

## *„Heil dir, oh Oldenburg“*

Bericht zur

### **9. Fachtagung „Datenbanksysteme in Büro, Technik und Wissenschaft“**

(BTW 2001, Universität Oldenburg, 7.-9. März 2001)<sup>1</sup>

*M. Bon, M. Flehmig, T. Härder, J. Kovse,  
W. Mahnke, U. Marder, N. Ritter, H.-P. Steiert, J. Zimmermann  
AG Datenbanken und Informationssysteme  
Fachbereich Informatik, Universität Kaiserslautern  
Postfach 3049, 67653 Kaiserslautern*

#### **Allgemeines zur BTW 2001**

Die BTW 2001, 9. Fachtagung der Reihe „Datenbanksysteme in Büro, Technik und Wissenschaft“, die alle zwei Jahre vom GI-Fachausschuss 2.5 (Rechnergestützte Informationssysteme) veranstaltet wird, fand vom 7. - 9. März 2001 in Oldenburg statt. Das Organisationskomitee um H.-J. Appelrath leistete glänzende Arbeit, so dass ein reibungsloser Ablauf gewährleistet werden konnte. Sowohl Tagungs- und Rahmenprogramm als auch die Stadt Oldenburg selbst werden wohl allen Teilnehmern der Konferenz lange in positiver Erinnerung bleiben. Beim Empfang im Alten Landtag von Oldenburg gelang es Herrn Horst Milde, Landtagspräsident a. D., mit einem gelungenen Kurzvortrag zur Geschichte und Gegenwart der Stadt Oldenburg den Konferenzteilnehmern den besonderen Charme der Stadt Oldenburg zu vermitteln. Musikalisch wurde dann mit der alten Oldenburger Hymne „Heil dir, oh Oldenburg“ zu einem deftigen Grünkohllessen übergeleitet, denn „wer keinen Grünkohl gegessen hat, ist nie in Oldenburg gewesen“. Auch in den Pausen während des Konferenzprogramms, das in dem modernen Hörsaalzentrum der *Carl von Ossietzky Universität Oldenburg* stattfand, wurde die angenehme Atmosphäre durch dezente Klaviermusik untermalt.

Das Konferenzprogramm setzte sich aus einem wissenschaftlichen Programm und einem Industrieprogramm zusammen. Das wissenschaftliche Programm wurde vom Programmkomitee um A. Heuer zusammengestellt. Von 69 eingereichten Beiträgen (53 Lang-, 16 Kurzbeiträge) wurden 15 Lang- und 14 Kurzbeiträge (6 als Langbeiträge eingereichte Papiere wurden als Kurzbeiträge angenommen) akzeptiert. Während sich das Wissenschaftliche Programm über zwei volle Tage (von Mittwoch Mittag bis Freitag Mittag) erstreckte, gab es während des gesamten Donnerstags parallel Industriesitzungen, so dass das Industrieprogramm einen vorher nie dagewesenen Anteil hatte. Letzteres wurde von einem eigenen Programmkomitee um F. Leymann zusammengestellt (2 Beiträge wurden aus eingereichten Beiträgen ausgewählt, 12 Vortragende wurden eingeladen) und erfuhr einen hohen Zuspruch durch die Konferenzteilnehmer. So wurde der größere Hörsaal für das Industrieprogramm genutzt. Insgesamt hatte die

---

1. <http://www.btw2001.de>

BTW 2001 deutlich über 200 Teilnehmer, etwa ein Viertel aus der Industrie und drei Viertel aus dem Hochschulbereich. Damit ist es wohl gelungen, dem Anspruch der BTW-Tagungsreihe, Datenbankforscher und -praktiker zusammenzubringen und den Entwicklungsstand und die Perspektiven neuer Technologien für Informationssysteme in aktuellen Einsatzgebieten zu diskutieren, gerecht zu werden. Inhaltlich waren die Schwerpunkte der Tagung konform mit der neuen Bedeutung des Kürzels *BTW: Business, Technology, Web, Warehouse, and Workflow Support*.

Wesentliche Neuerungen bei der BTW 2001 waren die Verleihung von Dissertationspreisen sowie die Durchführung eines Studi-Programms. Letzteres sowie die traditionell zusammen mit der BTW ausgerichteten Datenbank-Tutorientage werden in eigenen Berichten beschrieben. Im Folgenden werden wir etwas genauer auf die inhaltlichen Schwerpunkte der einzelnen Programmteile eingehen.

### **Eingeladene Vorträge**

Das eigentliche Tagungsprogramm wurde mit einem eingeladenen Vortrag von Michael Rys (Microsoft SQL Server XML Technology Group) begonnen. Die Inhalte des Vortrags orientierten sich sehr stark an den Möglichkeiten, die von Microsoft's SQL Server hinsichtlich einer XML-Unterstützung geboten werden. Es wurde deutlich gemacht, wie XML an der DB-Schnittstelle genutzt werden kann, um relational verwaltete Daten anzufragen bzw. zu manipulieren. Auf diese Weise soll die Grundlage der Einbindung relationaler Datenbanken in über Web-Technologie integrierte, komplexe Systemverbunde geschaffen werden. Während dieser erste eingeladene Vortrag doch sehr stark die Seite der kommerziellen Anbieter repräsentierte, wurde der zweite eingeladene Vortrag von einem Vertreter aus dem Hochschulbereich gehalten. Tamer Özsu (University of Waterloo) berichtete in interessanter Weise über Möglichkeiten und Herausforderungen der Gestaltung von Dokument- und Multimedia-Datenbanken. In diesem Vortrag wurde deutlich, dass Datenbanksysteme auf der einen Seite die geeignete Technologie darstellen, um moderne Anwendungen wie digitale Bibliotheken, Bild- und/oder Video-Repositories zu realisieren, dass aber auf der anderen Seite noch sehr viel zu tun ist, um alle Anforderungen zu erfüllen. Insbesondere Anfragen, deren Suchbereich mehrere unterschiedliche Arten von Dokument- und Multimediadaten umfasst, erfordern die Entwicklung neuer Konzepte und Mechanismen.

### **Wissenschaftliches Programm**

*Sitzung 1: Datenbank-Föderation.* Die erste Sitzung des wissenschaftlichen Programms umfasste zwei Langbeiträge. Diese machten deutlich, dass das Gebiet der Datenbank-Föderationen weiterhin ein anspruchsvolles Forschungspotential aufweist, obwohl bereits eine ganze Reihe von Basismechanismen zur Integration (teil-)autonomer Datenquellen existieren. So beschäftigte sich der Beitrag von der Uni Magdeburg mit der Informationsfusion. Dieser Begriff bezeichnet die Integration und Interpretation heterogener Daten mit dem Ziel der Gewinnung neuer Information einer höheren Qualität. Der Beitrag beschreibt die Architektur einer Workbench, die den gesamten Prozess der Informationsfusion von der Integration heterogener Datenquellen bis zur Ableitung neuer Informationen abdecken soll. Der anschließende

Vortrag von Susanne Busse (TU Berlin) setzte sich mit der Unterstützung der Schemaevolution in föderierten Systemen auseinander. Es wurde dargelegt, wie sichergestellt werden kann, dass eine nachträgliche Schemaevolution nicht die der (initialen) Integration der Teilsysteme zugrundeliegenden Annahmen invalidiert.

*Sitzung 2: Workflow und Alerting.* Der Ulmer Langbeitrag lehrte uns, dass sich Skalierbarkeit, erzielt durch verteilte Workflow-Ausführung, und eine effiziente Handhabung dynamischer Modifikation, die man eher zentral realisieren würde, nicht widersprechen müssen. Es folgten vier Kurzbeiträge. Der erste (Uni Zürich) versuchte zu motivieren, dass organisationsübergreifende Abläufe nicht nur auf festen Kooperationsbeziehungen beruhen müssen, sondern dass vielmehr auch ein Marktplatz für Workflows eine interessante Variante darstellen kann. Workflows werden wie andere Dienstleistungen angeboten. Agenten tummeln sich auf dem Marktplatz und suchen passende Angebote oder bieten ihre Dienste an. Im zweiten der Kurzbeiträge (Uni Jena, IBM Böblingen) wurde eine Methode vorgestellt, wie man ausgehend von vorhandenen (Ablauf-)Protokolldaten zu einem Prozessmodell für ein WfMS gelangt. Der Ansatz eignet sich wohl besonders gut für das Re(verse)-Engineering von Prozessen und damit zur Optimierung. Der nachfolgende Kurzbeitrag (ETH Zürich) motivierte die Notwendigkeit der Überwachung von Abläufen, die nicht im Kern-WfMS laufen, und stellte mit den „Schattenprozessen“ einen vielversprechenden Ansatz vor, wie Zwischenzustände dieser externen Abläufe überwacht werden können. Der letzte (Kurz-)Beitrag (vor dem Grünkohlessen) kam von der FU Berlin und brachte uns die Probleme näher, die bei „Alerting Services“ alleine durch die Beobachtungsstrategie entstehen. Selbst wenn die lokalen Uhren der Systeme von Beobachter und Ereignismelder synchronisiert sind, kann das tatsächliche Ereignis schon geraume Zeit zurückliegen und damit hinfällig sein.

*Sitzung 3: Datenbanken im Internet.* Unter diesem recht allgemeinen Titel wurden zwei Lang- und zwei Kurzbeiträge vorgestellt. Der erste Langbeitrag (Uni Passau, TU München) beschäftigte sich mit der verteilten Metadatenverwaltung für die Anfragebearbeitung auf Internet-Datenquellen. Durch einen Publish/Subscribe-Mechanismus und der damit einhergehenden Dynamik bietet sich das vorgestellte Konzept insbesondere im Bereich E-Commerce an. Der zweite Langbeitrag (Uni Hannover) zeigte, wie objekt-relationale Datenbanktechnologie dazu eingesetzt werden kann, geographische Daten zu verwalten. Daneben wurden eine XML-basierte Sprache für den Austausch von kartographischen Daten entwickelt und generelle Regeln für die Definition von DTDs für stark strukturierte Daten eines objekt-relationalen Datenbankschemas aufgezeigt. In dem darauf folgenden Kurzbeitrag (FH Oldenburg) wurde eine Architektur vorgeschlagen, die der XML-basierten Repräsentation von bewegten geographischen Objekten gilt. Der Schwerpunkt lag hier also bei den beweglichen und damit zeitbezogenen geographischen Daten. Abgeschlossen wurde die Sitzung mit der Vorstellung eines an der Uni Zürich entwickelten Verfahrens, in dem Anfragen auf heterogenen, dynamisch eingebundenen Datenquellen ausgeführt werden können. Der Benutzer stellt dafür lediglich unscharfe Anfragen, die durch „Query Preprocessing“ umgewandelt werden.

*Sitzung 4: Data Warehousing und Datenintegration.* Wieso sind Beiträge zu Data Warehousing und Datenintegration seit Jahren in fast jedem Tagungsband zu finden? Beide Gebiete bieten eine Vielzahl noch offener Fragen und sind von hoher praktischer Relevanz. Für Data Ware-

housing gilt weiterhin, dass nahezu jede Fragestellung aus dem Bereich der Datenbanksysteme berührt wird, von der Bereitstellung geeigneter Benutzungsoberflächen bis zur optimalen Nutzung der Externspeicher. Dies spiegelt sich auch in den Beiträgen (2 Lang-, 2 Kurzbeiträge) dieser Sitzung wider.

Während im ersten Langbeitrag die Verbesserung der Antwortzeiten für OLAP-Anwendungen durch eine Optimierung der internen Verarbeitung behandelt wurde, war der Inhalt des zweiten Langbeitrags die Qualitätssicherung von Daten in Data Warehouse-Systemen. Im ersten Vortrag stellte T. Stöhr (Uni Leipzig) eine Fragmentierungsstrategie für große Datenmengen in OLAP-Anwendungen vor. Auf Basis eines fragmentbasierten Anfrageverarbeitungs- und Datenallokationsmodells hat der Autor ein Kostenmodell und eine Metrik entwickelt, mit denen sich der Aufwand und die Antwortzeiten für OLAP-Anfragen unter der entwickelten Fragmentierungsstrategie abschätzen lassen. Der zweite Langbeitrag (OFFIS) beschäftigt sich mit Konzepten für ein umfassendes Qualitätsmanagement. Er stellt formale Datenqualitätsmetriken und ein wohldefiniertes Vorgehensmodell für das Datenqualitätsmanagement vor. Die auf Metadaten basierende, generische Auslegung des Ansatzes erlaubt es, das Konzept an unterschiedliche Domänen anzupassen.

Der zweite Block der Sitzung, Datenintegration, war mit zwei Kurzbeiträgen etwas schwächer gewichtet. Die weite Verbreitung des Internet und die Vielzahl der darin zu findenden heterogenen Datenquellen werden zukünftig aber eher noch verstärkt für interessante Fragestellungen bezüglich Datenintegration sorgen. In der Einleitung des ersten Kurzbeitrags (Uni Freiburg) wird darauf hingewiesen: das vorgestellte Konzept „promises to be able to easily solve all classical integration problems along with LDAP advantageous network behavior features“. Der zweite Kurzbeitrag (Uni Zürich, ABB) befasst sich mit der Versionskontrolle in föderierten Entwurfsdatenbanken. Diese Problematik dürfte in allen großen Entwurfprojekten aufgrund der vielen dort verwendeten Werkzeuge auftreten. Es wird ein Konzept zur Lösung des Problems der Versionspropagierung beschrieben, wobei besonderer Wert auf ein flexibles und anpassungsfähiges Konzept gelegt wird, das sich an die Bedürfnisse der Domäne anpassen lässt.

*Sitzung 5: Technische Aspekte und Leistungsmessung.* Das zeitlose Thema des Leistungsverhaltens bleibt auch nach der neuen Deutung des BTW-Kürzels von zentraler Bedeutung. In dieser Sitzung wurden neue Ansätze der Leistungsbeurteilung von auf neuer Technologie (ORDBVS, MoM, XML) beruhenden Systemen vorgestellt. Im ersten Langbeitrag (Uni Kaiserslautern) wird der Versuch unternommen, ORDBVS hinsichtlich ihrer Nutzbarkeit in der objektorientierten Systementwicklung zu beurteilen. Dabei wurde auch eine gewisse Kritik am SQL:1999-Standard laut, da dieser zumindest teilweise eine zu halbherzige Integration objektorientierter Konzepte vornimmt, denkt man beispielsweise an Unterstützung von Navigation oder strukturierte Anfrageergebnisse. Der zweite Langbeitrag (Dresdner Bank, Uni des Saarlandes) beschäftigte sich mit der Leistungsbeurteilung von Systemen, die Message-orientierte Middleware einsetzen. Es wurde ein analytisches Modell vorgestellt, auf dessen Basis Antwortzeitmomente vorhergesagt und damit eine geeignete Prioritätenvergabe (für MoM-Komponenten) aus Antwortzeitvorgaben (für das Gesamtsystem) abgeleitet werden können. Der erste der zwei nachfolgenden Kurzbeiträge, beide aus Leipzig, führte den Benchmark-Ansatz XMach-1 ein, mit dessen Hilfe spezielle XML-Systeme bzw. XML-Unterstützung anbietende

relationale DBVS evaluiert werden können. Der zweite Kurzbeitrag beschäftigte sich ausgehend vom Bucky-Benchmark mit der Evaluierung von Verfahren der parallelen Join-Verarbeitung. Insbesondere die sich durch die Nutzung von Referenzattributen, wie sie mittlerweile von ORDBVS unterstützt werden, ergebenden neuen Möglichkeiten sollen mit diesem Ansatz evaluiert werden.

*Sitzung 6: Erweiterbarkeit und Verteilung.* Diese thematisch sehr durchmischte Sitzung umspannte die Themengebiete Daten-Replikation, inhaltsbasierte Indexierungsstrategien, komponentenbasierte Middleware und Untersuchung des praktischen Nutzens objekt-relationaler Erweiterungen. In dem ersten (Hybris, TU Dresden) der beiden Langbeiträge dieser Sitzung wurde ein theoretisches Modell zur Analyse von Replikationsstrategien vorgestellt, welches sich durch ein geschickt gewähltes Kostenmodell mit umfassender Parametrisierbarkeit auszeichnete. Der praktische Nutzen wurde durch Simulationen validiert. Im zweiten Langbeitrag (TU Chemnitz) wurde eine multidimensionale Indexstruktur vorgestellt, die auf einer intelligenten, auf neuronalen Netzwerken basierenden Methode zur Clusterbildung von inhaltlich ähnlichen Daten beruht. Dieses Verfahren erlaubt auch konkrete Aussagen über eine optimale Partitionierung von Daten auf mehreren, parallelen Sekundärspeichern. Der Vorteil der vorgestellten Methode wurde durch Messungen evaluiert. Allerdings waren die Parameter der Vergleichsmessungen sehr zugunsten des neuen Verfahrens gewählt (B\*-Baum vs. R-Baum). Von komponentenbasierter Middleware zur Unterstützung von Interoperabilität über Anwendungsbereiche hinweg handelte der erste Kurzbeitrag (DaimlerChrysler, Uni Stuttgart). Der Komponentenbegriff beinhaltet hier sogar ganze Anwendungspakete, deren Interoperabilitätsmöglichkeiten durch die vorgeschlagene Middleware-Lösung verbessert werden soll. Ein zentraler Aspekt dieses Architekturvorschlags ist der Begriff des „Corporate Repository“ (CR), bei dessen Realisierung das „Simple Object Access Protocol“ (SOAP) eingesetzt wurde. Der letzte Vortrag (Uni Kaiserslautern) dieser Sitzung befasste sich kritisch mit dem praktischen Nutzen der Anwendungsintegration in ORDBS. Zu diesem Zweck wurde der TPC-W-Benchmark für Messungen des Leistungsdurchsatzes verschiedener Architekturvarianten verwendet. Es konnten allerdings keine signifikanten Vorteile der DB-internen Ausführung des Anwendungsservers gegenüber klassischen Architekturvarianten nachgewiesen werden. Erst die sinnvolle Kombination mit anderen objekt-relationalen Erweiterungskonzepten lässt auf einen Nutzen hoffen.

*Sitzung 7: Multimedia und XML.* In dieser Sitzung waren drei sehr unterschiedliche Langbeiträge zusammengefasst, von denen sich allerdings keiner direkt mit Multimedia und nur einer mit dem Thema XML befasste. Im ersten Vortrag stellte Knut Stolze von der Uni Jena den Teil 5 des im Entstehen begriffenen SQL/MM-Standards vor, welcher sich um eine Standardisierung der Verwaltung von Bildern in ORDBVS bemüht. Hierbei wurde deutlich, dass der Standard letztendlich vor der Komplexität und Vielschichtigkeit des Themas Bildverarbeitung und -verwaltung kapituliert und nur eine ausgesprochen minimalistische Lösung propagiert. Der Vortrag machte jedoch Hoffnung auf baldige Erweiterungen und Verbesserungen. Im zweiten Vortrag (Uni des Saarlandes) wurde eine Anfragesprache für XML-Dokumente vorgestellt, die im Gegensatz zu Sprachen wie XPATH oder XML-QL eine ähnlichkeitsbasierte Suche mit relevanzbasiertem Ranking des Ergebnisses unterstützt. Die Autoren legen dar, dass eine auf dieser XXL genannten Sprache basierende Web-Suchmaschine wesentlich präzisere

Ergebnisse liefern könne als heutige Suchmaschinen. Auch der dritte Vortrag (Uni Halle, Uni Leipzig) befasste sich mit ähnlichkeitsbasierter Suche. Der klassische merkmalsbasierte Ansatz wurde hier auf 3D-Objekte, wie sie hauptsächlich in der Computergrafik vorkommen, angewandt. Hierbei wurde nicht nur gezeigt, wie die Merkmalsvektoren gewonnen werden, sondern zudem die Wirksamkeit des Verfahren durch eine kurzweilige Prototyp-Vorführung demonstriert.

*Sitzung 8: Finale.* Drei Themen, wie sie unterschiedlicher kaum sein können, beleuchteten in dieser abschließenden Sitzung noch einmal schlaglichtartig die ganze Bandbreite der BTW 2001. Der letzte Vortrag, in dem Rüdiger Reinecke (IBM) die ganze Problematik der kommenden Euro-Währungsumstellung ausbreitete, repräsentierte hierbei das B wie „Business“. Rainer Mantheys (Uni Bonn) aufschlussreicher Vortrag über die Untiefen und Fallgruben der bekannten und scheinbar wohl durchdachten Integritätssicherungsmechanismen von SQL, stand für das T wie „Technology“. Und der mittlere Vortrag, in dem Ulrich Marder (Uni Kaiserslautern) ein Datenmodell für die Verwaltung und Verarbeitung multimedialer Daten in Web-basierten Informationssystemen vorstellte, konnte schließlich zumindest noch ein Drittel des W, nämlich das „Web“, abdecken. Als kleine Anregung für die nächste BTW: Würde man das W umdrehen, dann hätten wir auch noch ein M wie „Mining, Metadata, und Multimedia“...

## **Industrieprogramm**

*Sitzung A: Datenintegration.* Das Thema Datenintegration war auf der BTW sowohl im wissenschaftlichen als auch im Industrieprogramm vertreten. Der erste Vortrag war durch die zunehmende informationstechnische Integration und Zusammenarbeit von Unternehmen motiviert. Ronald Grzybowski (Oracle) gab eine sehr gute Klassifikation der dabei auftretenden und zu lösenden Probleme. Leider ist der Vortrag nicht mit einem Beitrag im Tagungsband vertreten. Aus der Rolle des Kunden wurde im zweiten Vortrag (Klaudia Hergula, DaimlerChrysler) die Integrationsproblematik in einem weltweit operierenden Unternehmen mit seinen vielen heterogenen Werkzeugen und Datenhaltungssystemen erläutert. Da viele Datenquellen den Zugriff auf die von ihnen verwalteten Daten nur über funktionale Schnittstellen erlauben, wird eine Integration auf der Ebene dieser Schnittstelle angestrebt, ohne dabei die Vorteile einer Anfragesprache mit generischen Operatoren aufgeben zu wollen. Im Vortrag wurden zwei Integrationstechnologien analysiert: Wrapper in SQL/MED und Konnektoren in J2EE. Da beide für sich allein die Anforderungen nicht erfüllen, wurde die „silver bullet“ der Datenintegration gefordert: Konnektoren, mit denen sich sowohl die Simulation von SQL-Operatoren als auch die funktionale Integration einfach realisieren lassen. Der letzte Vortrag (Telekom) dieser Sitzung war ein Erfahrungsbericht zu „DARWIN“, einem Data-Warehouse-Projekt der Deutschen Telekom. Es wurden, ausgehend von der verwendeten Hardware, die vier Ebenen des Systems vorgestellt.

*Sitzung B: XML und Datenbanken.* Diese zweite Industriesitzung wurde von Walter Waterfeld (Software AG) mit einem Vortrag über die Funktionalität und die Architektur von Tamino, das „native“ XML-Datenbanksystem der Software AG, eröffnet. Anschließend hielt Susan Malaika einen Vortrag über den „XML Support in IBM's DB2“, also über den XML-Extender von IBM. Sie wusste ihr Publikum bereits zu Beginn zu überzeugen, und zwar mit original „Californian

Candies“, die sie an die Zuhörerschaft verteilte. Der XML-Extender soll laut Susan Malaika eine hohe Akzeptanz bei den Kunden haben. Nach dem herstellerseitigen Blick auf XML-Datenbanken rundete der anwendungsseitige Blick vertreten durch die Dresdner Bank diese Session ab. Im Vortrag über Vision und Realität XML-basierter Content-Analyse wurden die Möglichkeiten und Probleme beschrieben, die bei der Verwendung von XML in einem großen Unternehmen auftreten. Die Vision ist, mit Hilfe von XML relevante Geschäftsinformationen, die in Textdokumenten gespeichert sind, effizient auffinden zu können. In der Realität ist aber laut Dresdner Bank die datenbankseitige Unterstützung noch verbesserungswürdig. Insbesondere wurde die Forderung nach einer Unterstützung zur automatischen Abbildung und Konvertierung stark strukturierter Dokumente laut.

*Sitzung C: Mining und Dokument-Management.* Data Mining beschäftigt sich mit der Suche nach sogenannten „Patterns“ in vorhandenen Daten, um damit gegenwärtige Entscheidungsfindung zu unterstützen. Der Vortrag von Andreas Arning (IBM) stellte die grundlegenden Methoden des Data Mining (Ermittlung von Assoziationsregeln, Segmentierung und Klassifizierung) vor und umriss damit die Möglichkeiten des entsprechenden Data-Mining-Produkts der IBM. Die beiden folgenden Vorträge der Session waren ebenfalls sehr produktorientiert. Zunächst legte Henning Hinrichs dar, welchen Nutzen das Produkt von InfoOffice für auf Java und XML basierende Web-Anwendungen haben soll, wobei hier besonderer Wert auf Flexibilität und Anpassbarkeit gelegt wird. Danach erläuterte Pal Rujan „SERbrainware“. Es handelt sich hierbei um ein intelligentes, lernendes Werkzeug zur Dokumentverwaltung, dessen Fähigkeiten insbesondere der Analyse und Klassifikation unstrukturierter Texte (z. B. Webseiten, eMails, Papierdokumente) dienen.

*Sitzung D: Workflows und Kommunikation.* Zunächst stellte Johannes Klein (Microsoft) das Produkt „BizTalk Server 2000“ vor, das zur Integration unternehmensweiter Geschäftsprozesse durch Austausch von Dokumenten dient. Wesentlich waren dabei die Vorstellung eines Editors zur Modellierung und die Tatsache, dass XML als Austauschformat zum Zuge kommt. Für Verwunderung sorgte die Aussage, ein Rollenkonzept sei zwar angedacht, aber noch nicht umgesetzt. Verträge über das Internet abschließen ist schnell und effizient. Leider treten dabei eine Reihe von Problemen auf, die nicht zu unterschätzen sind, beginnend bei einer sicheren elektronischen Unterschrift bis hin zur Geheimhaltung vertraulicher Daten. In einem praktischen Erfahrungsbericht präsentierte Michael Lawo (select IT Consultant Services) seine Erfahrungen mit der DiViO-Plattform, die zur Abwicklung von Maklerverträgen entwickelt wurde. Interessant waren auch seine Ausführungen zur Akzeptanz der Sicherheitsvorrichtungen durch die Benutzer: der Wunsch nach Komfort übersteigt den nach Sicherheit bei weitem. Ist eine Maßnahme zu kompliziert anzuwenden, wird sie nicht benutzt, egal wie sinnvoll sie zu sein scheint. Im letzten Vortrag (Dresdner Bank, Uni des Saarlandes) wurde ein Werkzeug zur automatischen Umsetzung von Geschäftsprozessen in ausführbare Workflows vorgestellt. Dabei werden Geschäftsprozesse, die mit Hilfe von „ARIS“ modelliert wurden, in FDL, die Beschreibungssprache des WfMS „MQSeries Workflow“, transformiert. Eines der herausragenden Probleme dabei ist, die recht freie Struktur von Schleifen in ARIS in das Block-Konzept von FDL zu überführen.

*Sitzung E: Portale und Data Warehouses.* Am späten Donnerstag fand das Industrieprogramm seinen Abschluss in dieser lediglich aus einem Vortrag bestehenden Sitzung. Dieser Beitrag (Uni Oldenburg, IONA Professional Services, Materna) betrachtet den Prozess der Entwicklung von Portalen als Vorgang der Integration bestehender Informationsdienste. Es werden Anforderungen an Portale aufgestellt, kommerzielle Systeme (zur Entwicklung von Internet-Portalen) entsprechend beurteilt und neue Mechanismen für die Kopplung mehrerer Portale gefordert.

### **BTW-Dissertationspreise**

Als eine der Neuerungen wurden die BTW-Dissertationspreise vergeben. Die IBM Deutschland war der Sponsor für zwei Preise von je 5.000 DM, mit denen hervorragende Dissertationen aus dem Bereich „Datenbanken und Informationssysteme“ ausgezeichnet wurden. Ein Preis wurde an Achim Kraiß für seine Dissertation „Hierarchische Speicherverwaltung für Informationssysteme mit Tertiärspeicher“ verliehen, die er an der Fakultät für Mathematik und Informatik, Universität des Saarlandes, anfertigte. In dieser Arbeit setzte er sich in umfassender Weise mit dem Problem hoher Latenzzeiten beim Zugriff auf Tertiärspeicher auseinander. Die zentrale Frage war, wie potentielle Antwortzeiteinbußen soweit wie möglich eliminiert werden können. Achim Kraiß untersuchte mit strenger und neuartiger mathematischer Modellierung zwei Problemkreise, von denen jeder für sich für eine ausgezeichnete Dissertation ausgereicht hätte. Die grundlegenden Ideen betrafen dabei zum einen Prefetching/Caching von Speicherobjekten und zum anderen Pipelining für Datenströme. Zur Lösung des ersten Problems entwickelte er, vom LRU-K-Verfahren ausgehend, eine Methode, die „temperaturorientiertes“ Caching und Prefetching in einer mehrstufigen Speicherhierarchie integriert. Ein typisches Beispiel für die Anwendung dieser Methode stellt die navigierende Zugriffsfolge auf Daten eines Web-Servers, der eine digitale Bibliothek verwaltet, und ihre Optimierung durch spekulatives Prefetching dar. Zur Lösung des zweiten Problems optimierte er Pipelining für Operationen auf Datenströmen. Eine mögliche Anwendung dafür ist der Band-Platte-Verbund (Tape-Disk-Join), wie er häufig in Data-Warehouse-Systemen abzuwickeln ist. Bei dieser Anwendung kommt der „äußere“ Datenstrom von Magnetband, der mit dem „inneren“ Datenstrom von Magnetplatte synchronisiert werden muss.

Ebenfalls ausgezeichnet wurde Frank Wietek für seine Dissertation „Intelligente Analyse multidimensionaler Daten in einer visuellen Programmierumgebung und deren Anwendung in der Krebsepidemiologie“, die am Fachbereich Informatik, Universität Oldenburg entstanden ist. In seiner Arbeit beschäftigte er sich mit der Analyse multidimensionaler Daten, einem Themengebiet, das heute oft durch Schlagworte wie OLAP, Data Mining, Data Warehousing oder Decision Support charakterisiert wird. Typische Datenanalysen erfordern lange Auswertungsketten, bei denen die einzelnen Analyseschritte aufeinander aufbauen und vom Benutzer gesteuert werden. Im Vergleich zu herkömmlichen Verfahren erreicht sein Ansatz zur explorativen Datenanalyse eine verbesserte Abwicklung der Auswertungsketten bei ausgeprägter Interaktivität. Als wesentliche Aspekte seines Ansatzes entwickelte er eine datenflussbasierte (visuelle) Programmierumgebung VIOLA, ein mehrdimensionales Datenmodell MADEIRA und eine ver-



besserte semantische Beschreibung der mehrdimensionalen Klassifikation von Objektmengen. Am Beispiel des Anwendungsgebietes der Krebsepidemiologie zeigte er, wie die explorative Datenanalyse als kooperatives Zusammenwirken von Mensch und Rechner unterstützt werden kann.

### **Fazit und Ausblick**

Schon aufgrund verschiedener Neuerungen (separates Industrieprogramm, Dissertationspreis und Studi-Programm) setzte die Oldenburger BTW neue Akzente. Das wissenschaftliche Programm zeigte ein durchgehend hohes Niveau. Damit gelang es dem Programmkomitee, der Neudeutung der Buchstaben im Kürzel „BTW“ eine dazu passende, gelungene Programmgestaltung folgen zu lassen. Auch das Industrieprogramm war ansprechend. Dennoch ist es anzustreben, die Industrievertreter beim nächsten Mal zu einer höheren Zahl von Einreichungen zu bewegen, so dass auch hier ein Begutachtungsprozess stattfinden kann. Positiv anzumerken ist, dass aufgrund der (fast) gleichen Anteile von Industrie- und wissenschaftlichen Beiträgen eine konstruktive „Spannung“ zwischen Industrie und Forschung zu spüren war. Als Beispiel sei hier der sehr ansprechende Vortrag von Rainer Manthey genannt, dem es gelang, die Wichtigkeit von Integritätsmechanismen in relationalen DBVS dem mangelnden Interesse der DBVS-Hersteller, diese Mechanismen in die Standards oder sogar in die Produkte einfließen zu lassen, gegenüberzustellen. In der zugehörigen Diskussion wurde deutlich, dass das (DBVS-)Hersteller-Argument, Integritätssicherung ginge in der Meinung ihrer Kunden zu sehr auf Kosten der Effizienz, kein Grund dafür sein darf, keine geeigneten Mechanismen in den Systemen anzubieten, da sehr viele Anwendungen mit hohen, beispielsweise rechtlichen Anforderungen die Integritätsmechanismen oberhalb des DBS selbst implementieren müssen. Weiterhin wurde in einigen der Pausendiskussionen darüber spekuliert, inwieweit die Datenbankforschung derzeit ihrer ursprünglichen Vorreiterrolle gegenüber der Industrie noch gerecht wird. Betrachtet man beispielsweise den SQL:1999-Standard, so scheint es derzeit eher so zu sein, dass die Forschung die Vorgaben aus der Industrie zu kritiklos als Grundlage eigener Arbeiten übernimmt. Vielleicht gelingt es auch bei den zukünftigen Austragungen der BTW durch konstruktive Kritik zwischen Forschung und Industrie eine weitere positive „Befruchtung“ zu erreichen. Insgesamt ist die BTW 2001 als eine sehr gelungene Veranstaltung zu beurteilen, was natürlich vor allem der sehr guten Arbeit der Programmkomitees und Organisatoren zu verdanken ist. Damit wurde die „Messlatte“ für die nächste Tagung der BTW-Reihe, die 2003 in Leipzig stattfinden wird, sehr hoch gelegt.