

Databasemanagement

Herausforderung Datenmanagement

Erfolgsfaktoren für die Arbeit mit der Datenbank

Andrea Ahlemeyer-Stubbe

Management Summary

Beim Aufbau einer unternehmensweiten Datenbank gilt es, den Gesamtüberblick zu behalten - über Risiken, Chancen und Prozesse. Als Einstieg in diese komplexe Materie skizziert dieser Artikel die notwendigen Schritte von der Projektinitialisierung bis zur Inbetriebnahme. Der Fokus liegt dabei besonders auf den Bedürfnissen des Database Marketing und des CRM.

Leider ist es heute immer noch so, dass in vielen Firmen zwar viele Daten im operativen Geschäft entstehen, aber diese Daten sind in den unterschiedlichsten Formen in den einzelnen Bereichen und Abteilungen gespeichert und sind dadurch für andere Abteilungen z.B. Marketing und Vertrieb oder unternehmensübergreifend nicht oder nur schwer nutzbar. Ein weiteres Problem stellt die Tatsache dar, dass viele Daten aufgrund der Art und Weise, wie sie jetzt 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.

Es geht also darum, die unterschiedlichen Daten zum Kunden so zu sammeln, zu speichern, zu verwalten und zu analysieren, dass daraus Informationen entstehen, die helfen Kunden in differenzierte Zielgruppe einzuteilen und eine optimale Kundenkommunikation aufzubauen.

Eine Marketing Datenbank – Warum, für wen und wie?

Wenn ein Unternehmen sich mit dem Gedanken trägt ein Datawarehouse oder eine Marketing Datenbank aufzubauen, stehen Fragen nach dem **Warum**, nach dem „Für **wen**“ und nach dem **Wie** im Zentrum.

Warum?

Um 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 dem Kunden auch in Zukunft ein maßgeschneidertes Angebot machen, Streuverluste minimieren und eventuelle Abwanderungs-Absichten des Kunden rechtzeitig erkennen und abfangen.

Für Wen?

Für die Fachabteilungen und das Management.

Die Datenbank soll diesen entscheidungsrelevante Informationen zur Verfügung stellen. Besonders die einzelnen Fachabteilungen mit ihren Analysen (Data Mining, OLAP, ...) müssen sich mit der Datenbasis d.h. mit der Datenbank zurechtfinden. Für die einzelnen Informationsbedürfnisse und die unterschiedliche Beherrschung von Analysen und Analysewerkzeugen müssen unterschiedlich aufgearbeitete Daten bzw. Datenaggregate zur Verfügung stehen.

Ein Analyst benötigt z. B. für eine Warenkorbanalyse Zugriff auf die Einzelinformationen eines Kaufes, während ein Mitarbeiter im Marketing für einen Vorjahresvergleich einzelner Mailing-Aktionen am besten schon auf vorkumulierte Werte zurückgreifen können muss, um effizient zu arbeiten.

Wie?

Hier gibt es vier sehr unterschiedliche Möglichkeiten der Umsetzung.

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

Die optimale Lösung berücksichtigt die spezifische Unternehmenssituation, die Unternehmensgröße, das vorhandene Know-how sowie das Datenvolumen bzw. die angestrebte Komplexität der Datenbank

Am erfolgversprechendsten ist der Ansatz mit einer intern und extern besetzten Projektgruppe. Dadurch wird garantiert, dass neben den zusätzlichen Kenntnissen und Erfahrungen der externen Berater auch die unternehmens- und branchentypischen Besonderheiten in der Datenbank berücksichtigt werden.

Unabhängig von dieser Entscheidung ist jedoch folgendes zu beachten:

Datawarehouse oder Datamart?

Meistens wird unter einem Datawarehouse eine unternehmensweite Lösung verstanden, während Data Marts im Gegensatz dazu nur Teile der Aufgaben eines Datawarehouse erledigen können.

Unter einem Data Mart versteht man eine spezifische Datensammlung, in der nur die Bedürfnisse der jeweiligen Datensicht, z. B. des Marketing, abgebildet werden. Die Data Marts erfüllen zwar einen Teil ihrer jeweiligen Aufgaben im jeweiligen Bereich, aber eine unternehmensweite Strategie lässt sich daraus nur schwer ableiten, da die Ergebnisse nur schwer vergleichbar sind.

Ein Beispiel: Wie wollen Sie den Verkauf von bestimmten Artikeln innerhalb eines Vertriebssystems beurteilen, wenn die Artikelstatistiken im Einkauf nicht nach Gebieten und die im Vertrieb nur nach Warengruppen unterteilt werden? Dass dabei innerhalb der verschiedenen Bereiche auch noch unterschiedliche Definitionen vorliegen, was zum Beispiel unter Umsatz zu verstehen ist, merken viele Unternehmen häufig erst im Verlauf eines Datawarehouse-Projektes.

Auf der anderen Seite sind viele Datawarehouse-Projekte auch daran gescheitert, dass man versucht hat, gleich ein unternehmensweites Datawarehouse zu bauen. Diese Projekte brauchen häufig einen sehr langen Zeitraum (2-3 Jahre) bis sie abgeschlossen sind. Damit man diese Projekte überhaupt zu Ende führen kann, definiert man während des Projektes einen Zeitpunkt, ab dem keine der Veränderungen in den Geschäftsprozessen oder Daten bis zum Abschluss des Projektes mehr berücksichtigt werden können. Das heißt: Im ungünstigen Fall ist ein Teil der im Modell abgebildeten Geschäftsprozesse nicht mehr vorhanden und die neu hinzugekommen sind im Datawarehouse nicht abgebildet.

Aufbau einer Marketing Datenbank - Projektphasen

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.

Deshalb bietet sich als Lösung die Gliederung des Gesamtprojektes in einzelne Phasen, die sich an der Bereichs- bzw. Informationsstruktur eines Unternehmens orientieren. Von Anfang an wird dabei jedoch der Blick auf das Gesamte nicht vergessen. Das bedeutet, die Informationen, die alle Bereiche eines Unternehmens gleichzeitig benötigen, werden schon in Phase 1 realisiert. In späteren Phasen müssen so nur noch die bereichstypischen Informationen ergänzt werden.

Zum Beispiel startet ein Unternehmen mit dem Bereich Marketing und entwickelt eine

kleine Datenbank-Lösung, die sich vor allem am Informationsbedürfnis dieses Bereiches orientiert. So werden die Informationen und Prozesse rund um den Kunden und seine Kauftransaktionen sehr ausführlich abgebildet (z. B. Alter, Geschlecht, erhaltene Werbemittel/ Besuche, aktuelle Kaufdaten, historische Kaufdaten, usw.), Detailinformationen zum Produkt, soweit vom Marketing/Vertrieb nicht benötigt (besondere Herstellungs- oder Lieferdaten, Lieferantenartikelnummer, Lagerplatz, Gewicht, Volumen, u.ä.), jedoch vorerst einmal nicht.

Gleichzeitig werden im Datawarehouse Schnittstellen eingebaut - an Stellen, die eine spätere Erweiterung ermöglichen. Zum Beispiel in Richtung Einkauf, Produktion oder Logistik, um den speziellen Informationsbedarf dieser Bereiche zu einem späteren Zeitpunkt ins Datawarehouse einbinden zu können. Das Gesamtunternehmen kann also von Anfang an über einen „allgemeinen“ Teil der Informationen verfügen (z.B. Tagesvergleiche mit dem Vorjahr), und Marketing und Vertrieb erhalten schon die bereichstypischen Informationen.

Durch diese Vorgehensweise wird sicher gestellt, daß das Projekt in jeder Phase überschaubar und bezahlbar bleibt. Zudem können schon mit Abschluss der ersten Phase Informationen und Gewinn aus dem Datawarehouse gewonnen werden.

Damit ein Datawarehouse seine Aufgaben erfüllen kann, ist es besonders wichtig, dass man sich schon im Vorfeld des Datawarehouse-Projektes sehr genaue Gedanken über die Nutzung macht. Noch besser ist es, wenn sich der Unternehmensvorstand mit der strategischen Ausrichtung und den strategischen Erwartungen an das Datawarehouse selber auseinandersetzt. Wenn ein Mitglied der Geschäftsleitung die Patenschaft für das Projekt übernimmt, ist sichergestellt, das der Kontakt zwischen Projektteam und Geschäftsführung nicht verloren geht. Auch die Fachabteilungen müssen von Anfang an in das gesamte Projekt mit einbezogen werden.

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

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

Es empfiehlt sich deshalb neben der Hard- und Softwareauswahl folgendes schrittweises Vorgehen zusammen mit den Projektmitgliedern der Fachabteilungen, um ein Datawarehouse zu erstellen.

1. Erarbeiten der strategischen Zielen des Datawarehouse

2. Klärung des Informationsbedarfs aller späteren Anwender zur Bewältigung ihres Tagesgeschäftes und zur Umsetzung der strategischen Ziele
3. Analyse der bestehenden und zukünftigen Geschäftsprozesse des Unternehmens und der Datenbestände des operativen EDV-Systems sowie Analyse möglicher externer Daten
4. Definition der Datenqualität für das Datawarehouse
5. Erarbeiten und Implementierung des Datenmodelles und Definition der Metadaten (Metadaten sind vergleichbar mit einem sehr ausführlichen Inhalts-/Quellenverzeichnis des Datawarehouse). Dieses Datenmodell muss ausbaubar, flexibel und an die Geschäftsprozesse angepasst sein
6. Aufbereitung und Übernahme der Daten aus den operativen Systemen ins Datawarehouse für das Prototyp-Datawarehouse
7. Testen des Prototyp-Datawarehouse durch die Fachabteilungen, daraus resultierend gegebenenfalls Überarbeitung und Weiterentwicklung
8. Aufbereitung und Übernahme der Daten aus den operativen Systemen ins Datawarehouse für das Echt-System
9. Tuning des Datawarehouse für die zu diesem Zeitpunkt bekannten hauptsächlichen Abfragen/Analysen
10. Nach Inbetriebnahme des Datawarehouse muss eine kontinuierliche und konsequente Anpassung des Datawarehouse an die sich entwickelnden Geschäftsprozesse und Aufgaben/Analysen erfolgen.

Informationsbedarf der Unternehmens-Bereiche

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

Zwar variiert der Informationsbedarf von Branche zu Branche, die einzelnen Unternehmensbereiche benötigen branchenübergreifend jedoch typische Informationen, die in jedem Unternehmen vorhanden sein sollten, da sie die Datenbasis für Database Marketing und CRM bilden.

Die folgende, beispielhafte Zusammenstellung verzichtet zu Gunsten der Übersicht auf die Darstellung der Anforderungen einer konzernübergreifenden Informationssammlung sowie der besonderen Informationstiefe im BtoB-Bereich. Denn gerade im BtoB-Bereich werden neben den Informationen zum Kunden (z.B. Firma A) auch noch Informationen zur Person des Bestellers (Angestellter/Chef der Firma A) gesammelt – bekanntermaßen ein recht heikler und schwieriger Punkt, da die Ansprechpartner häufig wechseln.

Bereich Auftragsannahme/Kundenservice:

- 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.

- Fragen zum Angebot und Produktinformationen
- Beschwerden und Reklamationen

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.

- Anfragen zu Rechnungen, Gutschriften und Kontoauszügen

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

Herstellungsdaten, Lieferdaten, Lieferantenartikelnummer, Lagerplatz, Gewicht, Volumen, Preis usw.

Bereich Werbung/Vertrieb/Marketing

- alle Einzelinformationen zur Bewerbung des Kunden wie:
Datum, Art und Anzahl des überlassenen Werbematerials; Anzahl und Dauer der Außendienstbesuche; Informationen aus dem persönlichen Kontakt zwischen dem Außendienstler und dem Kunden (anstehende Jubiläen, runder Geburtstag, eventuell geäußerte Wünsche zur Produktveränderung, anstehende Pläne (Neubau nächstes Jahr, ...) etc.); Zielgruppe; Kundennummer; usw.

Datenqualität

Aber das Sammeln alleine hilft nur bedingt weiter, denn erst 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 z. B. Rechnungen ohne Kundenstammdaten (der Kundenstammdatensatz wurde gelöscht, da z.B. der Kunde seit x Jahren inaktiv ist, die Rechnung muß aber noch auf Grund der gesetzlichen Bestimmungen archiviert werden), Intradubletten, unvollständige oder inkonsistente Daten können die Aussagekraft der späteren Analysen in Frage stellen.

Deswegen ist es notwendig in diesen Phasen besonders 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 Projektphase ergibt sich deshalb häufig eine bereinigende Rückkoppelung aus dem Projekt in Richtung operative Systeme. Das kann z. B. so aussehen, dass für die weitestgehend dublettenfreie Kundenneuanlage eine Verbesserung des Matchcodesystem durchgeführt wird, endlich die Daten der Interessenten oder der Katalogretouren auch permanent gespeichert werden oder die Erfassungsmasken so gestaltet werden, dass Eingabefehler reduziert werden.

Besonders das Thema Intradubletten, der Umgang mit inaktiven Kunden und Interessenten, birgt hier Schwierigkeiten.

Im operativen Geschäft sind z. B. Dubletten oder inaktive Kunden häufig kein Problem, denn sie haben keinen oder nur wenig Einfluss auf den laufenden Betrieb. Somit fanden diese Themen bei vielen Firmen in der Vergangenheit wenig Beachtung. Für die Bewertung eines Kunden (z. B. in Bezug auf die Selektion eines Mailings) ist es aber sehr wohl wichtig, möglichst viele Informationen zu einem inaktiven Kunden zu haben, oder Detailinformationen zum Interessenten in die Bewertung mit einbeziehen zu können. Der

mehrfache Versand von ein und demselben Mailing an den gleichen Kunden kostet zum einen nur unnötiges Geld und kann zum anderen den Kunden verärgern.

Man sollte also ein solches Projekt auch dazu nutzen, die bestehenden Prozesse des operativen Geschäftes zu untersuchen und gegebenenfalls zu verbessern.

Befüllung und tägliche Nutzung des Datawarehouse

Die Befüllung der Datenbank erfolgt zum ersten Mal während des Projekts als Initialbefüllung. Alle für die Datenbank als relevant definierten Daten werden mit dem Stand XY in die Datenbank geladen. Während des Ladens wird ein Teil der Daten noch gemäß der Definitionen transformiert bzw. aggregiert (z. B. wird aus der Anrede eine Variable „Geschlecht“ erstellt).

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

Damit aus Daten relevante Informationen werden: Analysen

Neben den Selektionen sind die verschiedensten Analysen die Haupteinsatzgebiete der Datenbank. Bei den Analysen sollte man grundsätzlich unterscheiden zwischen den mehr kreativen, forschenden Analysen des Data Mining und den standardisierbaren regelmäßig wiederkehrenden Analysen wie z.B. Werbeerfolgskontrollen, Tagesstatistiken oder typische Fragestellungen des Berichtswesens, die häufig mit sogenannten OLAP-Tools durchgeführt werden.

Data Mining

Die Anzahl oder das Durchschnittsalter neuer Kunden oder Interessenten auszugeben ist kein Problem. Die Nutzung der Daten zur Beantwortung entscheidungsrelevanter Fragen gestaltet sich schon aufwändiger – stellt aber einen großen strategischen Vorteil dar. Beispiele für solche Fragen 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-Potential 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 bspw. 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.

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 ja Wettbewerbsvorteile erreichen und die Mitbewerber sollen nicht unbedingt wissen, dass man schon Data Mining betreibt.

Data Mining ist heute die Geheimwaffe im Wettlauf um Marktanteile. Über 80% der heute 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/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/Regeln in den Daten mit Hilfe der geeigneten Verfahren zu finden. Wobei natürlich grundsätzlich gilt: Nur Informationen, die den Analyseverfahren zur Verfügung stehen, können auch bei der Entwicklung der neuen Regeln berücksichtigt werden.

Als Beispiel: Kein Verfahren kann einen irgendwie gearteten 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 Personen leichter, die häufig sehr komplexen Zusammenhänge zu verstehen und daraus die richtigen Database Marketing Strategien und Umsetzungen zu entwickeln. Technisch ist man heute soweit, daß man Kundenverhalten genau vorhersagen kann und somit die Grundlage für die Aktionen schafft, um bereits bestehende Kunden zu halten und neue zu gewinnen.

Database Marketing und Data Mining

Database-Marketing und Data Mining gehören unverzichtbar zusammen. Häufig werden die Begriffe auch so verwendet, dass sie die Bedeutung des jeweils anderen implizieren. So war es besonders in der Anfangszeit des Database Marketing (80er und frühe 90er Jahre) 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 bzw. 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. Diese Entwicklung wurde nicht zuletzt dadurch erleichtert, dass die heute in großer Zahl vorhandenen Softwarepakete zu OLAP und Data Mining die früher in den Firmen mit Database Marketing üblichen, selbst geschriebenen Anwendungen/Analyseprogramme nach und nach ersetzen.

Ein zweiter positiver Aspekt der entwickelten Softwarepakete ist, dass die Analytiker und Anwender heute mehr Zeit haben, sich mit den Analysen und deren Ergebnissen auseinanderzusetzen und so dieses Wissen zur Optimierung der Geschäftsprozesse genutzt werden kann.

Fazit:

Unternehmen, die Informationen über ihre Kunden sammeln, analysieren und die Ergebnisse strategisch nutzen, schaffen sich eine solide Basis für profitable Kundenbeziehungen.