bothe, H. -H. 1995]**

Fuzzy Logic. Einführung in Theorie und Anwendungen, 2. Auflage, Berlin [u. a.]: Springer 1995,.

**[Buck-Emden, Rüdiger/Zencke, Peter 2004]**

mySAP CRM 4.0, 1. Aufl., Galileo 2004.

**[Bullinger, Hans Jörg/Wagner, Kristina/Ohlhausen, Peter 2000]**

Intellektuelles Kapital als wesentlicher Bestandteil des Wissensmanagement, in: Hermann Krallmann (Hrsg.): Wettbewerbsvorteile durch Wissensmanagement: Methodik und Anwendungen des Knowledge Management, Stuttgart: Schäffer-Poeschel Verlag 2000, S. 73-90.

**[Cabena , P./Hadjinian, P./Stadler, R./Verhees, J./Zanasi, A. 1998]**

Discovering Data Mining – From Concept to Implementation, Upper Saddle River, NJ: Prentice Hall 1998.

**[Cantor, Georg 1966]**

Gesammelte Abhandlungen mathematischen und philosophischen INhalts. Herausgegeben von E. Zermelo. (Reprographischer Nachdruck der Ausgabe Berlin 1932). Hildesheim: Georg Olms Verlagsbuchhandlung.

**[Chamoni, P. 1998]**

Entwicklungslinien und Architekturkonzepte des On-Line Analytical Processing, in: Chamoni, P.; Gluchowski, P. (Hrsg.): Analytische Informationssysteme – Data Warehouse, On-line Analytical Processing, Data Mining, 1. Auflage, Berlin: Springer Verlag 1998, S. 231-250.

**[Chamoni, Peter/Gluchowski, Peter 2006]**

Analytische Informationssysteme, 3. Aufl., Berlin: Springer 2006.

**[Cornelsen, Jens 2000]**

Kundenwertanalysen im Beziehungsmarketing, Nürnberg: GIM-Verlag 2000, S. 38.

**[Dangelmaier, W./Helmke, S./Uebel, M. 2004]**

Grundrahmen des Customer Relationship Management-Ansatzes, in: Dangelmaier, W; Helmke, S.; Uebel, M. (Hrsg.): Praxis des Customer Relationship Management, 2. Aufl., Wiesbaden: Gabler 2004, S. 2-16.

**[Davenport, Thomas H./Prusak, Lawrence 1998]**

Wenn Ihr Unternehmen wüsste, was es alles weiß: Das Praxis Handbuch zum Wissensmanagement, Landsberg: Moderne Industrie 1998.

**[Davenport, Thomas H./Laurence Prusak 2000]**

Working Knowledge. How Organizations Manage What They Know, 2. Aufl., Boston: Harvard Business School Press 2000.

**[Davis, Stanley M. 1987]**

Future Perfect. 1. Aufl., Reading, Massachusetts: Addison Wesley 1987.

**[Diller, H./Müllner, M. 1998]**

Kundenbindungsmanagement, in: A. Meyer (Hrsg.): Handbuch Dienstleistungs-Marketing, Stuttgart 1998, S. 1220.

**[Dittmar, Carsten/Gluchowski, Peter 2002]**

Synergiepotenziale und Herausforderungen von Knowledge Management und Business Intelligence, in: Uwe Hannig (Hrsg.): Knowledge Management und Business Intelligence, Berlin [u. a.]: Springer Verlag 2002, S. 27-43.

**[Drösser, Cristoph 1994]**

Fuzzy Logic, Methodische Einführung in krauses Denken. Hamburg: Rowohlt Taschenbuch Verlag 1994.

**[Dubois, D./Prade, H. 1980]**

Fuzzy Sets and Systems. Theory and Applications, New York: Academic Press 1980.

**[Dubois, D./Prade, H. 1988]**

Possibility Theory. An Approach to Computerized Processing of Uncertainty, New York [u. a.]: Plenum Press 1988.

**[Eckey, Hans-Friedrich /Kosfeld, Reinhold/Rengers, Martina 2002]**

Multivariate Statistik, 1. Aufl., Wiesbaden: Gabler 2002.

**[Ematinger, Reinhard/Sommer, Renate/Dr. Stengl, Britta 2001]**

CRM mit Methode, 1. Aufl., Bonn: Galileo Press 2001.

**[Eschenbach, Sebastian/Geyer, Barbara 2004]**

Wissen & Management – 12 Konzepte für den Umgang mit Wissen im Management, Wien: Linde 2004.

**[Fayyad, U. M./Piatetsky-Shapiro, G./Smyth, P. 1996a]**

From Data Mining to Knowledge Discovery: An Overview, in: Fayyad, U. M; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R. (Hrsg.): Advances in Knowledge Discovery and Data Mining, AAAI Press / The MIT Press, Menlo Park, California 1996, S. 1-34.

**[Förster, Anja 2001]**

Klassisches Marketing versus Online Marketing, in: Kreuz, Peter; Förster, Anja; Schlegelmilch, Bodo B. (Hrsg.): Customer Relationship Management im Internet: Grundlagen und Werkzeuge für Manager, BoD, Books on Demand 2001, S. 33-40.

**[Gawlik, Tom/Kellner, Joachim/Selfert, Dirk 2002]**

Effiziente Kundenbindung mit CRM, 1. Aufl., Bonn: Galileo Press 2002, S. 39-99.

**[Gluchowski, Peter /Gabriel, Roland /Dittmar, Carsten 2008]**

Management Support Systeme und Business Intelligence, 2. Aufl., Berlin: Springer 2008.

**[Grothe, Martin /Gentsch, Peter 2000]**

Business Intelligence – Aus Informationen Wettbewerbsvorteile gewinnen, München: Addison Wesley 2000.

**[Güldenber,Stefan 2003]**

Wissensmanagement und Wissenscontrolling in Lernenden Organisationen, 4.Aufl., Wiesbaden: Deutscher Universitätsverlag DUV 2003.

**[Günter, B./Helm, S. 2006]**

Kundenbewertung im Rahmen des CRM, in: Hippner, H.; Wilde, K. (Hrsg.): Grundlagen des CRM – Konzepte und Gestaltung, 2. Aufl., Wiesbaden: Gabler 2006, S. 360-366.

**[Hannig, Uwe 2002]**

Knowledge Management und Business Intelligence (Hrsg.), Berlin [u. a.]: Springer 2002, S. 3-25.

**[Helmke, S./Dangelmaier, W. 2002]**

CRM-Audit – Grundstein für eine erfolgreiche Einführung von CRM, in: Helmke, S.; Uebel, M.; Dangelmaier, W. (Hrsg.): Effektives Customer Relationship Management. Instrumente – Einführungskonzepte – Organisation, 2. Aufl., Wiesbaden: Gabler 2002, S. 295-306.

**[Hettich, S./Hippner, H. 2001]**

Assoziationsanalyse, in: Hippner, Hajo; Küsters, Ulrich; Meyer, Matthias; Wilde, Klaus D. (Hrsg.): Handbuch Data Mining im Marketing – Knowledge Discovery in Marketing Databases, Wiesbaden: Vieweg 2001, S. 459 – 493

**[Hippner, H./Wilde, K. 2001]**

Der Prozess des Data Mining im Marketing, in: Hippner, Hajo; Küsters, Ulrich; Meyer, Matthias; Wilde, Klaus D. (Hrsg.): Handbuch Data Mining im Marketing – Knowledge Discovery in Marketing Databases, Wiesbaden: Vieweg 2001, S. 53-123.

**[Hippner, H./Wilde, K. 2003a]**

CRM – Ein Überblick, in: Helmke, S.; Uebel, M.; Dangelmaier, W.; (Hrsg.): Effektives Customer Relationship Management. Instrumente – Einführungskonzepte - Organisation, 3. Aufl., Wiesbaden: Gabler 2003, S. 4-37.

**[Hippner, H./Wilde, K. 2003b]**

Data Mining im CRM, in: Helmke, S.; Uebel, M.; Dangelmaier, W.; (Hrsg.): Effektives Customer Relationship Management. Instrumente – Einführungskonzepte - Organisation, 3. Aufl., Wiesbaden: Gabler 2003, S. 211- 231.

**[Hippner, H. 2006]**

CRM – Grundlagen, Ziele und Konzepte, in: Hippner, H.; Wilde, K. (Hrsg.): Grundlagen des CRM – Konzepte und Gestaltung, 2. Aufl., Wiesbaden: Gabler 2006, S. 17-39.

**[Hippner, H./ Rentzmann, Rene/ Wilde, Klaus D. 2006]**

Aufbau und Funktionalitäten von CRM-Systemen, in: Hippner, H.; Wilde, Klaus D. (Hrsg.): Grundlagen des CRM – Konzepte und Gestaltung, 2. Aufl., Wiesbaden: Gabler 2006, S. 45-74.

**[Hippner, H./ Hoffmann, O./Rimmelspacher, U./Wilde, K. D. 2006]**

IT-Unterstützung durch CRM-Systeme am Beispiel von mySap CRM, in: Hippner, H.; Wilde, Klaus D. (Hrsg.): Grundlagen des CRM – Konzepte und Gestaltung, 2. Aufl., Wiesbaden: Gabler 2006, S. 75-95.

**[Holland, Heinrich/Huldi, Christian/Kuhfuß, Holger/Nietsche, Martin 2001]**

CRM im Direktmarketing – Kunden gewinnen durch interaktive Prozesse, 1. Aufl., Wiesbaden: Gabler 2001.

**[Holland, H. 2004a]**

Direkt-Marketing, 2. Aufl., München: Franz Vahlen Verlag 2004, S. 177-194, 107-148.

**[Holland, H. 2004b]**

CRM erfolgreich einsetzen. Warum CRM-Projekte scheitern und wie sie erfolgreich werden, Edition Praxiswissen, Göttingen: Business Village 2004, S. 21-26

**[Holthuis, Jan 1998]**

Multidimensionale Datenstrukturen - Modellierung, Strukturkomponente, Implementierungsaspekte, in: Muksch, Harry; Behme, Wolfgang (Hrsg.): Das Data Warehouse Konzept. Architektur – Datenmodelle – Anwendungen, 3. Aufl., Wiesbaden: Gabler 1998, S. 143-193.

**[Holthuis, Jan 1998]**

Der Aufbau von Data Warehouse-Systemen. Konzeption – Datenmodellierung – Vorgehen, 2. Aufl., Wiesbaden: Deutscher Universitätsverlag DUV 1998.

**[Homburg, C./Bruhn, M. 1999]**

Kundenbindungsmanagement: Begriff und Grundlagen des Kundenbindungsmanagements, in: Bruhn, Manfred; Homburg, Christian (Hrsg.): Handbuch Kundenbindungsmanagement, Grundlagen – Konzepte – Erfahrungen, 2. Aufl., Wiesbaden 1999.

**[Homburg, C./Sieben, F. G. 2000]**

Customer Relationship Management (CRM) – Strategische Ausrichtung statt IT-getriebenem Aktivismus, in: Bruhn, M.; Homburg, C. (Hrsg.): Handbuch Kundenbindungsmanagement, 3. Aufl., Wiesbaden: Gabler 2000, S. 472-501.

**[Inmon W. H. 1996]**

Building the Data Warehouse, 2nd Edition, New York u. a.: Wiley Computer Publishing 1996. (aktuellste Auflage: 3. Aufl. 2002).

**[Janetzko, D./Steinhöfel, K. 1997]**

Lotsen los! Data Mining: Verborgene Zusammenhänge in Datenbanken aufspüren, in: c't, (3) 1997, S. 294-300.

**[Kehl, R. E./Rudolph B. J. 2001]**

Warum CRM-Projekte scheitern, in: Jörg Link (Hrsg.): Customer Relationship Management – Erfolgreiche Kundenbeziehungen durch integrierte Informationssysteme, 1. Aufl., Berlin: Springer 2001, S. 253-275.

**[Klir, George J. /Folger, Tina A. 1988]**

Fuzzy Sets, Uncertainty, and Information. Englewood Cliffs, New Jersey: Prentice Hall 1988.

**[Knox, Simon/Maklan, Stan/Ryals, Lynette. 2000]**

Customer Relationship Management (CRM) – Building the Business Case, London: Pearson Education 2000.

**[Kolbe, Lutz M./Österle, Hubert/Brenner, Walter/Geib, Malte 2003]**

Grundlagen des Customer Knowledge Management, in: Kolbe, L.; Brenner, W.; Österle, H. (Hrsg.): Customer Knowledge Management, Berlin et al.: Springer 2003, S. 3-21.

**[Kosko, Bart 1993]**

Fuzzy Logisch. Eine neue Art des Denkens, Hamburg: Carlsen 1993. Übersetzung aus dem Amerikanischen Originaltitel: *Fuzzy Thinking. The new science of fuzzy logic.*

**[Kotler, Philip/ Bliemel, Friedhelm 1999]**

Marketing-Management: Analyse, Planung, Umsetzung und Steuerung, 9. Aufl., Stuttgart: Schäffer-Poeschel 1999.

**[Kotler, Philip 2001]**

Einschätzungen und Erfahrungen an der Schwelle des neuen Jahrtausends, in: Marketing & Kommunikation (Hrsg.), Jahrbuch CRM Schweiz, Berneck: Rheintaler Druckerei und Verlag 2001.

**[Köhler, Richard 2001]**

Customer Relationship Management – Interdisziplinäre Grundlagen der systematischen Kundenorientierung, in: Klein, S.; Loebbecke, C. (Hrsg.): interdisziplinäre Managementforschung und Lehre, Wiesbaden: Gabler 2001, S. 79-107.

**[Kreuz, Peter/Förster, Anja 2001]**

Was Sie schon immer über Customer Relationship Management wissen wollten, in: Kreuz, Peter; Förster, Anja; Schlegelmilch, Bodo B. (Hrsg.): Customer Relationship Management im Internet: Grundlagen und Werkzeuge für Manager, BoD, Books on Demand 2001, S. 8-19.

**[Kruse, Rudolph/Gebhardt, Jörg/Klawonn, Frank 1993]**

Fuzzy-Systeme, Stuttgart: Teubner 1993.

**[Link, J./Hildebrand, V. G. 1993]**

Database-Marketing und Computer Aided Selling – Strategische Wettbewerbsvorteile durch neue informationstechnologische Systemkonzeptionen, München: Vahlen 1993, S. 93 ff.

**[Link, J./Hildebrand, V. 1995]**

EDV-gestütztes Marketing im Mittelstand: Wettbewerbsvorteile durch kundenorientierte Informationssysteme, in: Link, J.; Hildebrand, V. (Hrsg.): EDV-gestütztes Marketing im Mittelstand, München: Franz Vahlen Verlag 1995, S. 1- 21.

**[Link, Jörg 2001]**

Customer Relationship Management – Erfolgreiche Kundenbeziehungen durch integrierte Informationssysteme (Hrsg.), 1. Aufl., Berlin: Springer 2001, S. 1-33.

**[McNeill, Daniel/Freiberger, Paul 1996]**

Fuzzy Logic. Die unscharfe Logik erobert die Technik, aktualisierte Taschenbuchausgabe, München: Droemer Knauer-Verlag 1996

**[Mentzl, Ronald/Ludwig, Cornelia 1998]**

Das Data Warehouse als Bestandteil eines Database Marketing-Systems, in: Muksch, Harry; Behme, Wolfgang (Hrsg.): Das Data Warehouse Konzept. Architektur – Datenmodelle – Anwendungen, 3. Aufl., Wiesbaden: Gabler 1998.

**[Mertens, Peter/Wieczerek, Hans Wilhelm 2000]**

Data X Strategien: Data Warehouse, Data Mining und operationale Systeme für die Praxis, Berlin/Heidelberg: Springer 2000.

**[Meyer, Matthias,/Weingärtner, S./Döring, F. 2001]**

Kundenmanagement in der Network Economy – Business Intelligence mit CRM und e-CRM, 1. Aufl., Vieweg 2001, S. 53-92, 93-126.

**[Muksch, Harry/Behme, Wolfgang 1998]**

Das Data Warehouse-Konzept als Basis einer unternehmensweiten Informationslogistik, in: Muksch, Harry; Behme, Wolfgang (Hrsg.): Das Data Warehouse Konzept. Architektur – Datenmodelle – Anwendungen, 3. Aufl., Wiesbaden: Gabler 1998.

**[Muther, Andreas 2001]**

Electronic Customer Care: Die Anbieter-Kunden-Beziehung im Informationszeitalter, 3. überarbeitete Aufl., Berlin: Springer 2001, S. 13-18, 67-79.

**[Nonaka, Ikujiro 1991]**

The Knowledge-Creating Company, in: Harvard Business Review Vol. 69, Nr. 6, S. 96-104.

**[Nonaka, Ikujiro/Takeuchi, Hirotaka 1997]**

Die Organisation des Wissens. Wie japanische Unternehmen eine brachliegende Ressource nutzbar machen, Frankfurt/Main [u. a.]: Campus 1997.  
Original: The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovation, New York: Oxford Univ. Press 1995.

**[Oberweis, Andreas/Paulzen, Oliver/Sexauer, Hagen J. 2002]**

Die Integration des Wissensmanagement in den Entwicklungsprozess von CRM-Systemen, in: Meyer Matthias (Hrsg.): CRM-Systeme mit EAI: Konzeption, Implementierung und Evaluation, 1. Aufl., Wiesbaden: Vieweg 2002, S. 117-135.

**[Overlack, Jochen 1988]**

Wettbewerbsvorteile durch Informationstechnologie. Frankfurt am Main [u.a.]: Peter Lang Verlag 1988, S. 9.

**[Peppers, Don/Rogers, Martha 1996]**

Strategien für ein individuelles Marketing. München: Knauer 1996.

**[Pffiffer, M. 1999]**

Wissen wirksam machen: wie Kopfarbeiter produktiv werden, 2. Aufl., Bern [u. a.]: Paul Haupt Verlag 1999.

**[Picot, A./ Reichwald, R./Wigand, R. 2001]**

Die grenzlose Unternehmung: Information, Organisation und Management. Lehrbuch zur Unternehmensführung im Informationszeitalter, 4.Aufl., Wiesbaden 2001.

**[Piller, Frank/Schaller, Christian 2004]**

Innovative Kundenbindungsstrategien im E-Business – ein ganzheitlicher Ansatz, in: Hagen J. Sexauer (Hrsg.): Konzepte des Customer Relationship Management (CRM): Strategien - Instrumente - Umsetzung, Wiesbaden: Denk!Institut 2004, S. 57-81.

**[Pofahl, Ulrich /Viertl, Reinhard 2006]**

Exakte Zahlen genügen nicht: Mathematik und Ungewissheit, Wien: Technische Universität 2006.

**[Polanyi, Michael 1997]**

The Tacit Dimension, in Prusak, L. (Hrsg.): Knowledge in Organizations, Boston, MA: Butterworth-Heinemann 1997, S. 135-146, auszugsweiser Nachdruck aus Polanyi (1996): The Tacit Dimension, New York: Doubleday 1966.

**[Probst, Gilbert/Raub, Steffen/Romhardt, Kai 1999]**

Wissen managen, 3. Aufl., Frankfurt am Main/Wiesbaden: Gabler 1999.

**[Rapp, Reinhold 2005]**

Customer Relationship Management: Das Konzept zur Revolutionierung der Kundenbeziehungen; 3. Aufl., Campus 2005, S. 10-39.

**[Rehäuser, J./Krcmar, H. 1996]**

Wissensmanagement im Unternehmen, in: Schreyögg, G.; Conrad, P. (Hrsg.): Managementforschung 6 (Band): Wissensmanagement, Berlin, New York: de Gruyter 1996, S. 1-40.

**[Reichheld, F. F. 1997]**

Der Loyalitätseffekt – Die verborgene Kraft hinter Wachstum, Gewinnen und Unternehmenswert, Frankfurt: Campus 1997, S. 52.

**[Reinke, H./Schuster, H. 1999]**

OLAP verstehen. Unterschleißheim: Microsoft Press Deutschland 1999.

**[Schätzler, D./Eilingsfeld, F. 1997]**

Intranets: Firmeninterne Informationssysteme mit Internet-Technologie, Heidelberg: dpunkt Verlag 1997.

**[Schaller, C./Stotko, C./Piller, F. 2006]**

Mit Mass Customization basiertem CRM zu loyalen Kundenbeziehungen, in: Hippner, H.; Wilde, Klaus D. (Hrsg.): Grundlagen des CRM – Konzepte und Gestaltung, 2. Aufl., Wiesbaden: Gabler 2006, S. 121-128.

**[Scharnbacher, Kurt/Guido, Kiefer 2003]**

Kundenzufriedenheit, 3. Aufl., Oldenburg 2003, S. 1-16.

**[Schildhauer, Thomas/Grothe, Martin/Braun, Matthias/Schultze, Matthias 2004]**

Business Intelligence – Durch eBusiness Strategien und Prozesse verbessern, 1. Aufl., Göttingen: Business Village Verlag 2004.

**[Schildt, Gerhard-Helge /Kastner, Wolfgang 1998]**

Prozessautomatisierung. Wien: Springer 1998.

**[Schirmeister, R./Kreuz, C. 2003]**

Der investitionsrechnerische Kundenwert, in: Günter, B.; Helm, S. (Hrsg.): Kundenwert, 2. Aufl., Wiesbaden: Gabler 2003, S. 347.

**[Schmidt, Michael P./Grimm, Monica 2002]**

Data Marts als Basis des Wissensmanagements, in: Hannig, Uwe (Hrsg.): Knowledge Management und Business Intelligence, Berlin [u. a.]: Springer 2002, S. 247-257.

**[Schneider, Ursula 2003]**

Wissensmanagement ist vor allem Kontextmanagement: Lektionen aus den bisherigen Erfahrungen mit Wissensmanagement in Unternehmen, in: Thom, Norbert/Harasymowicz-Birnbach, J. (Hrsg.): Wissensmanagement im privaten und öffentlichen Sektor. Was können beide Sektoren voneinander lernen? Zürich 2003, S. 55-76.

**[Schulte, Ulrich 1993]**

Einführung in Fuzzy Logik. München: Franzis-Verlag 1993.

**[Schulze, Jens 2002]**

CRM erfolgreich führen. Springer-Verlag 2002, S. 7-66.

**[Schumacher, Jörg/Meyer, Matthias 2004]**

Customer Relationship Management strukturiert dargestellt - Prozesse, Systeme, Technologien, Springer 2004.

**[Schüppel, Jürgen 1996]**

Wissensmanagement - Organisatorisches Lernen im Spannungsfeld von Wissens- und Lernbarrieren, Wiesbaden: Gabler 1996.

**[Sexauer, Hagen J. 2004]**

Konzepte des Customer Relationship Management (CRM): Strategien - Instrumente -Umsetzung / Hagen J. Sexauer (Hrsg.), Wiesbaden: Denk!Institut 2004.

**[Smithson, Michael/Verkuilen, Jay 2006]**

Fuzzy Set Theory, Applications in the Social Sciences, Thousand Oaks: Sage Publications 2006.

**[Stäger, C. 1999]**

Multi Channel Management - Mehrdimensionale Optimierung der Kundenbeziehung zur nachhaltigen Steigerung der Profitabilität im Retail Banking, Bern: Haupt Verlag 1999.

**[Steinmüller, W. 1993]**

Informationstechnologie und Gesellschaft. Einführung in die angewandte Informatik, Darmstadt: WBG (Wiss. Buchgesellschaft) 1993.

**[Stewart, T. A. 1998]**

Der vierte Produktionsfaktor: Wachstum und Wettbewerbsvorteile durch Wissensmanagement, München, Wien: Carl Hanser Verlag 1998.

**[Thieme, K. H./Steffen, W. 1999]**

Call Center - Der professionelle Dialog mit dem Kunden, Landsberg/Lech: Verlag Moderne Industrie AG 1999.

**[Tilli, Thomas 1992]**

Fuzzy Logic. Grundlagen, Anwendungen, Hard- und Software, 2. Aufl., München: Franzis Verlag 1992.

**[Tishoosh, Hamid R. 1997]**

Fuzzy-Bildverarbeitung. Einführung in Theorie und Praxis, Berlin: Springer 1997.

**[Töpfer, Armin/Bauer, Ewald/Beckett, Nigel P./Bieletzki, Hubert 1999]**

Kundenzufriedenheit - Messen und Steigern, 2. Aufl., Neuwied: Luchterhand 1999, S. 59.

**[Vavra, Terry. G. 1995]**

Aftermarketing: How to keep customers for life through relationship marketing. Chicago etc.: Irwin-Verlag 1995.

**[Werli, H. P./Jüttner, U. 1994]**

Beziehungsmarketing: Konzepte und Konsequenzen. Arbeitspapier des Instituts für betriebswirtschaftliche Forschung an der Universität Zürich. Zürich.

**[Wiencke, W./Koke, D. 1999]**

Call Center Praxis – Den telefonischen Kundenservice erfolgreich organisieren, 2. Aufl., Stuttgart: Verlag Schäffer-Poeschel 1999.

**[Wieken John-Harry 1998]**

Meta-Daten für Data Marts und Data Warehouses, in: Muksch, Harry; Behme, Wolfgang (Hrsg.): Das Data Warehouse Konzept. Architektur – Datenmodelle – Anwendungen, 3. Aufl., Wiesbaden: Gabler 1998.

**[Wilke, Helmut 2001]**

Systemisches Wissensmanagement, 2. Aufl., Stuttgart: Lucius & Lucius 2001

**[Wirtz, Bernd W. 2000]**

Medien- und Internetmanagement, 1. Aufl., Wiesbaden: Gabler Verlag 2000.

**[Wirtz, Bernd W. 2001]**

Electronic Business, 2. Aufl., Wiesbaden: Gabler Verlag 2001.

**[Zadeh, Lotfi A. 1973]**

The concept of a linguistic variable and its application to approximate reasoning. Memorandum ERL-M411 Berkeley, October 1973.

**[Zahn, E./Greschner, J. 1996]**

Strategische Erneuerung durch organisationales Lernen, in: Bullinger, H. J. (Hrsg.): Lernende Organisationen: Konzepte, Methoden und Erfahrungsberichte, HAB-Forschungsberichte, Band 8, Stuttgart: Schäffer-Poeschel Verlag 1996, S. 41-74.

**[Zahn, Erich/Foschiani, Stefan/Tilebein, Meike 2000]**

Nachhaltige Wettbewerbsvorteile durch Wissensmanagement, in: Hermann Krallmann (Hrsg.): Wettbewerbsvorteile durch Wissensmanagement: Methodik und Anwendungen des Knowledge Management, Stuttgart: Schäffer-Poeschel Verlag 2000, S. 239-270.

**[Zimmermann, Hans-Jürgen 1991]**

Fuzzy Set Theory and its Applications. Second Edition, Boston: Kluwer Academic Publishers 1991.

**[Zimmermann, Hans-Jürgen 1993]**

Fuzzy Technologien: Prinzipien, Werkzeuge, Potenziale. Düsseldorf: VDI-Verlag 1993.

**[Zimmermann, Hans-Jürgen 1995]**

Datenanalyse: Anwendung von DataEngine mit Fuzzy Technologien und Neuronalen Netzen. Düsseldorf: VDL-Verlag 1995.

**[Zingale, Alfredo/Arndt, Matthias 2002]**

Das E-CRM Praxisbuch: Was Sie über Customer Relationship im Internet wissen müssen, Weinheim: Wiley 2002, S. 155-160.

**[Zipser, Andreas 2001]**

Business Intelligence im CRM: Die Relevanz von Daten und deren Analyse für profitable Kundenbeziehungen in: Link, Jörg (Hrsg.): Customer Relationship Management: Erfolgreiche Kundenbeziehungen durch integrierte Informationssysteme, Springer 2001.

**WISSENSCHAFTLICHE FACHZEITSCHRIFTEN****[Bea, F. X. 2000]**

Wissensmanagement, in: WiSt- Das wirtschaftswissenschaftliche Studium, 7/2000, S. 362-367.

**[Bollinger, T. 1996]**

Assoziationsregeln - Analyse eines Data Mining Verfahrens, in: Zeitschrift Informatik-Spektrum, Jg. 19, Heft 5, S. 257 – 261.

**[Bullinger, Hans-J./Warschat, Joachim /Prieto, Juan/Wörner, Kai 1998]**

Wissensmanagement - Anspruch und Wirklichkeit: Ergebnisse einer Unternehmensstudie in Deutschland, in: IM-Fachzeitschrift für Information Management, 13. Jg., 1998, Heft 1, S. 7-23.

**[Diller, H./Kusterer, M. 1988]**

Beziehungsmanagement, in: Marketing, ZfP-Zeitschrift für Forschung und Praxis, 10. Jg., 1988, Heft 3, S. 211-220.

**[Diller, H. 1995]**

Beziehungs-Marketing, in: WiSt-Wirtschaftswissenschaftliches Studium, 24. Jg., 1995, Heft 9, S. 442-447.

**[Diller, H. 1996]**

Kundenbindung als Marketingziel, in: Marketing, ZfP-Zeitschrift für Forschung und Praxis, 18. Jg., 1996, Nr. 2, S. 81-94.

**[Dittmar, Carsten 2000]**

Vom Data zum Knowledge Warehouse – Wissen sichtbar machen, in: Computerwoche extra, Beilage zur Zeitschrift Computerwoche, 16.06.2000, S. 14-17.

**[Eck, C. D. 1997]**

Wissen – ein neues Paradigma des Managements, in: Die Unternehmung, 55. Jg., 1997, Heft 3, S. 155-179.

**[Frank, U./Schauer, H. 2001]**

Software für das Wissensmanagement, in: WISU-Das Wirtschaftsstudium, 6/2001, S. 718-726.

**[Hagedorn, J./Bissantz, N./Mertens, P. 1997]**

Data Mining (Datenmustererkennung): Stand der Forschung und Entwicklung, in: Wirtschaftsinformatik, Heft 6/1997, S. 601-612.

**[Hand, D. J. 1999a]**

Why data mining is more than statistics writ large? In: Bulletin of the International Statistical Institute ISI, 52nd Session, 1999, Vol. 1, S. 433-436.

**[Hasenkamp, U./Rossbach, P. 1998]**

Wissensmanagement, in: WISU-Das Wirtschaftsstudium, 8-9/1998, S. 956-963.

**[Hettich, S./Hippner, H./Wilde, K. D. 2000]**

Customer Relationship Management, in: WISU-Das Wirtschaftsstudium, 29. Jg. (2000), Nr. 10, S. 1346-1366.

**[Homburg, Christian/Daum, Daniel 1997]**

Die Kundenstruktur als Controlling-Herausforderung, in: Controlling, 9. Jg., 1997, Nr. 6, S. 394-405

**[Kantsperger, R. 2002]**

Aufbau und Pflege profitabler Kundenbeziehungen – 4 Prinzipien eines modernen Relationship Marketing, in: FGM Aktuell - Fördergesellschaft für Marketing, Nr. 31/2002, S. 1-2.

**[Krallmann, Hermann/Rieger, Bodo 1987]**

Vom Decision Support System (DSS) zum Executive Support System (ESS), in: Handwörterbuch der modernen Datenverarbeitung (HMD), 24. Jg., Heft 138, 1987, S. 28-38.

**[Krogh G. von/Roos, J. 1996]**

Five Claims on Knowing, in: European Management Journal, Vol. 14, Issue 4, August 1996, S. 423-426.

**[Leßweng, H.-P./Lanninger, V./Thome, R. 2004]**

Betriebliche Standardanwendungssoftware, in: WISU-Das Wirtschaftsstudium, 2004, Nr.2, S. 219-227.

**[Loebbecke, C. 2001]**

eCommerce: Begriffsabgrenzung und Paradigmenwechsel, in: Betriebswirtschaftliche Forschung und Praxis (BFuP), Heft 2/2001, S. S.93-108.

**[Pendse, Nigel/ Creeth, Richard 1995]**

The OLAP Report, in: Business Intelligence, London 1995

**[Rosemann, M./Rochefort, M./Behnck, W. 1999]**

Customer Relationship Management, in: HMD, Nr. 208, 1999, S. 105-116.

**[Schögel, M/Sauer, A 2002]**

Multi-Channel Marketing – Die Königsdisziplin im CRM, in: Thexis, Fachzeitschrift für Marketing der Universität St. Gallen. Nr. 1/2002, S. 26-31.

**[Schramm, Dirk 2001]**

Wie verwaltet man Inhalte? Anforderungen an XML-basierte Content Management Systeme im Electronic Publishing, in: WiSt - Das wirtschaftswissenschaftliche Studium , 30. Jg., 2001, Nr. 11, S. 615-620.

**[Scott, Morton/ Michael S. 1983]**

State of the Art of Research in Management Support Systems, Massachusetts Institute of Technology, Center for Information Systems Research, Working Paper CISR, No. 107, o. O. 1983.

**[Stojek, M. 2000]**

Customer Relationship Management – Software, Prozess oder Konzept?, in: IM – Fachzeitschrift für Information Management & Consulting, 15. Jg. (2000), Nr. 1, S. 37-42.

**[Trix, P. 1999]**

Interactive Selling-Systeme. Die neue Dimension des Verkaufens, in: CAS-Report, Das Salesprofi-Special für CAS und CRM, 2. Jg., Ausgabe 1999, S. 32-35.

**[Wiig, Karl M./De Hoog, Robert/Van der Spek, Rob 1997]**

Supporting Knowledge Management: A Selection of Methods and Techniques, in: Expert Systems with Applications, Publisher Elsevier Ltd., Vol. 13, Issue 1, July 1997, S. 15-27.

**[Wilke, Helmut 2002]**

Trauma Wissensgesellschaft und seine Folgen, in: Wissensmanagement-Zeitschrift für Führungskräfte, Heft 2/2002, 54-55.

**WISSENSCHAFTLICHE ARBEITEN IM INTERNET****[Andreß, Hans-Jürgen 2001]**

Glossar zur Datenerhebung und statistischen Analyse, eine Deutsche Übersetzung der gleichnamigen Arbeit von Nikos Drakos und Ross Moore, auf URL: <http://psydok.sulb.unisaarland.de/volltexte/2004/268/html/index.html>, [Abruf: 25-10-2008].

**[Bensberg, F./Grob, H. L. 1999]**

Das Data-Mining-Konzept, Arbeitsbericht Nr. 8 zu computergestütztes Controlling, Juni 1999, S. 1-20., auf URL: [http://www.wi.uni-muenster.de/aw/forschen/ab\\_cc.html](http://www.wi.uni-muenster.de/aw/forschen/ab_cc.html)  
Direkter Zugriff: <http://www.wi.uni-muenster.de/aw/download/publikationen/CGC8.pdf>, [Abruf: 25-10-2007].

**[CRISP-DM 1.0 2000]**

Step by step Data Mining Guide: Document of the partners of the CRISP-DM consortium: NCR Systems Engineering Copenhagen (USA and Denmark), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringen en Bank Groep B.V (The Netherlands), August 2000, auf URL: <http://www.crisp-dm.org/CRISPWP-0800.pdf>, [Abruf: 15-07-2008].

**[Codd, E. F./ Codd, S. B./ Salley, C. T. 1993]**

Providing OLAP (On-Line Analytical Processing) to User-Analyst: An IT Mandate, Technical Report. E. F. Codd & Associates, 1993, auf URL: [http://dev.hyperion.com/resource\\_library/white\\_papers/providing\\_olap\\_to\\_user\\_analysts.pdf](http://dev.hyperion.com/resource_library/white_papers/providing_olap_to_user_analysts.pdf), [Abruf: 25-10-2007]

**[Conrad, K. W. /Newman, B. 1999]**

A Framework for Characterizing Knowledge Management Methods, Practices, and Technologies 1999, first publication in: The Introduction for Knowledge Management, George Washington University Cours EMGT 298. T1, Spring 1999, S. 1-19, auf URL: <http://www.bruegge.informatik.tu-muenchen.de/static/contribute/Lehrstuhl/documents/KM-Characterization-Framework.pdf>, [Abruf: 10-04-2008].

**[Dangelmaier, Wilhelm/Helmke, Stefan/Uebel, Matthias 2001]**

„Trends im Customer Relationship Management (CRM)“ – Artikel veröffentlicht am 05/2001 auf der Webseite Competence SITE, auf URL: <http://www.competence-site.de/crm.nsf/fbfca92242324208c12569e4003b2580/ebcfddf93ec70f42c1256a6900566f77?OpenDocument&Highlight=0,Kundenbindung>, [Abruf: 24-10-2007].

Direkter Zugriff: [http://www.competence-site.de/crm.nsf/EBCFDDF93EC70F42C1256A6900566F77/\\$File/crm-trends\\_tdbm2001.pdf](http://www.competence-site.de/crm.nsf/EBCFDDF93EC70F42C1256A6900566F77/$File/crm-trends_tdbm2001.pdf)

Dieser Artikel ist ein Auszug aus dem Jahrbuch der Zeitschrift DATABASE MARKETING 2001.

**[Decker/Focardi 1995]**

Technology overview: a report on data mining. Technical Report CSCS TR-95-02, CSCS-ETH, Swiss Scientific Computing Center, 1995, auf URL:

<http://citeseer.ist.psu.edu/73088.html>

Direkter Zugriff:

<http://citeseer.ist.psu.edu/cache/papers/cs/739/http:zSzzSzwww.cs.utk.edu/zSzzSz~cs494zSzreading-listzSzdata-mining-ETH.pdf/decker95technology.pdf>, [Abruf: 25-10-2007].

**[Dilger, Werner 2007]**

Data Mining (Kompaktkurs an der Berufsakademie Mannheim), Vorlesungsskript 2007 der Technischen Universität, Fakultät für Informatik: Chemnitz.

Auf URL: <http://www.tu-chemnitz.de/informatik/KI/skripte.php>

Direkter Zugriff: <http://www.tu-chemnitz.de/informatik/KI/scripts/DM-skr-2.doc>, [Abruf: 01-06-2008].

**[Dittmar, Carsten 2000]**

Vom Data zum Knowledge Warehouse – Wissen sichtbar machen, in: Computerwoche extra, Beilage zur Zeitschrift Computerwoche, 16.06.2000, S. 14-17.

Dieser Artikel ist am 16.05.2001 online veröffentlicht, auf URL:

<http://www.computerwoche.de/index.cfm?pid=707&pk=524090>, [Abruf: 20-12-2007]

**[Fayyad, U. M./Piatetsky-Shapiro, G./Smyth, P. 1996b]**

„From Data Mining to Knowledge Discovery“, in: AI - Zeitschrift für Künstliche Intelligenz, Fall 1996. Artikel, veröffentlicht auf der Webseite AAAI (The American Association for Artificial

Intelligence), auf URL: <http://www.aaai.org/AITopics/assets/PDF/AIMag17-03-2-article.pdf>, [Abruf: 26-10-2007].

**[Fayyad, U. M./Piatetsky-Shapiro, G./Smyth, P. 1996c]**

„The KDD Process for Extracting Useful Knowledge from Volumes of Data“, in: Communications of the ACM, November 1996/Vol. 39, Issue No. 11, auf URL:

<http://citeseer.ist.psu.edu/fayyad96kdd.html>

Direkter Zugriff:

[http://citeseer.ist.psu.edu/cache/papers/cs/20907/http:zSzzSzwwwhome.cs.utwente.nl/zSzzSz~mpoelzSzcollegeszSzwdmzSZACM\\_artikelenzSzfayyad2.pdf/fayyad96kdd.pdf](http://citeseer.ist.psu.edu/cache/papers/cs/20907/http:zSzzSzwwwhome.cs.utwente.nl/zSzzSz~mpoelzSzcollegeszSzwdmzSZACM_artikelenzSzfayyad2.pdf/fayyad96kdd.pdf), [Abruf: 25-10-2007].

**[Gaßen, Helga 1999]**

Wissensmanagement - Grundlagen und IT-Instrumentarium, Arbeitspapiere WI, Nr. 6/1999, (Hrsg.): Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Universität: Mainz 1999. Herausgeber: Univ. Prof. Dr. Axel C. Schwickert, Professur für BWL und Wirtschaftsinformatik, Justus-Liebig-Universität Gießen,

auf URL: [http://geb.uni-giessen.de/geb/volltexte/2004/1628/pdf/Apap\\_WI\\_1999\\_06.pdf](http://geb.uni-giessen.de/geb/volltexte/2004/1628/pdf/Apap_WI_1999_06.pdf), [Abruf: 22-12-2007].

**[Gengler, Barbara 2006]**

„CRM-Trends 2006 - ein Blick in die Kristallkugel“ - Artikel, veröffentlicht auf der Webseite Competence-Site am 06.03.2006, auf URL:

<http://www.dienstleistungsmanagement-competence-center.de/crm.nsf/f1b7ca69b19cbb26c12569180032a5cc/13189e10b417e225c125713e0048587a!OpenDocument>

Direkter Zugriff:

<http://www.sap.info/public/DE/de/index/Category-12613c61affe7a5bc-de/0/articlesVersions-32061440c27210d59b>, [Abruf: 14-11-2007].

**[Glymour, Clark/Madigan, David/Pregibon, Daryl/Smyth, Padhraic 1996]**

Statistical Themes and Lessons for Data Mining, in: Data Mining and Knowledge Discovery, 1, Kluwer Academic Publishers, Boston, Manufactured in The Netherlands 1996, S. 25-42, auf URL: <http://citeseer.ist.psu.edu/77634.html>

Direkter Zugriff:

<http://coblitz.codeen.org:3125/citeseer.ist.psu.edu/cache/papers/cs/6200/http:zSzzSzwww.stat.washington.eduzSzbayeszSzPAPERSzSzdami.pdf/glymour96statistical.pdf>, [Abruf: 26-10-2007].

**[Hafner, Nils 2003]**

„Lernendes Marketing ohne Streuverluste“ - Artikel, veröffentlicht am 01/2003 auf der Webseite Competence SITE, auf URL: <http://www.competence-site.de/marketing.nsf/fbfa92242324208c12569e4003b2580/58121d0d94b00164c1256d780039f859!OpenDocument&Highlight=0,pricing,preispolitik>, [Abruf: 20-05-2007].

Direkter Zugriff: [http://www.competence-site.de/marketing.nsf/58121D0D94B00164C1256D780039F859/\\$File/marketing%20ohne%20streuverluste.pdf](http://www.competence-site.de/marketing.nsf/58121D0D94B00164C1256D780039F859/$File/marketing%20ohne%20streuverluste.pdf)

**[Hand, D. J. 1999b]**

„Statistics and Data Mining: Intersecting Disziplines“, in: SIGKDD Explorations, June 1999, vol. 1, issue 1. S. 16-19, auf URL: <http://citeseer.ist.psu.edu/hand99statistics.html>

Direkter Zugriff:

[http://citeseer.ist.psu.edu/cache/papers/cs/20907/http:zSzzSzwwwhome.cs.utwente.nlSz~mpoelzSzcollegeszSzwdmzSzACM\\_artikelenzSzhand.pdf/hand99statistics.pdf](http://citeseer.ist.psu.edu/cache/papers/cs/20907/http:zSzzSzwwwhome.cs.utwente.nlSz~mpoelzSzcollegeszSzwdmzSzACM_artikelenzSzhand.pdf/hand99statistics.pdf), [Abruf: 25-10-2007].

**[Hand, D. J. 1998]**

„Data Mining: Statistics and More?“, in: The American Statistician, May 1998 Vol. 52, No. 2, S. 112-118, auf URL: <http://citeseer.ist.psu.edu/264891.html>

Direkter Zugriff:

<http://citeseer.ist.psu.edu/cache/papers/cs/12591/http:zSzzSzwww.amstat.orgzSzpublicationszSztaszSzhand.pdf/hand98data.pdf>, [Abruf: 25-10-2007].

**[Helmke, Stefan/Uebel, Matthias 2001]**

„CRM-Grundlagen“ - Artikel veröffentlicht am 02.05.2001 auf der Webseite Competence SITE, auf URL: <http://www.competence-site.de/crm.nsf/44ed936957de26d7c1256911003d7e42/62a30e204a580c89c125694a004db1a5!OpenDocument>, [Abruf: 24-03-2007].

Direkter Zugriff: [http://www.competence-site.de/crm.nsf/62A30E204A580C89C125694A004DB1A5/\\$File/grundlagenartikel.pdf](http://www.competence-site.de/crm.nsf/62A30E204A580C89C125694A004DB1A5/$File/grundlagenartikel.pdf)

**[Hudec, Marcus o. J.]**

Data Mining – Ein neues Paradigma der angewandten Statistik, Institut für Statistik und Decision Support Systems Universität Wien, S. 1-9, auf URL:

<http://www.statistik.tuwien.ac.at/oezstat/ausg021/papers/hudec.doc>, [Abruf: 21-09-2007].

**[Kolbe, Lutz M./Österle, Hubert/Brenner, Walter/Geib, Malte 2003]**

Grundlagen des Customer Knowledge Management, in: Kolbe, L.; Brenner, W.; Österle, H. (Hrsg.): Customer Knowledge Management, Berlin et al.: Springer 2003, S. 3-21, auf URL:

[http://web.iwi.unisq.ch/org/iwi/iwi\\_pub.nsf/wwwAuthorPublGer/8995A7BB72FBB633C1256FE8006210CC/\\$file/CKM\\_Grundlagen.pdf](http://web.iwi.unisq.ch/org/iwi/iwi_pub.nsf/wwwAuthorPublGer/8995A7BB72FBB633C1256FE8006210CC/$file/CKM_Grundlagen.pdf), [Abruf: 07-04-2008].

**[Stein, Benno M./Lettmann, Theodor. 2005]**

Fuzzy Logic, Skript WS 04/05. Universität Paderborn, Institut für Informatik, auf URL:

<http://wwwcs.uni-paderborn.de/cs/aq-klbue/de/courses/ws04/awis/awis04-09-09.pdf>, [Abruf: 20-09-2008]

**[o. V. 2007]**

„Die acht Grundlagen von CRM: Ein Handbuch für Führungskräfte zu den acht unabdingbaren Elementen jeder erfolgreichen CRM-Initiative“ - Whitepaper (2007) von Salesforce.com, auf URL:

[http://s3.amazonaws.com/S3Karma\\_demo\\_A779DC8/31113/eight\\_crm\\_essentials.pdf](http://s3.amazonaws.com/S3Karma_demo_A779DC8/31113/eight_crm_essentials.pdf), [Abruf: 14.11.2007].

**[Pendse, Nigel 2005]**

What is OLAP? (letztlich aktualisiert am 15-10-2005), auf URL:

<http://www.olapreport.com/fasmi.htm>, [Abruf: 25-10-2007].

**[Piller, Frank T./Schoder, Detlef 1999]**

Mass Customization und Electronic Commerce, auf URL:

[http://wim.uni-koeln.de/fileadmin/publikationen/downloads/Piller-Schoder\\_1999\\_Mass-Customizing\\_Empirie.pdf](http://wim.uni-koeln.de/fileadmin/publikationen/downloads/Piller-Schoder_1999_Mass-Customizing_Empirie.pdf), [Abruf: 30-03-2007].

Vorabdruck einer gleichnamigen Veröffentlichung in der ZfB - Zeitschrift für Betriebswirtschaft, Heft 10/1999

**[Puschmann, Thomas/Alt, Rainer R. 2002]**

Benchmarking Customer Relationship Management. Arbeitsbericht, S. 1-42, auf URL:

<http://www.alexandria.unisg.ch/Publikationen/22904>, [Abruf: 10-11-2007].

**[Rehäuser, J./Krcmar, H. 1996]**

Wissensmanagement im Unternehmen, Arbeitspapier von TU-München/Lehrstuhl für Wirtschaftsinformatik, Februar 1996, S. 1-45, auf URL:

[http://www.winfobase.de/lehstuhl/publikat.nsf/intern01/FC0F0EC41403EF3D412566500029C4A5/\\$FILE/96-14.pdf](http://www.winfobase.de/lehstuhl/publikat.nsf/intern01/FC0F0EC41403EF3D412566500029C4A5/$FILE/96-14.pdf), [Abruf: 20-12-2007].

**[Rapp, R./Decker, A. 2000]**

„Herausforderungen und Trends im Customer Relationship Management“, Artikel entnommen aus: Wilde, K. D./Hippner, H. (Hrsg.): Marktstudie CRM 2000 – Customer Relationship Management – So binden Sie Ihre Kunden, Herausgeber: absatzwirtschaft - Zeitschrift für Marketing u. Vertrieb, Düsseldorf, S. 73-77; auf URL:

<http://www.competencesite.de/dienstleistung.nsf/f1b7ca69b19cbb26c12569180032a5cc/0bc434b98247dc95c1256af5003aa797!OpenDocument>, [Abruf: 20-03-2007].

Direkter Zugriff:

[http://www.competencesite.de/dienstleistung.nsf/0BC434B98247DC95C1256AF5003AA797/\\$File/herausforderungen\\_und\\_trends\\_im\\_crm.pdf](http://www.competencesite.de/dienstleistung.nsf/0BC434B98247DC95C1256AF5003AA797/$File/herausforderungen_und_trends_im_crm.pdf), S. 2.

**[Schneider, Lars 2004]**

„Die häufigsten Stolpersteine bei der CRM-Einführung“ - Artikel, veröffentlicht am 11/2006 auf der Webseite Competence-Site, auf URL:

<http://www.competencesite.de/crm.nsf/f1b7ca69b19cbb26c12569180032a5cc/2c439eb83c26caa3c125723000526a77!OpenDocument>

Direkter Zugriff:

[http://www.competencesite.de/crm.nsf/2C439EB83C26CAA3C125723000526A77/\\$File/schneider\\_s\\_tolpersteine\\_einfuehrung\\_crm.pdf](http://www.competencesite.de/crm.nsf/2C439EB83C26CAA3C125723000526A77/$File/schneider_s_tolpersteine_einfuehrung_crm.pdf), [Abruf: 10-11-2007]

**[Schreyögg, Georg/Geiger, Daniel 2003]**

Kann die Wissensspirale die Grundlage des Wissensmanagements sein? Arbeitspapier von der Freien Universität Berlin/Institut für Management 04/2003, S. 1-31, auf URL: <http://www.fu-berlin.de/wiwiss/institute/management/schreyoegg/publikationen/arbeitspapiere/index.html>

Direkter Zugriff:

[http://www.fu-berlin.de/wiwiss/institute/management/schreyoegg/publikationen/pdf/Wissensspirale\\_Schreyoegg\\_Geiger.pdf](http://www.fu-berlin.de/wiwiss/institute/management/schreyoegg/publikationen/pdf/Wissensspirale_Schreyoegg_Geiger.pdf), [Abruf: 25-12-2007.]

**[Schwetz, Wolfgang 2007]**

„Kundenorientierung ist Chefsache“ - Artikel, erschien am 03/2007 auf der Webseite

Computerwoche, auf URL: <http://www.computerwoche.de/heftarchiv/2007/03/1217439/>, [Abruf: 05-11-2007].

**[Schwetz, Wolfgang o. J.]**

„CRM - Der Weg zum kundenorientierten Unternehmen“ - Artikel, auf URL:

[http://www.isis-specials.de/profile\\_pdf/1s359\\_ed\\_crm0207.pdf](http://www.isis-specials.de/profile_pdf/1s359_ed_crm0207.pdf), [Abruf: 05-11-2007].

**[Sexauer, Hagen J. 2002]**

Entwicklungslinien des Customer Relationship Mngement (CRM), in: Wirtschaftswissenschaftliches Studium (WiSt), 31. Jg. (2002), Nr. 4, S. 218-222.,

auf URL: <http://www.denkinstytut.de/CRMportal/crmportal.html>

Direkter Zugriff:

[http://www.denkinstytut.de/CRMportal/Entwicklung\\_CRM\\_Sexauer\\_WiSt.pdf](http://www.denkinstytut.de/CRMportal/Entwicklung_CRM_Sexauer_WiSt.pdf), [Abruf: 25-11-2007].

**[Winkelmann, Peter 2002]**

„Zehn Bausteine für den CRM-Erfolg“ - Artikel, erschien am 07.11.2003 auf der Webseite

Computerwoche, Autor(in): Karin Quack, auf URL:

<http://www.computerwoche.de/index.cfm?pid=387&pk=542333>, [Abruf: 15-11-2007].

**[Zellner, Georg 2002]**

Beziehungsmanagement im Fokus – Ergebnisse einer empirischen Untersuchung. Arbeitsbericht, S. 1-23, auf URL: [http://www.alexandria.unisg.ch/publications/person/Z/Gregor\\_Zellner/24181](http://www.alexandria.unisg.ch/publications/person/Z/Gregor_Zellner/24181), [Abruf: 10-11-2007].

**DIPLOMARBEIT****[Scheir, Peter 2002]**

Wissensmanagement zur Unterstützung von Kundenbeziehungsmanagement: Konzepte, Technologien und ein Prototyp. Magisterarbeit an der Technischen Universität Graz, Institut für Informationsverarbeitung und Computergestützte neue Medien, Graz 2002.

**[Schmidt, Jens o. J.]**

Methoden des Wissensmanagement zur Unterstützung des Customer Relationship Management. Diplomarbeit, Fachbereich Wirtschaftswissenschaften - Johann Wolfgang Goethe Universität, Frankfurt am Main, o. J., auf URL: <http://www.marketing.ch/wissen/crm/diplomcrm.pdf>, [Abruf: 10-12-2007].

**[Schramm, D. 2000]**

Content Management mit SGLM-Datenbanken, Diplomarbeit an der Technischen Universität Dresden, Fakultät Wirtschaftswissenschaften, Dresden 2000, S. 3.

**DISSERTATION****[Reichold, Annette 2006]**

Prozesse des analytischen CRM: Fallbeispiele aus der Finanzdienstleistungsbranche, Architekturvorschlag und Methodenelemente, Bamberg 2006, Dissertation.

**[Zeller, Andreas 2003]**

Technologiefrühaufklärung mit Data Mining – Informationsprozessorientierter Ansatz zur Identifikation schwacher Signale, Wiesbaden: Deutscher Universitäts-Verlag 2003. Dissertation Universität Stuttgart, 2003: D 93 - Informationsgewinnung und -verarbeitung zur Identifikation schwacher Signale mit Data Mining im Rahmen der Technologiefrühaufklärung.

**[Heiko, Timm 2002]**

Fuzzy-Clusteranalyse: Methoden zur Exploration von Daten mit fehlenden Werten sowie klassifizierten Daten, Doktorarbeit an der Fakultät für Informatik der Otto von Guericke Universität-Magdeburg, angenommen am 21.06.2002, auf URL: <http://diqlib.uni-magdeburg.de/Dissertationen/2002/heitimm.pdf>, [Abruf: 01-10-2008].

**LIZENTIATSARBEIT****[Knechtenhofer, Oliver 2001]**

Möglichkeiten und Grenzen einer Integration von CRM- und ERP-Systemen. Institut für Wirtschaftsinformatik der Universität Bern, Lizentiatsarbeit, Oktober 2001, auf URL: <http://www.ie.iwi.unibe.ch/forschung/erp/resource/Knechtenhofer.pdf>, [Abruf:02-03-2007].

**[Mühlethaler, Barbara 2005]**

Wissensmanagement. Institut für Organisation und Personal. Lizentiatsarbeit eingereicht der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Bern am 7. Februar 2005, auf URL: <http://www.iop.unibe.ch/lehre/lizentiatsarbeiten/Liz-Muehlethaler-Barbara.pdf>, [Abruf: 02-01-2008].

**MARKTSTUDIEN****[Englbrecht, A./Hippner, H./Wilde, K. D. 2005]**

Customer Relationship Management: Strategie und IT-Systeme, in: Marktstudie CRM 2005, S. 11-30, auf URL: [http://www.marktstudie.de/downloads/leseprobe\\_crm2005.pdf](http://www.marktstudie.de/downloads/leseprobe_crm2005.pdf), [Abruf: 25-03-2007].

## SEMINARARBEIT

### **[Kempf, Björn 2001]**

Grundlagen der Fuzzy Logik, Seminararbeit an der Technischen Universität Clausthal WS 01/02, auf URL: [http://www2.in.tu-clausthal.de/~reuter/ausarbeitung/Bjoern\\_Kempf - Fuzzy-Logik.pdf](http://www2.in.tu-clausthal.de/~reuter/ausarbeitung/Bjoern_Kempf_-_Fuzzy-Logik.pdf) , [Abruf: 20-09-2008]

## NÜTZLICHE WEBSEITEN

*Scientific Commons*

<http://de.scientificcommons.org/>

*CRM-Marketing Management Portal*

[www.themanagement.de](http://www.themanagement.de)

*CiteSeer.IST – Scientific Literatur Digital Library*

<http://citeseer.ist.psu.edu/cs>

*KDnuggets – Data Mining & Knowledge Discovery*

<http://www.kdnuggets.com/>

*The Data Warehouse Institute*

<http://www.dw-institute.com/>

*COMPUTERWOCHE.de*

[www.computerwoche.de](http://www.computerwoche.de)

*absatzwirtschaft online*

[www.absatzwirtschaft.de](http://www.absatzwirtschaft.de)

*Competence-SITE*

<http://www.competence-site.de/>

*The OLAP Report – Webseite für die Analysen über das Thema OLAP von N. Pendse*

<http://www.olapreport.com/>

*Business Intelligence – Webseite von Business Intelligence Ltd.*

<http://www.busintel.com/>

*The Data Warehousing Information Center*

<http://www.dwinfocenter.org:80/>

*The ACM Digital Library*

<http://portal.acm.org/portal.cfm>

*IEEE-Xplore Digital Library*

<http://ieeexplore.ieee.org/Xplore/questhome.jsp>

*CRISP-DM (Cross Industry Standard Process for Data Mining)*

<http://www.crisp-dm.org/>

## **10 ANHANG**

In dem beiliegenden CD-ROM sind die Quellcodes des selbst geschriebenen Programms in Java für das Fallbeispiel, die Datentabelle für 200 fiktive Kunden und die hergestellten Diagramme für die Analyse in MS-Excel-Format zu finden.