

Data Warehousing



DATA WAREHOUSING - Building the Corporate Knowledgebase;
Tom Hammergren; Sybase Press / International Thomson Press
ISBN 1-85032-856-0, 470 Seiten, 1996

Helmut Schluderbacher

S

Zeit geistert ein neuer Begriff durch die EDV: Data Warehouse - Das Datenwarenhäus. Doch was ist ein Datenwarenhäus?



Eine gute transaktionsorientierte Datenbank ist schlank und hat so wenig wie möglich Redundanzen. Anders ein Datenwarenhäus, das die Informationen in verschiedenen Verdichtungen verspeichert hält.

In früheren Jahren war es für die Firmen notwendig, alle Daten so einfach und komprimiert wie möglich zu speichern. Platzsparen war angesagt. Daher wurden alle Daten in den Basisinformationen verspeichert. Wenn nun eine Auswertung über die nicht unbe-

trächtlichen Daten gefahren wurde, hat dies immer einen Menge Rechenzeit beansprucht. Solche Informationen dienen dann aber der weiteren strategischen Planung!

Zudem sind die herkömmlichen Datenbanken prozeßorientiert aufgebaut. Das rasche Verspeichern und die rasche Abfrage von Detailinformation steht im Vordergrund.

Im Gegensatz dazu stehen die Auswertungen. Bei ihnen muß die Information nach bestimmten Kriterien zu meist verdichtet, sprich, komprimiert sein. Und dies auch noch in verschiedener Hinsicht. Zum Beispiel ein Artikel steckt auch in einer Artikelgruppe und in einer Warengruppe. Diese werden verdichtet auf das einzelne Geschäft (Filiale), das Bundesland und dann auf den einzelnen Staat. Dies alles wird dann noch verdichtet auf den Tag, die Woche, das Monat und dann auf das Jahr. Damit erhält man bei diesem einfachen Beispiel einen dreidimensionalen Würfel, mit den Dimensionen Artikel, Zeit und Ort.

Es sollte also möglich sein, die Informationen so aufzubereiten, daß eine Aus-

wertung nicht die Hälfte des Tages den Rechner blockiert und ein einzelner Fehlerfall die mittlere Katastrophe darstellt. Bei uns gibt es viele Firmen, deren Verarbeitungsprozesse an diese Zeiten herankommen. Die Abfragen sollten auch flexibel sein, damit das Hinzufügen von einfachen Zusatzinformationen nicht einen großen Programmieraufwand bedeutet. Dies ist jedoch nur eine Motivation für ein Data Warehouse.

Nach einer Einführung wird in dem Buch auf die Analyse, das Design und die Erstellung beschrieben. Natürlich werden die Begriffe Star-Schema, Snowflake-Schema, und viele andere sehr anschaulich erklärt. Das Buch von Tom Hammergren geht aber über die einfache Beschreibung des Data Warehouse weit hinaus. Als zusätzliches Zucker wird sehr ausführlich auf Data Warehouse-Projekte als solche eingegangen. Die Bildung eines Projektteams, und das Projektmanagement werden beschrieben. Die beiliegende Diskette enthält einige Tips und Vorlagen, die zumindest als Motivation sehr zu empfehlen sind. Das Buch ist in englischer Sprache.

➤Datenbanken, d.s. Sammlungen von systematisch oder methodisch angeordneten, einzeln vor allem elektronisch zugänglichen Elementen, sind

- a) in den vollen urheberrechtlichen Schutz - zeitlich beschränkte ausschließliche Verwertungsrechte - einbezogen, wenn sie „infolge der Auswahl oder Anordnung des Stoffes eine eigene geistige Schöpfung ihres Urhebers“ sind .Sie unterliegen nicht den Bestimmungen über Computerprogramme;
- b) auch ohne die Voraussetzung „eigene geistige Schöpfung eines Urhebers“ allen wichtigen Schutzbestimmungen des Urheberrechtsgesetzes unterstellt, wenn für die Beschaffung, Überprüfung oder Darstellung ihres Inhaltes eine wesentliche Investition erforder-

lich war. Entsprechendes gilt auch für Bearbeitungen einer geschützten Datenbank (wesentliche Investition für die Bearbeitung). *Selbst die Verwertung unwesentlicher Teile einer Datenbank ist geschützt, wenn dies wiederholt und systematisch geschieht und der normalen Verwendung der Datenbank widerspricht oder berechnigte Interessen ihres Herstellers verletzt.*

Keine Verletzung der Schutzrechte bewirkt die private Kopie einer nichtelektronischen Datenbank - also auch keine private Kopie einer elektronischen Datenbank erlaubt ! - oder die Verwendung jeder Art von Datenbank für Zwecke des Unterrichtes oder der wissenschaftlichen Forschung. Das Schutzrecht erlischt nach 15 Jahren. Unbefugte Benutzung macht entgelt-

pflichtig und ist überdies - ausgenommen bei unbefugter Privatkopie für eigene Zwecke oder auf Bestellung für private Zwecke eines anderen, so wie bei Computerprogrammen - auch strafbar. Geschützt sind jedenfalls alle Datenbanken im Bereich der Europäischen Union, außerhalb dieser nach Maßgabe von Staatsverträgen. Dies gilt auch für Datenbanken, die vor dem 1.1.1998 hergestellt worden sind. Frühestes Ende der Schutzrechte ist der 31.12. 2012. Verwertungshandlungen vor Inkrafttreten des Gesetzes werden nicht erfaßt.

Wenn dieser Entwurf vom Parlament im Herbst als Gesetz beschlossen wird, sollte damit ein weiteres Schlupfloch der EDV-Piraterie geschlossen sein. Über die Gesetzwerdung werde ich zu gegebener Zeit berichten.