

Gerhard Weikum
Harald Schöning
Erhard Rahm
(Herausgeber)



Datenbanksysteme für
Business, Technologie und Web
(BTW)

10. GI-Fachtagung
Leipzig, 26.-28. Februar



Herausgeber

Gerhard Weikum
Universität des Saarlandes
Im Stadtwald
D-66123 Saarbrücken
weikum@cs.uni-sb.de

Harald Schöning
Software AG
Uhlandstraße 12
D-64297 Darmstadt
Harald.Schoening@softwareag.com

Erhard Rahm
Universität Leipzig
Institut für Informatik
Augustusplatz 10-11
D-04109 Leipzig
rahm@informatik.uni-leipzig.de

Vorwort

Die 10. BTW-Tagung der Gesellschaft für Informatik (GI) findet vom 26. bis zum 28. Februar 2003 im traditionsreichen Leipzig statt, der Stadt von Johann Sebastian Bach und Auerbachs Keller, der Stadt der Völkerschlacht anno 1813 und der Stadt, in der die deutsche Wiedervereinigung mit den Montagsdemonstrationen im Herbst 1989 ihren Anfang nahm. Die BTW'2003 wird in den Räumen der Universität Leipzig am Augustusplatz ausgerichtet, direkt im Zentrum der Stadt. Die Universität Leipzig wurde vor nahezu 600 Jahren gegründet (1409) und verbindet Tradition und Innovation in besonderer Weise. Die Informatik wurde nach der Wende aufgebaut und hat sich mittlerweile an der Universität und in der deutschen Informatiklandschaft sehr gut etabliert.

Die BTW ist die zentrale Tagung des GI-Fachbereichs Datenbanken und Informationssysteme: Wissenschaftler, Praktiker und Anwender treffen sich alle zwei Jahre auf dieser Tagung. Die Tradition der BTW hat ihren Anfang in der ersten Tagung 1985 in Karlsruhe zu einer Zeit, in der sich Datenbanksysteme von den klassischen betrieblichen Einsatzfeldern zu Anwendungen in Büro, Technik und Wissenschaft entwickelten, daher der Name BTW. Heute ist Datenbanktechnologie der wichtigste Stützpfeiler der Softwarebranche generell; sie ist unverzichtbar für globale Business-Portals und elektronische Prozesse, als Infrastruktur in der Telekommunikation und anderen eingebetteten Technologien, als skalierbares Rückgrat digitaler Bibliotheken und vieler Data-Mining-Werkzeuge sowie für viele Arten von Web-Anwendungen. Im Zeitalter der Informationsexplosion, Virtualisierung und Globalisierung kommen auf die Datenbanktechnologie aber weiterhin neue Herausforderungen zu, wenn es etwa um die Informationsintegration aus heterogenen, verteilten Datenquellen, Internet-weites kollaboratives Data-Mining oder die Gestaltung der Vision eines "Semantic Web" geht. Die Veränderungen bewegen den GI-Fachbereich Datenbanken und Informationssysteme, die Abkürzung BTW neu zu definieren: ab der diesjährigen 10. Ausgabe steht der Name für Datenbanksysteme für Business, Technologie und Web.

Das wissenschaftliche Programm der BTW'2003 umfasst die eingeladenen Beiträge dreier international führender Forscher - Frank Leymann, Alon Halevy und Peter Lokemann - sowie 14 Lang- und 17 Kurzbeiträge, die das Programmkomitee aus 78 Einreichungen ausgewählt hat. Das Industrieprogramm umfasst 7 Lang- und 4 Kurzbeiträge, die vom Industriekomitee aus den Einreichungen ausgewählt und teilweise aktiv angeworben wurden. Wie schon in früheren BTW-Jahren wird dieses Kernprogramm ergänzt durch die Datenbanktutorientage der Deutschen Informatikakademie, das Studierendenprogramm, bei dem hervorragende Diplom- und Studienarbeiten vorgestellt werden, zwei Workshops zu den Themen "Datenbanken und E-Learning" und "Knowledge Discovery" sowie die Verleihung der von der Firma IBM gestifteten Dissertationspreise des GI-Fachbereichs Datenbanken und Informationssysteme.

In Leipzig findet die BTW erstmals zeit- und ortsgleich mit einer zweiten großen Informatiktagung, der 13. GI-Fachtagung über Kommunikation in verteilten Systemen, kurz KiVS, statt. BTW und KiVS starten gemeinsam mit der feierlichen Eröffnung im Gewandhaus und dem ersten eingeladenen Vortrag über Web-Services. Wir hoffen, dass dieses Bündnis zweier traditionsreicher Konferenzen dazu ermutigt, über die Grenzen des engeren Fachgebiets zu schauen und den Blick für die spannenden Verflechtungen und potentiellen Synergien zwischen Kommunikations- und Datenbanktechnologien zu schärfen. Ein weiteres Novum in der BTW-Geschichte ist, dass es erstmals zusätzlich zu dem Ihnen vorliegenden Tagungsband in der LNI-Reihe der GI einen elektronischen Tagungsband gibt, der direkt vom Trierer DBLP-Server erreichbar ist. Wir hoffen, dass diese zeitgemäße Neuerung zusammen mit dem zunehmenden Trend, Beiträge in englischer Sprache zu verfassen, die internationale Sichtbarkeit der BTW noch weiter verbessert.

Zum Gelingen einer Konferenz tragen viele bei. Wir möchten uns herzlich bedanken bei allen Vortragenden, insbesondere den eingeladenen Rednern, und ihren Koautoren, sowie bei denjenigen, die Beiträge eingereicht haben, auch wenn dann vielleicht das letzte bisschen Fortune bei der sehr selektiven Auswahl gefehlt hat. Ebenso herzlich danken wir den Mitgliedern des Programmkomitees und des Industriekomitees sowie den zusätzlichen externen Gutachtern für ihre Sorgfalt und Mühe, sowie den Mitgliedern des Organisationskomitees und den Verantwortlichen für Tutorien, Studierendenprogramm, Dissertationspreisen und Workshops. Besonderer Dank gebührt Ulrike Greiner, Thomas Kudraß, Katrina Leyking und Carsten Simon, die den Tagungsleiter und den Programmkomiteevorsitzenden mit großem Engagement organisatorisch unterstützt haben. Last but not least möchten wir den Sponsoren der diesjährigen BTW-Tagung für die finanzielle Unterstützung sehr herzlich danken. Es ist sehr beachtlich und erfreulich, dass es gerade in den schwierigen Zeiten der aktuellen Rezession gelungen ist, beim Sponsoring ein Rekordergebnis in der Geschichte der BTW zu erzielen.

Wir begrüßen Sie, liebe Tagungsteilnehmer, herzlich zur BTW'2003 in Leipzig und hoffen, dass Sie Vorträge und Rahmenprogramm spannend und erkenntnisreich finden. Um die in wenigen Tagen beginnende Zeit bis zur nächsten BTW-Konferenz 2005 zu überbrücken, empfehlen wir eingefleischten BTW-Fans und denjenigen, die es mit der Leipziger Konferenz hoffentlich werden, die Web-Seite mit der URL <http://www.btw-konferenz.de> auch in Zukunft häufiger zu besuchen.

Erhard Rahm, Tagungsleiter
Harald Schöning, Vorsitzender des Industriekomitees
Gerhard Weikum, Vorsitzender des Programmkomitees

Tagungsleitung:

Erhard Rahm, Universität Leipzig

Programmkomitee:

Gerhard Weikum (Vorsitz), Universität des Saarlandes
Hans-Jürgen Appelrath, Universität Oldenburg
Thomas Brinkhoff, FH Oldenburg
Stefan Conrad, Ludwig-Maximilian-Universität München
Stefan Deßloch, Universität Kaiserslautern
Klaus Dittrich, Universität Zürich
Peter Fankhauser, Fraunhofer IPSI und Infonyte, Darmstadt
Norbert Fuhr, Universität Duisburg
Torsten Grust, Universität Konstanz
Andreas Heuer, Universität Rostock
Stefan Jablonski, Universität Erlangen
Gerti Kappel, Technische Universität Wien
Alfons Kemper, Universität Passau
Arnd Christian König, Microsoft Research, Redmond
Donald Kossmann, Technische Universität München
Thomas Kudraß, HTWK Leipzig
Klaus Küspert, Universität Jena
Alexander Maedche, FZI Karlsruhe
Volker Markl, IBM Almaden Research Center, San Jose
Bernhard Mitschang, Universität Stuttgart
Erhard Rahm, Universität Leipzig
Kai Rannenber, Universität Frankfurt
Manfred Reichert, Universität Ulm
Norbert Ritter, Universität Hamburg
Kai-Uwe Sattler, Univ. Magdeburg
Bernhard Seeger, Universität Marburg
Agnes Voisard, Freie Universität Berlin
Gottfried Vossen, Universität Münster
Mechtild Wallrath, DZ-Bank AG, Karlsruhe
Walter Waterfeld, Software AG
Roger Weber, ETH Zürich
Mathias Weske, Hasso-Plattner-Institut Potsdam
Robert Winter, Universität St. Gallen

Industriekomitee:

Harald Schöning (Vorsitz), Software AG
Achim Kraiß, SAP
Frank Leymann, IBM
Thomas Ruf, GfK Marketing Services
Uta Störl, Dresdner Bank

Studierendenprogramm:

Andreas Oberweis, Universität Frankfurt
Gunter Saake, Universität Magdeburg

Tutorienprogramm:

Klaus Küspert, Universität Jena

Organisationskomitee:

Erhard Rahm, Universität Leipzig
Timo Böhme, Universität Leipzig
Ulrike Greiner, Universität Leipzig
Thomas Kudraß, HTWK Leipzig
Dieter Sosna, Universität Leipzig
Myra Spiliopoulou, HHL Leipzig

**Leitungsgremium des GI-Fachbereichs Datenbanken und Informationssysteme
und Gutachter für die Dissertationspreise:**

Hans-Jürgen Appelrath, Universität Oldenburg
Stefan Conrad, Ludwig-Maximilian-Universität München
Peter Dadam, Universität Ulm
Klaus Dittrich, Universität Zürich
Nobert Fuhr, Universität Duisburg
Theo Härder, Universität Kaiserslautern
Andreas Heuer, Universität Rostock
Klaus Küspert, Universität Jena
Andreas Oberweis, Universität Frankfurt
Erhard Rahm, Universität Leipzig
Gunter Saake, Universität Magdeburg
Gottfried Vossen, Universität Münster
Mechtild Wallrath, DZ-Bank AG, Karlsruhe
Gerhard Weikum, Universität des Saarlandes
Mathias Weske, Hasso-Plattner-Institut Potsdam

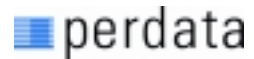
Weitere Gutachter

Thomas Aden
Evguenia Altareva
Philipp Bender
Bettina Berendt
Martin Bernauer
Timo Böhme
Dietrich Boles
Ilvio Bruder
Zhiyuan Chen
Michael Christoffel
Suo Cong
Peter Dadam
Daniela Damm
Hai Hong Do
Anca Dobre
Ruxandra Domenig
Stefan Figge
Michael Gesmann
Christoph Gollmick
Marco Grawunder
Ulrike Greiner
Farshad Hakimpour
Arne Harren
Christoph Hartwich
Axel Herbst
Holger Hinrichs

Thomas B. Hodel
Andreas Hotho
Jens Hündling
Markus Keidl
Maurice van Keulen
Toralf Kirsten
Thomas Klement
Meike Klettke
Birgitta König-Ries
Frank Köster
Gerhard Kramler
Jens Lechtenböcker
Wolfgang Lehner
Hans-Peter Leidhold
Jens Lufter
Eitel von Maur
Christian Meiler
Jürgen Meister
Sergey Melnik
Holger Meyer
Boris Motik
Jutta Mülle
Robert Müller
Felix Naumann
Frank Oldenettel
Ilia Petrov

Denny Priebe
Andreas Rann
Ralf Rantza
Joachim Schelp
Jürgen Schlegelmilch
Frank Schön
Gregor Schrott
Hilmar Schuschel
Andre Seifert
Stefan Seltz
Markus Sinnwell
Steffen Skatulla
Marko Smiljanic
Günther Specht
Bernhard Stegmaier
Knut Stolze
Heiko Tapken
Thomas Tesch
Raphael Volz
Gunnar Weber
Christian Wiesner
Xiangru Yuan
Bernhard Zeller
Andre Zeitz
Patrick Ziegler

Sponsoren



**BTW-Dissertationspreis
gestiftet von der IBM Deutschland GmbH**

Die bei der BTW'2001 in Oldenburg erstmals durchgeführte Auszeichnung hervorragender Dissertationen im Gebiet der Datenbanken und Informationssysteme wird in Leipzig dank des erneuten finanziellen Engagements der Firma IBM fortgeführt. Für das Auswahlverfahren wurden nur Arbeiten akzeptiert, die im Zeitraum 10/2000 bis 9/2002 mit der Promotionsprüfung abgeschlossen wurden. Die Kandidatinnen und Kandidaten wurden von ihren betreuenden Professoren nominiert, eine Selbstbewerbung war nicht möglich. Gutachter über die eingereichten Arbeiten waren die Mitglieder des Leitungsgremiums des GI-Fachbereichs Datenbanken und Informationssysteme unter der Leitung von Theo Härder (Universität Kaiserslautern). Bewertungskriterien waren die wissenschaftliche Qualität und die Bedeutung für das Gebiet.

Die Preisverleihung findet zum Abschluss der BTW'2003 statt. Die beiden Preisträger halten jeweils einen Kurzvortrag über ihre Dissertation. Schriftliche Kurzfassungen sind in diesem Tagungsband enthalten. Die Dissertationspreise gehen in diesem Jahr an

Reinhard Braumandl (Universität Passau) für die Dissertation
"Quality of Service and Optimization in Data Integration Systems"

und

Jens Lechtenböcker (Universität Münster) für die Dissertation
"Data Warehouse Schema Design".

Inhaltsverzeichnis

Eingeladene Beiträge

- Frank Leymann (IBM Germany, Böblingen): *Web Services: Distributed Applications Without Limits* 2
- Alon Halevy (University of Washington, Seattle): *Data Integration: A Status Report* 24
- Peter Lockemann (Universität Karlsruhe): *Information System Architectures: From Art to Science* 30

Wissenschaftliches Programm

Anfrageverarbeitung

- Goetz Graefe (Microsoft): *Executing Nested Queries* 58
- Kai-Uwe Sattler, Ingolf Geist, Rainer Habrecht, Eike Schallehn (Uni Magdeburg): *Konzeptbasierte Anfrageverarbeitung in Mediatorsystemen* 78
- Pedro Jose Marron, Georg Lausen, Martin Weber (Uni Freiburg): *Catalog Integration Made Easy* (Kurzbeitrag) 98

XML

- Wolfgang Lehner, Florian Irmert (TU Dresden, Uni Erlangen): *XPath-Aware Chunking of XML-Documents* 108
- Bernhard Zeller, Axel Herbst, Alfons Kemper (Uni Passau, SAP): *XML-Archivierung betriebswirtschaftlicher Datenbank-Objekte* 127
- Wolfgang May, Dimitrio Malheiro (Uni Göttingen, Uni Freiburg): *A Logical, Transparent Model for Querying Linked XML Documents* (Kurzbeitrag) 147
- Markus Kalb, Kerstin Schneider, Günther Specht (Uni Ulm, EML): *T-XPath: ein zeitliches Modell für XML-Datenbanken* (Kurzbeitrag) 157

Information Retrieval

- Andreas Henrich, Günter Robbert (Uni Bayreuth): *Ein Ansatz zur Übertragung von Rangordnungen bei der Suche auf strukturierten Daten* 167
- Nadine Schulz, Ingo Schmitt (Uni Magdeburg): *Relevanzziehung in komplexen Multimediaanfragen* (Kurzbeitrag) 187

Stefan Siersdorfer, Sergej Sizov (Uni Saarbrücken): <i>Konstruktion von Featureräumen und Metaverfahren zur Klassifikation von Webdokumenten</i> (Kurzbeitrag)	197
Alexander Hinneburg, Dirk Habich (Uni Halle): <i>Ähnlichkeitssuche in Musik-Datenbanken mit Hilfe von Visualisierungen</i> (Kurzbeitrag)	207
Anja Theobald (Uni Saarbrücken): <i>An Ontology for Domain-oriented Semantic Similarity Search on XML Data</i> (Kurzbeitrag)	217
 <i>Konsistenz und Standards</i>	
Martin Kempa, Volker Linnemann (Uni Lübeck): <i>Type Checking in XOBÉ</i>	227
Knut Stolze (Uni Jena): <i>SQL/MM Spatial - The Standard to Manage Spatial Data in a Relational Database System</i>	247
Olaf Buck, Volker Linnemann (Fa. Pietsch, Uni Lübeck): <i>Verbalisierung von Datenbanktransaktionen</i> (Kurzbeitrag)	265
Jens Lufter (Uni Jena): <i>Integritätsbedingungen für komplexe Objekte in objektrelationalen Datenbanksystemen</i> (Kurzbeitrag)	275
 <i>Indexstrukturen</i>	
Hans-Peter Kriegel, Martin Pfeifle, Marco Pötke, Thomas Seidl (LMU München, sd&m AG, RWTH Aachen): <i>The Paradigm of Relational Indexing: a Survey</i>	285
Michael G. Bauer, Frank Ramsak, Rudolf Bayer (TU München): <i>Multidimensional Mapping and Indexing of XML</i>	305
Carsten Kleiner, Udo Lipeck (Uni Hannover): <i>OraGiST - How to Make User-Defined Indexing Become Usable and Useful</i> (Kurzbeitrag)	324
Erik Buchmann, Klemens Boehm (Uni Magdeburg): <i>Effizientes Routing in verteilten skalierbaren Datenstrukturen</i> (Kurzbeitrag)	334
 <i>Verteilte Systeme und Web Services</i>	
Markus Keidl, Stefan Seltzsam, Christof König, Alfons Kemper (Uni Passau): <i>Kontext-basierte Personalisierung von Web Services</i>	344
Erich Müller, Peter Dadam, M. Feltes (Uni Ulm, Daimler-Chrysler): <i>Efficient Assembly of Product Structures in Worldwide Distributed Client/Server Environments</i>	364
Jens Graupmann, Michael Biwer, Patrick Zimmer (Uni Saarbrücken): <i>Towards Federated Search Based on Web Services</i> (Kurzbeitrag)	384

Dietrich Boles (Uni Oldenburg): *Von digitalen Bibliotheken zu Online-Shops mit digitalen Produkten* (Kurzbeitrag) 394

Konsistenz, Replikation, Workflow

Lutz Schlesinger, Wolfgang Lehner (Uni Erlangen, TU Dresden): 404
Konsistenzquantifizierung in Grid-Datenbanksystemen

Ulrike Greiner, Erhard Rahm (Uni Leipzig): *WebFlow: Ein System zur flexiblen Ausführung webbasierter, kooperativer Workflows* (Kurzbeitrag) 423

Markus Bon, Norbert Ritter, Hans-Peter Steiert (Uni Kaiserslautern, Uni Hamburg, Daimler-Chrysler): *Modellierung und Abwicklung von Datenflüssen in unternehmensübergreifenden Prozessen* (Kurzbeitrag) 433

Heiko Niemann, Wilhelm Hasselbring, Michael Huelsmann, Oliver Theel (OFFIS, Uni Oldenburg): *Realisierung eines adaptiven Replikationsmanagers mittels J2EE-Technologie* (Kurzbeitrag) 443

Christoph Gollmick (Uni Jena): *Nutzerdefinierte Replikation zur Realisierung neuer mobiler Datenbankanwendungen* (Kurzbeitrag) 453

Datenbanktechnologie in den Lebenswissenschaften

Stephan Heymann, Katja Tham, Peter Rieger, Johann-Christoph Freytag (Humboldt-Uni Berlin): *Rechnergestützte Suche nach Korrelationen in komplexen Datensätzen der Biowissenschaften* 463

Hong Hai Do, Toralf Kirsten, Erhard Rahm (Uni Leipzig): *Comparative Evaluation of Microarray-based Gene Expression Databases* 482

Dissertationspreise

Reinhard Braumandl (Uni Passau): *Quality of Service and Optimization in Data Integration Systems* 503

Jens Lechtenbörger (Uni Münster): *Data Warehouse Schema Design* 513

Industrieprogramm

Industrieprogramm: Integration und Konfiguration

Udo Nink, Stefan Schäfer (CronideSoft AG): *The IOP Approach to Enterprise Frameworks* 524

Anja Schanzenberger, Colin Tully, D.R. Lawrence (GfK): *Überwachung von Aggregationszuständen in verteilten komponentenbasierten Datenproduktionssystemen* 544

Robert Marti (Swiss Re): *Information Integration in a Global Enterprise: Some Experiences from a Financial Services Company* (Kurzbeitrag) 558

Marc Bastien (Oracle): *Integration von ETL und OLAP in die relationale DWH-Technologie: mehr Lösung für weniger Aufwand?* (Kurzbeitrag) 568

Industrieprogramm: XML und Tuning

Thomas Tesch, Peter Fankhauser, Tim Weitzel (Infonyte): *Skalierbare Verarbeitung von XML mit Infonyte-DB* 578

Michael Gesmann (Software AG): *Manipulation von XML-Dokumenten in Tamino* 591

Ulrike Schwinn (Oracle): *XML in der Oracle Datenbank "relational and beyond"* (Kurzbeitrag) 611

Eva Kwan, Sam Lighthouse, Berni Schiefer, Adam Storm, Leanne Wu (IBM): *Automatic Database Configuration for DB2 Universal Database: Compressing Years of Performance Expertise into Seconds of Execution* (Kurzbeitrag) 620

Industrieprogramm: Data Warehousing und Indexierung

Michael Hahne (cundus AG): *Logische Datenmodellierung zur Abbildung mehrdimensionaler Datenstrukturen im SAP Business Information Warehouse* 630

Roland Pieringer, Klaus Elhardt, Frank Ramsak, Volker Markl, Robert Fenk, Rudolf Bayer (Transaction Software, IBM, FORWISS): *Transbase: a Leading-edge ROLAP Engine Supporting Multidimensional Indexing and Hierarchy Clustering* 648

Goetz Graefe (Microsoft): *Partitioned B-trees - a user's guide* 668