

Neuerscheinungen – Bücher, die mir auffielen

Folge 3

Gottfried Vossen, Universität Münster

In dieser Rubrik möchte ich Ihnen wieder Bücher vorstellen, die mir in letzter Zeit auf den Tisch gekommen sind, die mir lesenswert (oder zumindest erwähnenswert) erscheinen oder die mir dankenswerterweise von einem Verlag zur Rezension überlassen wurden. Die dabei jeweils präsentierte Auswahl erhebt weder Anspruch auf Vollständigkeit noch auf Einhaltung einer bestimmten Systematik; der Natur meiner persönlichen Interessen entsprechend entstammen allerdings die meisten der genannten Bücher den Bereichen Datenbanken und Informationssysteme.

Ich beginne heute mit einer nicht nur für die Informatiker (weiteren) interessanten Neuerscheinung zum Thema Data Mining. Das Buch von Jiawei Han und Micheline Kamber mit dem Titel *Data Mining – Concepts and Techniques* ist vor kurzem bei Morgan Kaufmann erschienen (ISBN 1-55860-489-8). Beim Data Mining geht es ja um das Erkennen relevanter Information oder die Herleitung von „Wissen“ aus in Anwendungen angesammelten Daten; solche Anwendungen sind z.B. wissenschaftlicher, medizinischer, demographischer oder finanzwirtschaftlicher Art. Das Gebiet selbst gewinnt erst seit etwa 1993 an Bedeutung, hat allerdings heute eine hohe Dynamik und stößt in immer neuen Kreisen auf Beachtung. Das Buch beginnt mit einer Motivation, warum Data Mining wichtig



ist und welche Funktionalität man hier erwartet. Als Vorbereitung werden Data Warehouse-Technologie und OLAP erläutert, sodann geht es um den wichtigen Aspekte des *Data Preprocessing*, also der Vorbereitung der Daten auf das eigentliche Mining. Interessanterweise wird bereits in Kapitel 4 eine Anfragesprache (DMQL) vorgestellt, in der sich Mining-Aufgaben formulieren lassen. Den zentralen Mining-Techniken sind dann jeweils eigene Kapitel gewidmet: Bei der Suche nach Assoziationsregeln (Kapitel 6)

geht um das Auffinden von Zusammenhängen z.B. im Käuferverhalten („wer Bier kauft, kauft auch Chips“); die zentrale algorithmische Technik ist hier das Apriori-Verfahren von Agrawal et al., welches ausführlich und in zahlreichen Varianten und Verbesserungen behandelt wird. Bei Klassifikation und Vorhersage (Kapitel 7) sowie bei der Cluster-Analyse (Kapitel 8) kommen neben rein algorithmischen auch statistische Techniken zur Anwendung, die ebenfalls ausführlich behandelt werden. Kapitel 9 beschäftigt sich mit neueren Mining-Anwendungen, in denen die Untersuchungsgegenstände komplexe Objekte (wie räumliche Daten, Multimedia-Datenbanken, Zeitreihen, Texte oder Weblogs) sind. Das Buch endet mit Anwendungen und Trends im Data Mining und geht, wenn auch kurz, auf die Frage der sozialen Implikationen einer derartigen Technologie ein. Mit gut 500 Seiten kann das Buch noch als „kurz gefasst“ gelten. Es gibt einen ausgezeichneten Überblick über den gegenwärtigen Stand der Kunst auf diesem Gebiet, und mit dem im Web (demnächst) verfügbaren Lehrmaterial kann (und sollte) man (als Hochschullehrer) dann eine Spezialvorlesung zu diesem Thema ankündigen ☺

Das nächste Buch dürfte zu den ersten deutschsprachigen Texten über SQL-99 gehören: Wolfgang Panny gibt zusammen mit Alfred Taudes eine *Einführung in den Sprachkern von*

SQL-99 (Springer 2000, ISBN 3-540-65547-6). SQL-99 wurde als Sprachstandard ja seit 1995 erwartet und dann doch erst im Sommer 1999 verabschiedet. Bei SQL-99 wurde gegenüber dem Vorgänger SQL-92 die Dreiteilung in Entry, Intermediate und Full SQL aufgegeben; Entry wurde quasi ersetzt durch einen *Core SQL* genannten Sprachkern, den ein



SQL-99-konformer Anbieter einzuhalten hat. Das Buch besteht aus drei größeren Teilen: Einführung und Grundlagen, SQL für Endbenutzer und SQL und Anwendungsentwicklung. Dabei wird die jeweilige Syntax ausführlich erläutert, jedoch nimmt die Intensität der Beispiele nach hinten ab. Eine wesentliche Neuerung von SQL-99, die objekt-relationale Erweiterung von SQL und damit die benutzerdefinierten Typen und Funktionen, werden nur sehr knapp abgehandelt; gerade hier wäre ein umfangreicheres Beispiel angemessen gewesen. Ansonsten erspart einem das Buch natürlich ein umständliches und zeitraubendes Durcharbeiten des Standards, zumal jeder Datenbank-Hersteller ohnehin auf seine spezifische Weise davon abweicht. Hilfreich wäre – das habe ich an Date und Darwen immer geschätzt – eine SQL-Grammatik gewesen, an

der sich die wesentlichen syntaktischen Konstrukte, die ja durchaus recht komplexe Ausdrücke zu formulieren gestatten, nachvollziehen lassen.

Beim nächsten Titel geht es ebenfalls um SQL, diesmal allerdings dessen Anbindung an die Java-Welt. Das Buch von Gunter Saake und Kai-Uwe Sattler über **Datenbanken & Java - JDBC, SQLJ und ODMG** (dpunkt.verlag 2000, ISBN 3-932588-54-1) ist in gewissem Sinne ein deutsches Pendant zu dem in der letzten Ausgabe dieser Rubrik beschriebenen Buch von Melton und Eisenberg. Es behandelt die Kopplung von Java und Datenbanken, genauer Möglichkeiten des Zusammenwirkens von objektorientierter Programmierung mit Java mit relationalen, auf SQL basierenden Datenbanken. Es beginnt mit kurzen Einführungen in die jeweiligen Gebiete (Java, relationale Datenbanken); allerdings wendet sich das Buch erklärtermaßen an Leser, denen beide Bereiche nicht gänzlich neu sind. Danach erreicht das



Buch eine erstaunliche Breite: Es behandelt ausführlich die etablierten Kopplungstechniken JDBC und SQLJ (statisches SQL), wendet sich dann aber auch Ansätzen zu, die derzeit noch nicht als stabil gelten können, etwa dem Java Binding der ODMG. Das Kapitel 7 *Von Objekten zu Relationen* ist für den Praktiker sicher von besonderem Interesse, denn in Anwendungen wird ja zunehmend objektorientiert programmiert, systemseitig werden die zu bearbeitenden Daten jedoch relational gespeichert; hier werden objektrelationale Abbildungen sowie das Sun Werkzeug *Java Blend* vorgeführt. Das Buch verwendet durchgehend eine im Anhang zusammenfassend dargestellte Beispielanwendung und enthält ferner zahlreiche Programmbeispiele. Es hat ferner mit gerade mal 300 Seiten einen angemessenen Umfang. Da es sich auf

dem aktuellen Stand der Technik befindet und auch einen Ausblick auf noch zu Erwartendes wagt, ist das Buch ausgesprochen empfehlenswert für eine Leserschaft, die in ihrer täglichen Arbeit Relationen- und Objektwelten miteinander zusammenbringen muss oder soll; dies kann auch für Studierende in einschlägigen Praktika gelten.

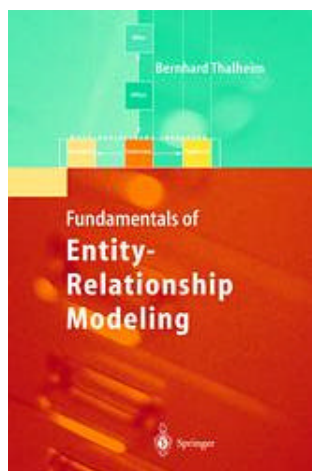
Der nächste Titel ist aus einer Initiative der Fachgruppe 5.12 Projektmanagement des Fachbereichs 5 (Wirtschaftsinformatik) der GI entstanden. Hans-Joachim Etzel, Heidi Heilmann und Reinhard Richter haben als Herausgeber Arbeiten zum Thema **IT-Projektmanagement – Fallstricke und Erfolgsfaktoren, Erfahrungsberichte aus der Praxis** (dpunkt.verlag 2000, ISBN 3-932588-83-5) zusammen getragen. Es geht um Projekte aus



dem IT-Bereich, die teilweise erfolgt waren und teilweise gescheitert sind. Die Autoren der insgesamt 6 Projektberichte geben ihre Erfahrungen in weitgehend anonymisierter Weise und zum Teil geraffter Form wider, und der Bogen ist weit gespannt: von der Softwareentwicklung und –einführung über Standardsoftware, Migration, Jahr-2000-Problematik bis hin zu Sanierung und Umsetzung einer (neuen) IT-Strategie. Was das Buch besonders lesenswert macht ist das einführende Kapitel von Heidi Heilmann über *Erfolgsfaktoren des IT-Projektmanagements*, in welchem das „magische Fünfeck“ als Bewertungsrahmen abgesteckt wird. Jedes der folgenden Kapitel endet mit einem Kommentar der Buchherausgeber, in dem das gerade behandelte Projekt kommentiert und seine Schwachstellen anhand dieses Bewertungsrahmens analysiert werden. Naturgemäß ist das Buch kein technisches, aber dennoch ist man ge-

neigt, es in einschlägigen Studiengängen, die sich in welcher Form auch immer mit IT befassen, als Standardlektüre zu empfehlen, denn man lernt hier viel Grundsätzliches über Projektmanagement und über Erfahrungen bei dessen Umsetzung.

Der als nächstes vorgestellte Titel ist noch einmal für den an (auch theoretischen) Grundlagen von Datenbanken und vor allem Datenbankentwurf Interessierten gedacht. Von Bernhard Thalheim stammt **Fundamentals of Entity-Relationship Modeling** (Springer 2000, ISBN 3-540-65470-4). Dieses Buch, das nach Aussage des Autors das Ergebnis von 10 Jahren Arbeit auf dem Gebiet der ER-Modellierung darstellt, geht aus vom klassischen Chenschen Modell und erweitert dieses schrittweise um eine Reihe von Aspekten: Modellierungskonstrukte wie



geschachtelte Attribute, disjunkte Vereinigungen von Typen („Cluster“), Beziehungen erster und höherer Ordnung, Integritätsbedingungen, Anfrage-, Update- sowie typspezifische Operationen und schließlich um Verhalten. Dabei wird durchgehend versucht, die Anschaulichkeit des ER-Modells mit seinen nur zwei grundlegenden Konstrukten zu erhalten, aber es liegt in der Natur der Sache, dass im Laufe der Entwicklung immer mehr formale Schreibweisen herangezogen werden müssen. Damit lässt sich natürlich die jeweils zugrunde liegende Theorie gut erfassen. Andererseits fragt sich der Praktiker stets, wie bzw. mit welchen Werkzeugen eine derart umfassende Theorie auch angewendet werden kann; zumindest sind hier Strategien sowie Vorgehensweisen gefragt. In dieser Hinsicht ist das Buch für den

Software-Entwickler, der im Rahmen einer größeren Aufgabe unter anderem ein ER-Verfahren braucht, vielleicht etwas enttäuschend, denn das in Kapitel 9 beschriebene RADD-System, welches ja zahlreiche der angesprochenen Modellierungsaspekte unterstützen soll, hat nach meinem Wissen das Entwurfsstadium nie verlassen. Man würde sich also noch so

etwas wie eine Marktübersicht verfügbarer Tools wünschen; andererseits ist die über 50 Seiten lange Bibliographie sicherlich das beste Kompendium zur ER-Modellierung, das derzeit vorliegt.

Der letzte Titel für heute widmet sich endlich dem Thema Unternehmensgründung und stammt vom wohl bekanntesten deutschen WI-Professor: August-Wilhelm Scheer schreibt ein Buch zum Thema *Unternehmen gründen ist nicht schwer ...* (Springer 2000, ISBN 3-540-41063-5). Ich mache es mir an dieser Stelle einfach und zitiere aus der Verlagsankündigung; ich selbst könnte den Inhalt des Buches nicht treffender wiedergeben:



„Unternehmensgründungen sind "in". Mit dem erfolgreichen Aufbau eines Unternehmens sind aber neben vielen Erfolgserlebnissen auch große Anstrengungen und Probleme verbunden. Die hohen Anforderungen an physische und psychische Stabilität der Unternehmensgründer in der New Economy werden erst in den kritischen Wachstumsphasen der Unternehmen deutlich. Der Autor, erfolgreicher Wissenschaftler, Unternehmensgründer und politischer Berater, fasst seine Erfahrungen zu spannenden Erzählungen und "Lebensregeln" zusammen. Die erfolgreiche Gründung der IDS Scheer AG, sowie die Entwicklung über 20 weiterer Spin-off-Unternehmen seines Instituts für Wirtschaftsinformatik an der Universität Saarbrücken bilden den Hintergrund seiner Ausführungen. Das Buch wendet sich an Unternehmensgründer, Studenten, Wissenschaftler und

Politiker. Die offene Schilderung der Gründungs- und Wachstumsphasen seiner Unternehmen lassen den Leser an dem Abenteuer des "Unternehmertums" teilhaben. Unternehmertum fordert er auch von den Professoren in der universitären Forschung und Lehre. Zur Verbesserung der Rahmenbedingungen für Unternehmensgründungen macht er kritische Anmerkungen und konstruktive Vorschläge zur deutschen Innovationspolitik. Er gibt Einblicke in das Lebensgefühl der Menschen in der New Economy, in der PC nicht nur für "Personal Computer", sondern auch für das Kultessen Pizza und Cola während der Nachtarbeit am Computer steht. Interessenten der High-Tech Welt lernen Interna kennen, die hinter den glitzernden Fassaden üblicherweise verborgen bleiben.“

Zum Schluss wieder einige Kurzhinweise auf weitere Neuerscheinungen, die ich mir noch nicht so intensiv angesehen:

- M. Goossens, S. Rahtz: *Mit LaTeX ins Web – Elektronisches Publizieren mit TeX, HTML und XML*; Addison-Wesley 2000, ISBN 3-8273-1629-4

Dieses Buch richtet sich an den Satztechniker, der nach Möglichkeiten sucht, seine LaTeX-Dokumente im Internet zu verwenden und zwischen HTML oder XML hin und her zu konvertieren. Es liegt an sich nahe, dass sich Markup-Sprachen vergleichsweise leicht ineinander konvertieren lassen; in diesem Buch ist praktisch alles zusammen getragen, was man dazu wissen muss.

- S. Voß, K. Gutenschwager: *Informationsmanagement*; Springer 2001, ISBN 3-540-67807-7

Dieses Buch behandelt Informationsmanagement aus einer wirtschaftswissenschaftlichen sowie aus einer wirtschaftsinformatischen Sicht. Im Einzelnen geht es um Modelle und Planung, Unternehmensmodellierung, Daten- und Wissensmanagement sowie um eine Reihe anderer relevanter Themen.

Die nächsten drei Titel drehen sich wieder um das Thema XML (vgl. auch die letzte Ausgabe dieser Rubrik); hier gibt es allmählich empfehlenswerte deutschsprachige Literatur, insbesondere seit Neuestem auch für XSL sowie für XSLT; außerdem ist der Wrox-Verlag dabei, sich als eine Art Spezialverlag für XML zu etablieren:

- H. Behme, S. Mintert: *XML in der Praxis*; Addison-Wesley, 2. erweiterte Auflage 2000, ISBN 3-8273-1636-7
- M. Bach: *XSL und XPath – verständlich und praxisnah*; Addison-Wesley 2000, ISBN 3-8273-1661-8
- M. Kay: *XSLT Programmer's Reference*; Wrox Press 2000, ISBN 1-861003-12-9