

# **Databasemanagement**

## **Erfolgsfaktoren für die Arbeit mit der Datenbank**

Autorin: Andrea Ahlemeyer-Stubbe

Erschienen in: DDV Jahrbuch 2001

### **Aufbau eines Datawarehouses oder einer Marketing Datenbank**

Leider ist es heute immer noch so, dass in vielen Firmen zwar viele Daten im operativen Geschäft entstehen, diese aber in unterschiedlichsten Formen in den einzelnen Bereichen und Abteilungen gespeichert sind. Dadurch sind sie für andere Abteilungen oder unternehmensübergreifend nicht oder nur schwer nutzbar.

Ein weiteres Problem stellt die Tatsache dar, dass viele Daten aufgrund der Art und Weise, wie sie derzeit gespeichert werden, zum Teil nur schwer analysierbar sind. Das liegt zum einen an der Datenhaltung, die zumeist für die Erledigung des operationalen Tagesgeschäftes optimiert wurde. Zum anderen liegt es häufig auch daran, dass der Wert einzelner Daten für die Optimierung der Kundenkommunikation nicht immer erkannt wird.

Ziel eines erfolgreichen Databasemanagements ist es, die unterschiedlichen Daten zum Kunden so zu sammeln und zu speichern, zu verwalten und zu analysieren, dass daraus Informationen entstehen, die helfen, den Kunden in differenzierte Zielgruppen einzuteilen und eine optimale Kommunikation aufzubauen.

Wenn sich ein Unternehmen mit dem Gedanken trägt, ein Datawarehouse oder eine Marketingdatenbank aufzubauen, stehen drei Fragen im Zentrum:

- **Warum?**

Das Ziel ist allgemein gültig: Mit Hilfe der verschiedenen Analysen unternehmenseigene und -fremde Daten in Informationen, Wissen und Strategien zu verwandeln. Denn nur verlässliche Daten und Informationen geben die notwendige Planungssicherheit. Erst wenn man genau weiß, was der Kunde, wann, wo und wie gekauft oder geäußert hat, kann man ihm auch in Zukunft ein maßgeschneidertes Angebot machen. Das minimiert Streuverluste und lässt eventuelle Abwanderungsgedanken rechtzeitig erkennen.

- **Für wen?**

Die Datenbank soll diese Informationen für die Fachabteilungen (Marketing) und das Management zur Verfügung stellen. Besonders die Fachabteilungen müssen sich für ihre Analysen mit der Datenbank zurechtfinden. Für die einzelnen Informationsbedürfnisse und die Beherrschung von Analysen und Analysewerkzeugen müssen unterschiedlich aufgearbeitete Daten und Datenaggregate zur Verfügung stehen.

Ein Beispiel: ein Analyst benötigt für eine Warenkorbanalyse Zugriff auf die Einzelinformationen eines Kaufes. Für einen Marketing-Mitarbeiter ist es bei einem Vorjahresvergleich einzelner Mailingaktionen hingegen einfacher, wenn er auf bereits kumulierte Werte zurückgreifen kann.

- **Wie?**

Hier gibt es vier unterschiedliche Möglichkeiten der Umsetzung.

- Umsetzung ausschließlich mit eigenen Leuten,
- Vergabe an einen Dienstleister,
- Kauf eines fertigen Produktes,
- oder Zusammenarbeit mit einer intern und extern besetzten Projektgruppe.

### **Die Suche nach der besten Lösung**

Welche Lösung die beste ist, hängt von der jeweiligen Unternehmenssituation, -größe, dem Know-how, dem Datenvolumen und der angestrebten Komplexität der Datenbank ab. Meist verspricht der Ansatz mit einer intern und extern besetzten Projektgruppe den größten Erfolg. Dieser garantiert, dass neben den zusätzlichen Kenntnissen und Erfahrungen der externen Berater auch die Unternehmens- und branchentypischen Besonderheiten in der Datenbank berücksichtigt werden.

Ein Datawarehouse ist eine unternehmensweite Lösung. Im Gegensatz dazu enthalten die sogenannten Datamarts nur Teile der Aufgaben eines Datawarehouse. Unter einem Datamart versteht man eine spezifische Datensammlung, in der nur die Bedürfnisse der jeweiligen Datensicht, zum Beispiel einer Abteilung, abgebildet werden. Mit den Datamarts lässt sich eine unternehmensweite Strategie nur schwer ableiten, da die Ergebnisse einzelner Bereiche kaum vergleichbar sind.

### **Ein Beispiel:**

Wie ist ein unternehmensweiter Vergleich möglich, wenn der Einkauf seine Artikelstatistiken nicht nach Gebieten unterteilt, der Verkauf aber nach Regionen strukturiert ist. Dass innerhalb der verschiedenen Bereiche auch noch unterschiedliche Ansichten darüber bestehen, was beispielsweise unter Umsatz zu verstehen ist, merken viele Unternehmen häufig erst, wenn sie versuchen, ein Datawarehouse aufzubauen.

Auf der anderen Seite sind viele Datawarehouse-Projekte auch daran gescheitert, dass versucht wurde, gleich ein unternehmensweites Datawarehouse zu bauen. Solche Projekte sind häufig erst nach 2 bis 3 Jahren abgeschlossen. Damit man sie überhaupt zu Ende führen kann, definiert man während des Projektes einen Zeitpunkt, ab dem keine Veränderung in den Geschäftsprozessen oder Daten mehr bis zum Abschluss des Projektes berücksichtigt werden.

Das heißt im ungünstigsten Fall, dass ein Teil der im Modell abgebildeten Geschäftsprozesse nicht mehr vorhanden sind und neue hinzukommen, die in der Datenbank nicht abgebildet werden. Neben diesem Problem ist es für ein Unternehmen auch wirtschaftlich sehr wichtig, dass zumindest in Teilbereichen so früh wie möglich Nutzen aus der entwickelten Datenbank gezogen wird.

### **Das Gesamtprojekt zerlegen**

Deshalb bietet sich folgende Lösung an: Man zerteilt das Gesamtprojekt in einzelne Phasen. Diese orientieren sich an der Bereichs- oder Informationsstruktur des Unternehmens, wobei das Gesamtprojekt schon in Phase I im Auge behalten werden sollte. Informationen, die alle Bereiche eines Unternehmens gleichzeitig benötigen, werden schon in Phase I realisiert, während in späteren Phasen die bereichstypischen Informationen ergänzt werden.

Konkret heißt das: Man startet mit dem Bereich Marketing und entwickelt eine kleine Lösung, die sich vor allen am Informationsbedürfnis des Marketing-Bereiches orientiert. Es werden also ausführliche Informationen und Prozesse rund um den Kunden und seine Kauftransaktionen abgebildet (zum Beispiel Alter, Geschlecht, erhaltene Werbemittel/ Besuche, aktuelle oder historische Kaufdaten). Dagegen werden detaillierte Produktinformationen wie besondere Herstellungs- oder Lieferdaten, (zum Beispiel Lieferantenartikelnummer, Lagerplatz, Gewicht, Volumen), erst einmal nicht berücksichtigt. Gleichzeitig werden im Datawarehouse Schnittstellen eingebaut, die eine spätere Erweiterung ermöglichen.

Das Gesamtunternehmen kann also von Anfang an über einen allgemeinen Teil der Informationen verfügen. Alle Abteilungen können zum Beispiel Tagesvergleiche mit dem Vorjahr erheben. Marketing und Vertrieb erhalten zusätzlich bereichstypischen Informationen. Durch diese Vorgehensweise bleibt das Projekt in jeder Phase überschaubar und bezahlbar. Zudem können schon mit Abschluss der ersten Phase Informationen und Gewinn aus dem Datawarehouse gezogen werden.

Damit ein Datawarehouse seine Aufgaben erfüllt, ist es hilfreich, wenn sich der Unternehmensvorstand mit der strategischen Ausrichtung und den strategischen Erwartungen an die Datenbank selber auseinandersetzt. Wenn ein Mitglied der Geschäftsleitung die Patenschaft für das Projekt übernimmt, ist sichergestellt, dass der Kontakt zwischen Projektteam und Geschäftsführung nicht verlorengeht. Andererseits müssen auch die Fachabteilungen von Anfang an einbezogen werden.

Ein Datawarehouse-Projekt kann nur dann Erfolg haben, wenn alle Beteiligten, besonders die späteren Nutzer, in das Projekt eingebunden sind beziehungsweise in den entsprechenden Phasen unterstützen. Denn der Erfolg eines Datawarehouse misst sich vor allem an der Zufriedenheit seiner Benutzer.

Leider steht viel zu oft die Auswahl der Produkte im Vordergrund der Projekte. Von einer koordinierten, abteilungsübergreifenden Planung unter Mitarbeit der Fachabteilungen und Anwender sind viele Datawarehouse Projekte weit entfernt.

### **Die Qualität der Daten**

Grundsätzlich wird ein Datawarehouse als eine Datensammlung zum Zweck der Analysen angesehen. Aus den unterschiedlichen operativen Systemen eines Unternehmens werden die relevanten Daten für das Datawarehouse auf ihre Qualität geprüft und dann (gegebenenfalls aufbereitet) geladen. Im Gegensatz zu den operativen Systemen spielt die Abbildung von Vergangenheitsdaten und externen Daten im Datawarehouse eine große Rolle.

Aber das Sammeln allein hilft nur bedingt weiter. Denn die Datenqualität des Datawarehouse bestimmt nachhaltig seinen Nutzen. Deshalb sind besonders die Phasen der Analyse der bestehenden Datenbestände und die Phase der Definition der Datenqualität von entscheidender Bedeutung. Fehler in den Datenbeständen wie Rechnungen ohne Kundenstammdaten (der Kundenstammdatensatz wurde gelöscht, da zum Beispiel der Kunde seit x Jahren inaktiv ist, die Rechnung muss aber noch auf Grund der gesetzlichen Bestimmungen archiviert werden), Dubletten, unvollständige und inkonsistente Daten können die Aussagekraft der späteren Analysen in Frage stellen.

Deswegen ist es notwendig, in diesen Phasen besonderes die anderen Fachabteilungen mit einzubeziehen und einen Qualitätsstandard zu erarbeiten, der nicht nur bei der Erstbefüllung sondern auch im späteren Betrieb der Datenbank Bestand hat.

In dieser Phase des Projektes ergibt sich häufig der Effekt, dass es eine bereinigende Rückkoppelung in Richtung der operativen Systeme gibt. Zum Beispiel kann eine Verbesserung des Matchcodesystems zur dublettenfreien Kundenneuanlage durchgeführt werden. Oder die Daten der Interessenten oder Katalogretouren werden künftig permanent gespeichert.

Besonders das Thema Dubletten, der Umgang mit inaktiven Kunden und Interessenten, birgt Schwierigkeiten. Im operativen Geschäft sind Dubletten häufig kein Problem. Das gleiche gilt für inaktive Kunden. Beide haben keinen oder nur wenig Einfluss auf den laufenden Betrieb. Deshalb haben viele Firmen diese in der Vergangenheit nicht weiter beachtet. Für die Selektion eines Mailings ist es aber sehr wohl wichtig, möglichst viele Informationen zu inaktiven Kunden zu haben. Bei Dubletten sind die Probleme noch offensichtlicher. Der mehrfache Versand von ein und demselben Mailing an den gleichen Kunden ist teuer und verärgert den Kunden. Man sollte also ein solches Projekt auch nutzen, um die bestehenden Prozesse des operativen Geschäftes zu untersuchen und gegebenenfalls zu verbessern.

### **Befüllung und tägliche Nutzung des Datawarehouses**

Die Befüllung der Datenbank erfolgt zum ersten Mal während des Projekts als Initialbefüllung. Das bedeutet, dass alle für die Datenbank als relevant definierten Daten in die Datenbank geladen werden. Während des Ladens wird ein Teil der Daten noch gemäß der Definitionen transformiert und aggregiert. Aus der Anrede wird beispielsweise eine zusätzlich Variable [Geschlecht] erstellt.

Darüber hinaus werden Routinen implementiert, die regelmäßige Befüllung sicherstellen. Je nach Einsatzgebiet entscheiden sich die Firmen meistens für einen täglichen oder wöchentlichen Rhythmus.

Bei den Analysen sollte man grundsätzlich zwischen den kreativ forschenden Analysen des Data-Mining und den standardisierbaren regelmäßig wiederkehrenden Analysen unterscheiden. Letztere, zum Beispiel Werbeerfolgskontrollen, Tagesstatistiken oder typische Fragestellungen des Berichtswesen, erfolgen häufig mit sogenannten OLAP-Tools.

So stellt es kein Problem dar, die Anzahl oder das Durchschnittsalter neuer Kunden oder Interessenten auszugeben, während die Nutzung der Daten zur Beantwortung entscheidungsrelevanter Fragen einen großen strategischen Vorteil darstellt.

Beispiele hierfür sind:

- Welchen Kunden sollte wann welches Angebot unterbreitet werden?
- Bei welchem Kundenprofil lohnt sich ein Außendienstbesuch?
- Welche Kunden sind gefährdet?
- Wie hoch ist das Cross-Selling- Potenzial für ein neues Produkt?
- Welcher Lifetime-Profit lässt sich mit welchem Kunden erzielen?
- Wie lassen sich Top-Interessenten mit hohen Lifetime-Values gewinnen?
- Welcher Umsatz wird im nächsten Jahr erzielt?
- Kernfrage ist: Warum blieben gerade diese, für das Management entscheidende, Fragen häufig unbeantwortet?

Betrachtet man den Charakter der Fragestellungen, so liegt die Antwort auf die aufgezählten Fragen nicht in einem einzelnen Datenfeld oder einem Kundenmerkmal, sondern in der richtigen Kombination unterschiedlicher Kundeninformationen. So kann die Angebotsaffinität eines Kunden von einer Vielzahl von Merkmalen wie Alter, Geschlecht, Familienstand,

demographischen Typologien, bisher gekauften Produkten, gezeigtem Produktinteresse, Zahlungsmoral und einer Reihe weiterer Eigenschaften abhängen.

### **Mit Data-Mining Wettbewerbsvorteile erreichen**

An dieser Stelle setzt das Data-Mining an. Data-Mining wird nicht nur von der absoluten Mehrheit der Unternehmen als strategisch angesehen, sondern ist in vielen Fällen sogar ein unternehmensinternes Geheimnis. Mit Data-Mining-Anwendungen will man Wettbewerbsvorteile erreichen. Dass man schon aktiv ist, will man seinen Mitbewerbern jedoch gern verschweigen. Data-Mining ist heute die Geheimwaffe im Wettlauf um Marktanteile. Über 80 Prozent der produktiv eingesetzten Data-Mining-Lösungen stammen aus dem Bereich Marketing.

Beim Data-Mining unterscheidet man grundsätzlich zwei Vorgehensweisen:

- Validieren von Hypothesen auf den Daten
- Entdecken und Entwickeln von bisher unbekanntem Mustern oder Regeln in den Daten

Während für die erste Vorgehensweise schon vorab eine entsprechende Hypothese entwickelt werden muss, versucht man in der zweiten Vorgehensweise, neue Muster und Regeln in den Daten mit Hilfe der geeigneten Verfahren zu finden. Dabei gilt: Nur Informationen, die in den Analyseverfahren zur Verfügung gestellt werden, können auch bei der Entwicklung der neuen Regeln berücksichtigt werden.

Als Beispiel: Kein Verfahren kann einen Zusammenhang zwischen Geschlecht und Kaufverhalten feststellen, wenn in dem zu analysierenden Datenbestand keine Variable für Geschlecht, Anrede oder die Vornamen der Kunden abgelegt sind, oder der gesamte Datenbestand nur Kunden eines Geschlechts enthält. Durch die entsprechende visuelle Aufbereitung fällt es vielen Nutzern leichter, die häufig sehr komplexen Zusammenhänge zu verstehen und daraus die richtigen Strategien zu entwickeln. Technisch ist man heute soweit, dass man Kundenverhalten genau vorhersagen kann.

Database-Marketing und Data-Mining gehören unverzichtbar zusammen. Häufig werden die Begriffe auch synonym verwendet. Besonders in der Anfangszeit des Database-Marketing (80er und frühen 90er Jahre) war es üblich, nur von Database-Marketing zu sprechen und die Analysen, wie Sie heute unter OLAP und Data-Mining fallen, als Teil des Database-Marketing zu betrachten. Heute geht die Entwicklung dahin, dass man Database-Marketing mehr als Anwendung beziehungsweise Strategie auf Basis der durch die Analysen gefundenen Ergebnisse für die jeweiligen Zielgruppen betrachtet, während sich OLAP und Data-Mining zu eigenständigen Disziplinen entwickelt haben.

### **Breites Angebot an Basissoftware**

Diese Entwicklung wurde nicht zuletzt dadurch erleichtert, dass mit den heute in großer Zahl vorhandenen Softwarepaketen zu OLAP und Data-Minings, die früher in den Firmen mit Database-Marketing üblichen selbst geschriebenen Anwendungen/Analyseprogramme nach und nach ersetzt werden. Ein zweiter positiver Aspekt der entwickelten Softwarepakete ist, dass die Analytiker und Anwender heute mehr Zeit haben, sich mit den Analysen und ihrer Ergebnisse auseinanderzusetzen, damit dieses Wissen zur Optimierung der Geschäftsprozesse genutzt werden kann.

### **Der Aufbau eines Datawarehouses**

Es empfiehlt sich neben der Hard- und Softwareauswahl schrittweises Vorgehen, um ein Datawarehouse zu erstellen. Dies gilt auch, wenn »nur« eine Marketing-Datenbank als Datamart angelegt werden soll.

1. Erarbeiten der strategischen Ziele des Datawarehouses
2. Klärung des Informationsbedarfs aller späteren Anwender zur Bewältigung ihres Tagesgeschäfts und zur Umsetzung der strategischen Ziele
3. Analyse der bestehenden und zukünftigen Geschäftsprozesse des Unternehmen und der Datenbestände des operativen EDV-Systems sowie die Analyse möglicher externer Daten
4. Definition der Datenqualität für das Datawarehouse
5. Erarbeiten und Implementierung des Datenmodells und Definition der Metadaten (Metadaten sind vergleichbar mit einem sehr ausführlichen Inhalts- oder Quellenverzeichnis des Datawarehouses). Dieses Datenmodell muss ausbaufähig, flexibel und den Geschäftsprozessen angepasst sein.
6. Aufbereitung und Übernahme der Daten aus den operativen Systemen ins Datawarehouse für des Prototyp-Datawarehouse
7. Testen des Datawarehouse-Prototyps durch die Fachabteilungen und gegebenenfalls Überarbeitung und Weiterentwicklung
8. Aufbereitung und Übernahme der Daten aus den operativen Systemen ins Datawarehouse für das Echt-System
9. Tuning des Datawarehouses für die jetzt bekannten hauptsächlichlichen Abfragen und Analysen
10. Nach Inbetriebnahme des Datawarehouse muss eine konstante und konsequente Anpassung des Datawarehouses an die sich entwickeln Geschäftsprozesse und Aufgaben/Analysen erfolgen.

### **Daten, die ins Datawarehouse gehören**

Im Folgenden finden sich typische Informationen für die einzelnen Unternehmensbereiche. Die Informationen variieren natürlich von Branche zu Branche, doch die hier aufgezählten sollten in jeder Firma vorhanden sein und die Datenbasis für Database-Marketing und CRM bilden.

Eine größere Informationstiefe ergibt sich besonders im BtoB-Bereich. da man hier neben den Informationen zum Kunden (zum Beispiel Firma A) auch noch Informationen zur Person des Bestellers (Angestellter/Chef der Firma A) sammeln kann. Bekanntermaßen ist dies ein recht heikler und schwieriger Punkt, da die Ansprechpartner häufig wechseln. Der Übersichtlichkeit halber wurde diese Ebene im folgenden Beispiel nicht weiter vertieft.

- **Der Bereich Buchhaltung/Finanzen/ Controlling:**
  - alle Einzelinformationen zur Zahlung wie:
    - Zahlungsbetrag,
    - Zahlungstermin,
    - Skontoabzug,
    - Vollständigkeit der Zahlung
    - Art und Anzahl der Mahnungen
  - Name und Anschrift des Zahlenden,
  - Kundennummer,
  - Rechnungsnummer, usw.

- **Der Bereich Auftragsannahme/Kundenservice:**
  - Fragen zum Angebot und Produktinformationen
  - Beschwerden und Reklamationen
  - Anfragen zu Rechnungen, Gutschriften und Kontoauszügen
  - alle Einzelinformationen zum Auftrag wie:
    - Auftraggeber (Kundennummer, Name, Adresse, Geburtstag),
    - Auftragsdatum,
    - Liefertermin,
    - Lieferadresse,
    - Auftragswert,
    - Stückzahlen,
    - Art und Anzahl der verschiedenen Produkte/Dienstleistungen,
    - vom Kunden gewünschte Zahlungsart,
    - in Anspruch genommene Rabatte,
    - Art der Bestellung (telefonisch, Fax, Email, Auftragsnummer), usw.
  
- **Der Bereich Logistik/Auslieferung/Einkauf**
  - alle Einzelinformationen zur Auslieferung wie:
    - Rechnungsdatum,
    - Rechnungsnummer,
    - Auftragsnummer,
    - Kundennummer,
    - Rechnungsnehmer,
    - Rechnungsbetrag,
    - Art und Anzahl der Posten.
    - Vollständigkeit/Teillieferung,
    - Auslieferungsart (Post. Eigenspedition, UPS), usw.
  - alle Einzelinformationen zum Produkt
  
- **Der Bereich Werbung/Vertrieb/Marketing**
  - alle Einzelinformationen zur Bewerbung des Kunden wie:
    - Datum,
    - Art und Anzahl des versandten Werbematerials,
    - Anzahl und Dauer der Außendienstbesuche,
    - Informationen aus dem persönlichen Kontakt zwischen dem Außendienstler und dem Kunden (anstehende Jubiläen, runder Geburtstag. Wünsche zu Produktveränderungen, anstehende Pläne (Neubau nächstes Jahr),
  - Zielgruppe,
  - Kundennummer, usw.