

EMISA Forum: Kolumnen

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 im Data Warehouse als der zentrale Zugang sowohl für die Forschung als auch für die Praxis hervorgehoben. In der Praxis sind Metadaten häufig in heterogenen Systemen vorhanden. In dieser Metadaten erscheint dringend erforderlich.

Weitere Informationen zu diesem Workshop im EMISA Forum finden Sie unter <http://SunSITE.Informatik.RWTHAACHEN.de>

Eine Workshop-Serie mit dem gleichen Akronym "Design and Management of Data Warehousing" findet seit der Jahrestagung der EMISA 2001 statt.



Data Warehousing Kolumne (3 Folgen, 2000 - 2001)



Business Process Management – Kolumne

Folge 5

Manfred Reichert (Universität Twente)

Stefanie Rinderle (Universität Ulm)

Barbara Weber (Universität Innsbruck)

Liebe Leser,

mit der aktuellen Ausgabe des EMISA Forums erhalten Sie Folge 5 unserer Kolumne zum Thema *Business Process Management (BPM)*. Diese berichtet über aktuelle Themen, Entwicklungen, wissenschaftliche Arbeiten und Veranstaltungen aus dem BPM-Umfeld.

Themen zur Modellierung und zum Management von Prozessen nehmen bei der diesjährigen EMISA-Fachtagung, die vom 8. – 9. Oktober 2007 in St. Goar stattfinden wird, wieder eine wichtige Rolle ein. Des weiteren findet vom 24. – 28. September 2007 in Brisbane zum fünften Mal und erstmals außerhalb Europas die BPM-Konferenz statt. Diese hat sich mittlerweile als die wichtigste wissenschaftliche Konferenz im BPM-Bereich etabliert. Das Programm der BPM'07 ist in diesem Heft abgedruckt.

Inhaltliche Schwerpunkte der heutigen Kolumne bilden Arbeiten zu Fehlern in Prozessmodellen und zur Definition von Prozessänderungen. Prozess-Management-Technologie wird in der Praxis nur dann sinnvoll einsetzbar sein, wenn Prozessmodelle fehlerfrei spezifiziert und später auf hoher Abstraktionsebene anpaßbar sind. Zu erstgenanntem Thema liefert die Dissertationsschrift von Jan Mendling mit dem Titel „*On the Detection and Prediction of Errors in EPC Business Process Models*“ einen wichtigen Beitrag. Eine Zusammenfassung ist in dieser Kolumne abgedruckt. Die Umsetzung von Prozessänderungen auf hoher Abstraktionsebene ist die Zielsetzung von *Process Change Patterns*. Sie bilden den Gegenstand unseres aktuellen Schlagworts.

Weitere Inhalte dieser Kolumne bilden ein Bericht zum IEEE ProGility'07-Workshop (*Agile Cooperative Process-Aware Information Systems*) sowie unser BPM-Veranstaltungskalender für die 2. Jahreshälfte 2007.

Liebe Leser, wir sind wie immer auch an Ihren Anregungen und Beiträgen zum Thema Business Process Management interessiert. Ihre Beiträge senden Sie bitte an Manfred



BPM Kolumne (11 Folgen, seit 2005)



Büchertipps (17 Folgen, seit 2000)

Neuerscheinungen – Bücher, die mir auffielen

Gottfried Vossen, Universität Münster

In dieser (wie die Rubrik mit Web-Tipps ebenfalls neuen) Rubrik möchte ich Ihnen Bücher vorstellen, die mir in letzter Zeit auf den Tisch gekommen sind und die mir lesenswert (oder zumindest erwähnenswert) erscheinen. Die dabei jeweils präsentierte Auswahl erhebt weder Anspruch auf Vollständigkeit noch auf Einhaltung einer bestimmten Systematik. Als Hochschullehrer ist man ja in der durchaus beneidenswerten Situation, dass man sich nur selten ein Fachbuch wirklich kaufen muss; die meisten erhält man als Dozentenfrei- oder Werbeexemplare von den Autoren bzw. Verlagen kostenlos. Daneben schicken einem auch viele Kollegen dankenswerter Weise ihre neuen Bücher. (Sollten Sie selbst gerade an einem Buch arbeiten, so verstehen Sie dies ruhig als den berühmten „Wink mit dem Scheunentor“, denn diese Rubrik soll ja nicht nur irgendwann fortgesetzt, sondern regelmäßig geschrieben werden.)

Die erste Neuerscheinung, ganz frisch, ist *Data on the Web – From Relations to Semistructured Data and XML* von Serge Abiteboul, Peter Buneman und Dan Suciu (Morgan Kaufmann Publishers 2000, ISBN 1-55860-622-X). Ausgehend von der Feststellung, dass heute immer mehr Daten in elektronischer Form vorliegen und im Web abrufbar sind und dass die Rolle von Datenbanken und Datenbanktechniken, aber auch die Austauschbarkeit von Daten zwischen unterschiedlichen Anwendungen in diesem Kontext immer wichtiger wird, wird der Stand der Forschung auf diesem Gebiet dargestellt.



Dabei geht es den Autoren um vier große Themenbereiche: Datenmodelle, Anfragen, Typen und Systeme. Bei Datenmodellen wird einerseits das Thema semistrukturierte Daten und andererseits die Sprache XML behandelt. Semistrukturierte Daten unterscheiden sich zwischen Inhalt und dessen Beschreibung und folgen insbesondere keinem einheitlichen, festen Daten-Schema. Um die Bestandteile einer Sammlung semistrukturierter Daten dennoch erfassen zu können oder sogar

bieten sich graphische Darstellungen
lichkeiten von XML an.

mit unterschiedlichen Erweiterungen von
(ausgesprochen wie das englische Wort
n Web-Daten und darüber hinaus) über
Einigungen hinsichtlich Anfragesprachen
Typisierung von semistrukturierten Daten
hier eine kompakte Übersicht, u.a. über
ng, Möglichkeiten der Extraktion von
bedingungen. Im Abschnitt über Systeme
verarbeitung behandelt (z.B. Mediatoren),
hentlich werden ein paar Bemerkungen zu
ht.

ie Web-Anwendungen entwickeln, wie an
an Studierende nach dem Vordiplom) mit
e aktuellen Techniken informiert werden
n noch nicht im Standardkanon einer
nahtlos darin integrieren bzw. daran

Für Sie gesurft – Neue (und alte) Tipps aus dem WWW

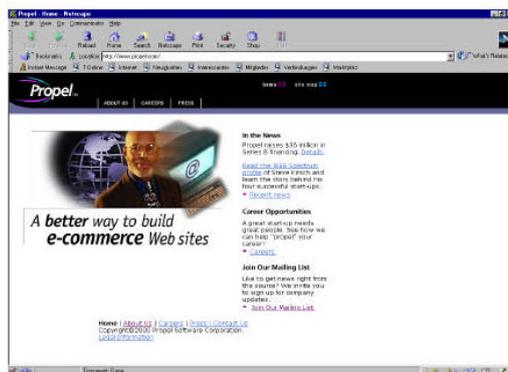
Folge 3

Gottfried Vossen, Universität Münster

In dieser Rubrik möchte ich Ihnen wieder Websites vorstellen, die mir in letzter Zeit aufgefallen sind. Dabei habe ich mir diesmal unter anderem einen thematischen Schwerpunkt vorgenommen, der sich Startups befasst, und zwar genauer mit solchen, an denen hinlänglich bekannte Forscher insbesondere aus der DB- und IS-Szene in geschäftsführender oder beratender Rolle beteiligt sind oder die diese (mit-) gegründet haben. Dabei kann ich natürlich keinen Anspruch auf Vollständigkeit erheben, schon allein deshalb nicht, weil in diesem Bereich ja täglich etwas hinzu kommt (oder auch wieder verschwindet).

In diesen Zusammenhang kann ich als erste Site eine bereits in der vorigen Ausgabe dieser Rubrik vorgestellte in Erinnerung bringen:

www.propel.com



Unter den Geldgebern findet man z.B. Marc Andreessen oder Andy Grove (Chef von Intel), gegründet wurde die Firma von Steve Kirsch, unter den Mitarbeitern und VPs ist u.a. **Mike Carey**, der früher an der University of Wisconsin, Madison, später am IBM Almaden Research Lab war und Datenbankern bestens bekannt ist. Unter den Firmenzielen ist eine Bereitstellung von „amazon in a box“ sowie „to create a revolutionary new way to build e-commerce websites“.

Jetzt aber zu den wirklich neuen Sites dieser Ausgabe: **Yannis Papakonstantinou**, nebenbei noch immer an der University of California in San Diego, gehört zu den Gründern von Enosys Markets in San Diego, einer Firma, die sich mit der Bereitstellung von B2B-Content beschäftigt. Mitgründer ist Vasilis Vassalos (ebenfalls Stanford-Absolvent) und im



Surftipps (22 Folgen, seit 2000)