

Data Warehouse: Unternehmensdaten richtig lagern

Erschienen in: acquisa, Ausgabe 04/2000

Unternehmensdaten richtig lagern

Egal, wer im Unternehmen gerade eine Auswertung benötigt, ob der Verkaufsleiter, der Marketingverantwortliche oder die Geschäftsführung, alle wollen gleich wissen, warum die Zahlen so sind wie sie sind. Doch schnelle Antworten kann nur derjenige erhalten, der auf eine strukturierte Datensammlung zurückgreifen kann.

Wer sich seine Informationen aus unterschiedlichen Quellen zusammensuchen muss, steht von vorneherein auf verlorenem Posten. Ganz schlimm, wenn er auf unterschiedlichste Datensammlung etwa in Access- oder Excel-Dateien angewiesen ist. Mehr Arbeit, aber dafür schnelle und präzise Antworten erwarten denjenigen, der seine Daten ordentlich in einem Datawarehouse sammelt. Dabei wird unter einem Datawarehouse eine unternehmensweite Lösung verstanden. Anders bei den Datamarts, die nur Teile der Aufgaben eines Datawarehouses erledigen können. Diese bilden eine spezifische Datensammlung, in der nur die Bedürfnisse der jeweiligen Datensicht, zum Beispiel einer Abteilung, abgebildet werden.

Insellösungen beseitigen

Über die Jahre haben sich in vielen Unternehmen eine Reihe von Datamarts, als Insellösungen der jeweiligen Bereiche, gebildet. Da diese Datamarts meistens ohne ein Gesamtkonzept entstanden sind, sind der Umfang der Datenpflege, die Definition der einzelnen Kennzahlen (zum Beispiel Umsatz), die Aktualisierungsintervalle (täglich oder wöchentlich oder in Echtzeit) häufig nicht vergleichbar. Die Datamarts erfüllen zwar einen Teil ihrer Aufgaben im jeweiligen Bereich, aber eine unternehmensweite Strategie lässt sich daraus nur schwer ableiten, da die Ergebnisse kaum vergleichbar sind. Ein Beispiel: Wie lässt sich der Verkauf eines Artikels innerhalb des Vertriebssystems beurteilen, wenn der Einkauf seine Artikelstatistiken nicht nach Gebieten unterteilt und die Zahlen nur wöchentlich aktualisiert, während der Vertrieb seine Gebietsstatistiken täglich aktualisiert, aber nur nach Warengruppen unterteilt? Noch schwieriger wird es, wenn unterschiedliche Definitionen verwendet werden, zum Beispiel was unter Umsatz zu verstehen ist. Viele Unternehmen merken erst innerhalb eines Datawarehouse-Projektes, dass die einzelnen Abteilungen mit ganz unterschiedlichen Definitionen arbeiten.

Aufbau in Teilprojekte zerlegen

Angesichts des großen Aufwands, der mit dem Aufbau eines Datawarehouses verbunden ist, ist es meist sinnvoll, das Projekt in einzelne Phasen zu unterteilen. Diese orientieren sich an der Bereichs- oder Informationsstruktur eines Unternehmens. Die Umsetzung beginnt also mit dem Aufbau einer einzelnen Datamart für eine ausgewählte Abteilung. Allerdings wird schon bei der Planung berücksichtigt, dass diese Datamart später in ein Gesamtkonzept integriert wird. Konkret heißt das, man startet zum Beispiel mit dem Bereich Marketing oder Vertrieb und entwickelt eine kleine Lösung, die sich vor allem am Informationsbedürfnis dieses Bereiches orientiert. Es werden also die Informationen und Prozesse rund um den Kunden und seine Kauftransaktionen sehr ausführlich abgebildet (Alter, Geschlecht, erhaltene Werbemittel/Besuche, aktuelle Kaufdaten, historische Kaufdaten und so weiter). Detailinformationen zum Produkt wie besondere Herstellungs- oder Lieferdaten (z.B. Lieferantenartikelnummer, Lagerplatz, Gewicht, Volumen), werden dagegen erst einmal nicht abgebildet. Gleichzeitig werden Schnittstellen eingebaut, die eine spätere Erweiterung zum Beispiel in Richtung Einkauf, Produktion oder Logistik ermöglichen, um den speziellen Informationsbedarf dieser Bereiche zu einem späteren Zeitpunkt ins Datawarehouse einbinden zu können. Durch diese Vorgehensweise wird sichergestellt, 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. Im Gegensatz zum früher häufig unkoordinierten Aufbau von bereichs- oder aufgabenbezogenen Datamarts werden durch das vorhandene Gesamtkonzept Inkonsistenzen im Datenbestand der einzelnen Datamarts vermieden. Damit ein Datawarehouse seine Aufgaben erfüllen kann, ist es besonders wichtig, dass man sich schon im Vorfeld des Projektes sehr genaue Gedanken über die Nutzung macht. Der Erfolg eines Datawarehouses 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 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 Ziele des Datawarehouses.
2. Klärung des Informationsbedarfs aller späteren Anwender zur Bewältigung ihres Tagesgeschäftes und zur Umsetzung der strategischen Ziele.
3. Analyse der bestehenden und zukünftigen Geschäftsprozesse und der Datenbestände des operativen EDV-Systems und 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-/Quellenverzeichnis des Datawarehouses, oft als »Daten über Daten« bezeichnet). Dieses Datenmodell muss ausbaubar, flexibel und den Geschäftsprozessen angepasst sein.
6. Aufbereitung und Übernahme der Daten aus den operativen Systemen für das Prototyp-Datawarehouse.
7. Testen des Prototyp-Datawarehouses durch die Fachabteilungen und gegebenenfalls Überarbeitung und Weiterentwicklung.
8. Aufbereitung und Übernahme der Daten aus den operativen Systemen für das Echt-System.
9. Tuning des Datawarehouses für die jetzt bekannten hauptsächlichen Abfragen/Analysen.
10. Nach Inbetriebnahme des Datawarehouses muss eine konstante und konsequente Anpassung des Datawarehouses an die sich entwickelnden Geschäftsprozesse und Aufgaben/Analysen erfolgen. Dabei sollten sich alle Verantwortlichen bewusst sein, dass der Aufbau eines Datawarehouses eine komplexe Aufgabe ist, die entsprechend Zeit und Ressourcen verlangt. Andererseits sollten die Daten, mit deren Hilfe die Verantwortlichen in Zukunft die richtigen Entscheidungen treffen wollen, den Aufwand auch Wert sein. Denn eine vernünftige Auswertung kann nur erhalten, wer vorher in die Qualität der Datenaufbereitung investiert.