

EMISA–Kolumne *Data Warehousing*

Nr. 1

Ziel

Diese Kolumne erscheint erstmals im Heft 1/2000 des EMISA–Forums. Mit dieser neuen Einrichtung soll den Lesern des EMISA–Forums ein Einblick in neue Entwicklungen, Veranstaltungen, Produkte und interessante Forschungsprojekte auf dem Gebiet des Data Warehousing eröffnet werden. Ein Anspruch auf Vollständigkeit wird nicht erhoben. Auch soll die Kolumne sich schwerpunktmäßig mit den "EMISA"–Fragestellungen beschäftigen:

- Wie entwickelt man ein Data Warehouse?
- Welche Modellierungsansätze sind erfolgreich?
- Wie wird ein Data Warehouse in der Praxis eingesetzt?
- Inwieweit befördert ein Data Warehouse die Integration von Informationsquellen in einer Unternehmung?

Falls Sie Neuigkeiten auf dem Gebiet des Data Warehousing haben, so senden Sie diese bitte an den Moderator dieser Kolumne. Beiträge sollten auf ca. 300 Worte begrenzt sein und als Einstieg dienen.

Bericht vom Workshop Design and Management of Data Warehouses

Im Juni fand in Heidelberg der erste internationale Workshop über den Entwurf und das Management von Data Warehouses statt. Gut dreißig Teilnehmer aus Forschung und Praxis beleuchteten verschiedene Probleme, mit denen eine Unternehmung konfrontiert ist, wenn es ein Data Warehouse installieren will. Dabei kristallisierten sich folgende Aspekte heraus:

Es sind Methoden zu entwickeln, die den Data Warehouse–Administrator bei der Auswahl der zu speichernden Datenwürfel (bzw. materialisierten Sichten) helfen. Schon bei einer relativ kleinen Anzahl von Datenquellen ist der Suchraum sehr groß. Zur Zeit werden Verfahren aus der Künstlichen Intelligenz untersucht.

Ein weiterer Schwerpunkt ist die konzeptuelle Modellierung eines Data Warehouse. Während sich bei Datenbanken das Entity–Relationship–Modell durchgesetzt hat, herrscht bei Data Warehouses noch eine große Unsicherheit über geeignete Modellierungssprachen. Unter anderem werden Konzeptlogiken untersucht, die sowohl die Beziehung zu Datenquellen als auch die Multidimensionalität berücksichtigen.

Schließlich sei das Metadaten–Management genannt. Die Metadatenbank eines Data Warehouse ist der zentrale Zugang sowohl für Administratoren als auch für Nutzer. In der Praxis sind Metadaten häufig in heterogenen Systemen untergebracht. Eine Integration all dieser Metadaten erscheint dringend erforderlich.

Weitere Informationen zu diesem Workshop inklusive der elektronischen Workshopbeiträge finden Sie unter <http://SunSITE.Informatik.RWTH–Aachen.DE/DMDW99/>.

Eine Workshop–Serie mit dem gleichen Akronym zum Thema "Data Mining und Data Warehousing" findet seit der Jahrestagung der GI im Jahre 1998 statt. Das Programm von

1999 ist zu finden unter <http://www-ai.cs.uni-magdeburg.de/~scheffer/dmdw99-programm.html>.

Interessante Adressen

Die Forschergruppe DWQ (Foundations of Data Warehouse Quality) hat Verfahren entwickelt, wie man Qualitätsanforderungen an Data Warehouses formuliert, ihren Erfüllungsgrad bestimmt, und ein Data Warehouse so entwirft, daß die Qualitätsanforderungen berücksichtigt werden. Weitere Informationen sind unter <http://www.dbnet.ece.ntua.gr/~dwq/> zu finden.

Ein guter Einstieg in Data Warehousing ist in <http://pwp.starnetinc.com/larryg/index.html> zu finden. Interessant sind besonders die Essays zu Themen wie der Nutzung von Data Warehouses zur strategischen Entscheidungsunterstützung. Es finden sich auch Verweise zu Produkten und Herstellern.

Eines der ersten akademischen Data Warehouse-Projekte ist WHIPS (<http://WWW-DB.Stanford.EDU/warehousing/>) . WHIPS hat einen Schwerpunkt auf der algorithmischen Seite, d.h. wie man ein Data Warehouse-System effizient implementiert und betreibt.

In Bochum wird unter <http://olap.winf.ruhr-uni-bochum.de/research.phtml> ein Überblick über Data Warehouse-Aktivitäten im deutschsprachigen Raum geboten.

Die Fachgruppe 2.5.1 hat einen Arbeitskreis "Konzepte des Data Warehousing" gegründet. Informationen unter <http://www6.informatik.uni-erlangen.de:8080/>.

Der wirtschafts-informatische Aspekt des Data Warehousing wird in dem Arbeitskreis "Modellierung und Nutzung von Data Warehouse-Systemen" (<http://ceus.uni-bamberg.de/ak5104/>) beleuchtet.

An der Universität Erlangen-Nürnberg wird ein Konsistenzbegriff für Data Warehouses untersucht. Die Frage ist, ob es ähnlich wie in der relationalen Datenbanktheorie Normalformen gibt, die "gute" Data Warehouse-Schemata beschreiben. Nähere Informationen unter <http://www6.informatik.uni-erlangen.de/research/konsistenz.html>.

An der Universität Oldenburg entstand im Rahmen einer Diplomarbeit eine konzeptuelle Modellierungsumgebung für Data Warehouses auf Basis von UML. Mehr Informationen dazu auf der Home Page von Arne Harren unter <http://www-is.offis.uni-oldenburg.de/personal/AHa.html>

Diese Liste ist sicher unvollständig. Falls Sie weitere interessante Adressen kenne, so senden Sie sie bitte an den Moderator dieser Kolumne:

Manfred A. Jeusfeld
Universität Tilburg, Infolab, Postbus 90153, NL-5000 LE Tilburg
Tel: +31-13-466-3119, Email: jeusfeld@kub.nl