



---

Mitteilungen der  
GI-Fachgruppe  
**Entwicklungsmethoden für  
Informationssysteme und  
deren Anwendung**

Jahrgang 27 Heft 2  
ISSN 1610-3351

August 2007

---

<b>Editorial</b> .....	1
<b>Aus der Fachgruppe</b> .....	2
<b>Fachbeiträge</b> .....	15
G. Vossen, D. Dahl: Share.loc – Transparenz und Integration von Lernobjekten an der Universität Münster .....	15
G. Thies, G. Vossen: DB2 goes SQL:2003 – Native SQL/XML-Unterstützung der IBM DB2-Datenbank .....	25
<b>Business Process Management Kolumne</b> .....	40
B. Weber, S. Rinderle, M. Reichert: Process Change Patterns .....	45
J. Mendling: On the Detection and Prediction of Errors in EPC Business Process Modells .....	52
<b>Für Sie gesurft</b> .....	60
<b>Berichte und Ankündigungen</b> .....	69

## **EMISA Homepage und EMISA Mailing-Liste**

Ergänzend zum EMISA FORUM gibt es die Möglichkeit, im World Wide Web Informationen mit Relevanz für die Fachgruppe bereitzustellen.

Ansprechpartner ist Dr. Dominik Kuropka [Dominik.Kuropka@hpi.uni-potsdam.de]. Die EMISA Homepage finden Sie unter <http://www.emisa.org>.

Weiter gibt es eine E-Mail-Liste (EMISA MAIL), über die Mitteilungen von allgemeinem Interesse für EMISA-Mitglieder an alle, die sich in die Liste eingetragen haben, verteilt werden können. Insbesondere ist an Ankündigungen und Programme von Tagungen, Workshops und Arbeitsgruppen gedacht, aber auch an (Verweise auf) andere Beiträge, die in das EMISA FORUM passen würden.

E-Mail-Verteiler:	emisa@informatik.uni-hannover.de
An/Abmeldung bitte an:	majordomo@informatik.uni-hannover.de
mit der Rumpfzeile:	[un]subscribe emisa [<E-Mail-Adresse>]
weitere Anfragen an:	emisa-owner@informatik.uni-hannover.de

Ein Archiv mit den letzten Mails findet sich im WWW unter:  
<http://www.informatik.uni-hannover.de/emisa>.

# Editorial

Liebe Fachgruppenmitglieder,

die Mitglieder des EMISA-Leitungsgremiums möchten den Kontakt zu Ihnen, den Fachgruppenmitgliedern, intensivieren. In einem ersten Schritt stellen wir uns Ihnen in dieser Ausgabe des EMISA-FORUM vor. Damit möchten wir Ihnen einen Eindruck von den Personen und ihren Motivationen vermitteln, die Sie im Leitungsgremium vertreten.

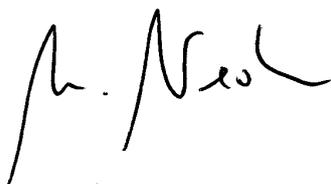
Die konkrete inhaltliche Arbeit von GI-Fachgruppen findet vor allem in Arbeitskreisen statt. Daher freue ich mich, Sie über die Gründung eines neuen EMISA-Arbeitskreises (Arbeitstitel: Effizientes Service-Discovery) informieren zu dürfen. Der Arbeitskreis beschäftigt sich mit pragmatischen Techniken für die Beschreibung von Diensten in serviceorientierten Architekturen, die das Auffinden von Diensten erleichtern. Der Arbeitskreis versammelt Experten aus der Industrie und Wissenschaftler von Hochschulen, um diese aktuelle und praxisnahe Fragestellung zu bearbeiten; in diesem Heft finden Sie die Einladung zum Gründungstreffen des Arbeitskreises.

Um die inhaltliche Arbeit der EMISA zu fördern, möchten wir gern weitere Arbeitskreise aufsetzen; dazu benötigen wir Ihre Mithilfe! Ich bin sicher, dass es eine Vielzahl von Ideen zu neuen Arbeitskreisen gibt; häufig ist es allerdings nicht einfach, eine kritische Masse von etwa zehn Personen zu einem Thema zu versammeln. Um die Gründung neuer Arbeitskreise zu unterstützen, schlage ich vor, dass wir mögliche Arbeitskreisthemen über die EMISA-Webseite zusammentragen. Bitte senden Sie Ihre Vorschläge an den Sprecher oder an Dominik Kuropka als Verantwortlichen für die Webseiten. Wir werden Ihre Vorschläge dann unter <http://emisa.org/neue-aks> veröffentlichen, um weitere Interessenten zu gewinnen. Auf diese Weise wollen wir versuchen, in der nahen Zukunft weitere EMISA-Arbeitskreise aufzusetzen. Wir freuen uns auf Ihre Ideen!

Das bereits angekündigte Fachgruppentreffen im Oktober bietet eine gute Gelegenheit, die Vorschläge zu erörtern. Das Fachgruppentreffen wird in diesem Jahr unter dem Titel *Enterprise Modelling and Information Systems Architectures (EMISA'2007)* gemeinsam mit der GI-Fachgruppe Mobis veranstaltet. Das Fachgruppentreffen ist international ausgerichtet und verspricht sehr interessant zu werden und anwendungsnahe Fragestellungen zu erörtern.

Ich würde mich sehr freuen, wenn ich Sie im Oktober auf der EMISA'2007 in St. Goar am Rhein begrüßen zu dürfen, und ich bin gespannt auf Ihre Ideen zu neuen Arbeitskreisen!

Mit herzlichen Grüßen



Mathias Weske

# Aus der EMISA-Fachgruppe:

- Aufruf zur Teilnahme: *EMISA'07 Fachgruppentreffen + 2nd Workshop on Enterprise Modeling and Information System Architectures*, 8.-9.10.2007, St. Goar
- Vorstellung der Mitglieder des neuen EMISA-Leitungsgremiums
- Ankündigung: Erstes Treffen des *EMISA-Arbeitskreises „Effizientes Service Discovery“*, Hasso-Plattner-Institut, Potsdam, 6. September 2007



## **Call for Participation**

**EMISA 2007**

### **Enterprise Modelling and Information Systems Architectures**

**October 8–9, 2007 — St. Goar / Rhine, Germany**

**<http://www.wi-inf.uni-duisburg-essen.de/emisa07/>**

The strategic importance of modelling is recognised by an increasing number of companies and public agencies. Enterprise modelling delivers the ‘blueprints’ for codesigning organisations and their information systems, so that they complement each other in an optimal way. Achieving this interplay requires a multi-perspective approach that takes into account technical, organisational and economic aspects. It also recommends the cooperation of researchers from different fields such as Information Systems, Business Informatics and Computer Science.

The EMISA’07 workshop is jointly organized by the GI Special Interest Group on Modelling Business Information Systems (GI-SIG MobIS) and the GI Special Interest Group on Design Methods for Information Systems (GI-SIG EMISA). The workshop will address all aspects relevant for enterprise modelling and for the design of information systems architectures. It will provide an international forum to explore new avenues by combining the contributions of different ‘schools’ of Information Systems, Business Informatics, and Computer Science. Therefore, the workshop is open for a broad range of subjects. Topics will include, but will be not limited to:

- enterprise modelling: languages, methods, patterns and tools
- reference models
- modeling services and service compositions in service-oriented architectures
- process modelling in process-aware information systems
- model-driven system development
- model analysis and simulation
- ontologies for enterprise modelling
- model evolution, model life cycle management
- management of model variants and versions
- model quality (e.g. compliance between the system and model level)
- modelling cross-organisational cooperation

## Organisation

### Peter Rittgen

University College of Borås

### Stefan Strecker

University Duisburg-Essen

## Programme Committee

### Manfred Reichert

University of Twente  
Co-Chair

### Klaus Turowski

University of Augsburg  
Co-Chair

Witold Abramowicz, Poland  
Pär Ågerfalk, Ireland  
Antonia Albani, Germany  
Colin Atkinson, Germany  
Peter Bøgh Andersen, Denmark  
Lars Bækgaard, Denmark  
Jörg Becker, Germany  
Martin Bertram, Germany  
Jörg Desel, Germany  
Werner Esswein, Germany  
Marc Favier, France  
Fernand Feltz, Luxembourg  
Ulrich Frank, Germany  
Andreas Gadatsch, Germany  
Claude Godart, France  
Ulrike Greiner, Germany  
Wilhelm Hasselbring, Germany  
Brian Henderson-Sellers, Australia  
Willem-Jan v.d. Heuvel, Netherlands  
Heinrich Jasper, Germany  
Fredrik Karlsson, Sweden  
Dimitris Karagiannis, Austria  
Roland Kaschek, New Zealand  
Ralf Klischewski, Egypt  
John Krogstie, Norway  
Dominik Kuropka, Germany

Susanne Leist, Germany  
Stephen W. Liddle, England  
Mikael Lind, Sweden  
Klaus-Walter Müller, Germany  
Markus Nüttgens, Germany  
Andreas Oberweis, Germany  
Erich Ortner, Germany  
Sven Overhage, Germany  
Hansjürgen Paul, Germany  
Erik Proper, Netherlands  
Michael Rebstock, Germany  
Stefanie Rinderle, Germany  
Peter Rittgen, Sweden  
Michael Rosemann, Australia  
Matti Rossi, Finland  
Gunter Saake, Germany  
Guttorm Sindre, Norway  
Elmar J. Sinz, Germany  
Stefan Strecker, Germany  
Juha-Pekka Tolvanen, Germany  
Gottfried Vossen, Germany  
Barbara Weber, Germany  
Hans Weigand, Germany  
Mathias Weske, Germany  
Roel Wieringa, The Netherlands  
Robert Winter, Switzerland

## Location

St. Goar (Map) is situated in the heart of the Middle Rhine region, the most picturesque part of the Rhine valley between Rüdesheim and Koblenz. Both banks are lined with fairy-tale castles and steep vineyards. The venue of our workshop, Burg Rheinfels, towers over St. Goar. It is the mightiest fortress that was ever constructed on the banks of the Rhine. The original castle was built in 1245 by Graf Dieter V von Katzenelnbogen, and its well-preserved ruins are still an impressive sight, offering spectacular views of the valley.



St. Goar is located between Frankfurt and Cologne with excellent connections by flight, train and car, e.g. to [Frankfurt Airport](#) (approx. 90 minutes by car and by train), to [Frankfurt Hahn Airport](#) (approx. 60 minutes by car), and to Cologne (approx. 2 hrs by car and by train).

The workshop will take place in the [Castle Rheinfels Hotel](#) set on a hilltop above the town of St. Goar. Recently given the title of the Best Event Hotel 2005, the 4 star hotel has been ranking among the best conference hotels in Germany for ten years. Hotel bookings will be available directly through the [conference hotel](#). Further links will be provided when [registration](#) starts in August 2007.



**Das wissenschaftliche Programm der EMISA'07 lag zum Zeitpunkt des Drucks dieses Heftes noch nicht vor. Bei mehr als 40 eingereichten Papieren ist aber mit einem qualitativ sehr guten und attraktiven Workshop zu rechnen. Das genaue Programm sowie detaillierte Informationen zur Registrierung finden sich unter**

**<http://www.wi-inf.uni-duisburg-essen.de/emisa07/>**

# Vorstellung des neuen EMISA-Leitungsgremiums

## Prof. Dr. Mathias Weske (Sprecher EMISA)



„Mit steigender Komplexität von Informationssystemen kommt der Modellierung dieser Systeme eine immer größere Bedeutung zu. Aus meiner Sicht nimmt die EMISA in diesem Kontext eine zentrale Rolle ein, weil sie versucht, eine breite, ganzheitliche Betrachtungsweise beim Entwurf von Informationssystemen einzunehmen, die unterschiedliche Methoden und Ansätze einschließt und miteinander in Beziehung setzt.“

Mathias Weske leitet das Fachgebiet Business Process Technology am Hasso Plattner Institut an der Universität Potsdam. Das Fachgebiet beschäftigt sich mit theoretischen und praktischen Fragestellung prozessorientierter und dienstbasierter Informations-systeme. Er ist seit Januar 2007 Sprecher der EMISA, nachdem er von 2001-2006 Mitglied des Leitungsgremiums und stellvertretender Sprecher der EMISA war.

<http://bpt.hpi.uni-potsdam.de/>

---

## Prof. Dr. Manfred Reichert (Stellv. Sprecher EMISA)



„Zukünftige Informationssysteme müssen sich immer rascher und kostengünstiger an Veränderungen anpassen lassen, um der zunehmenden Forderung nach mehr Agilität gerecht zu werden. Der Modellierung, der Entwicklung und dem Betrieb solcher agilen Informationssysteme kommt deshalb eine wichtige Bedeutung zu. Die Rolle der EMISA wird es sein, ganzheitliche Methoden und Ansätze hierfür zu betrachten, in Beziehung zueinander zu setzen und hinsichtlich ihrer praktischen Anwendbarkeit zu bewerten.“

Manfred Reichert ist seit 2005 als Associate Professor am Lehrstuhl für Informationssysteme (Computer Science Department) an der Universität Twente, Niederlande tätig. Hier koordiniert er auch den Forschungsschwerpunkt "Service-orientierte Architekturen und E-Health" und ist Mitglied des Leitungsgremiums des Instituts für Telematik und Informationstechnologie. Fachliche Schwerpunkte seiner aktuellen Arbeiten sind Business Process Management, Service-orientierte Architekturen und E-Health. Seit Januar 2007 ist er stellvertretender Sprecher der EMISA und Redaktionsleiter des EMISA Forums.

<http://wwwhome.cs.utwente.nl/~reichertm/index01.htm>

## **Prof. Dr. Jörg Desel**

„Als Sprecher der befreundeten GI-Fachgruppe, die sich primär mit Petrinetzen befasst, begleite ich die Arbeit der EMISA sehr gern, weil Prozessmodellierung bei dem Entwurf von Informationssystemen eine zentrale Rolle einnimmt.“



Jörg Desel hat einen Lehrstuhl für Angewandte Informatik an der Katholischen Universität Eichstätt-Ingolstadt. Zu seinen Forschungsgebieten gehört die Modellierung und Analyse von Informationssystemen und Geschäftsprozessen, insbesondere mit Petrinetzen und verwandten Modellierungssprachen.

<http://www.informatik.ku-eichstaett.de/mitarbeiter/desel/index.html>

---

## **Fernand Feltz**

„Die Praxis zeigt immer wieder die Schwierigkeit, neue Konzepte und Vorgehensweisen der Informatik so einzusetzen, dass auch die erhofften Vorteile zur Geltung kommen. Die EMISA ist eine ideale Plattform, in dieser Hinsicht weiterzukommen. Insbesondere die serviceorientierten Ansätze werden uns in den nächsten Jahren beschäftigen. Je intensiver diese Ansätze in der realen Welt der Unternehmen und der öffentlichen Verwaltungen umgesetzt werden, je wichtiger ist es zusätzlich zu der analytischen auch eine systemische Sichtweise mit sozio-ökonomischen Aspekten heranzuziehen. Auch eine Öffnung der EMISA hin zu anderen europäischen "Communities" induziert interessante Diskussionen, wie sie mit der Organisation der EMISA 2004 und der AIM 2006 in Luxemburg, wo vor allem germanophone und frankophone "Communities" zusammentrafen, erreicht wurden.“



Fernand Feltz leitet als wissenschaftlicher Direktor das "Research Department - Informatics, Systems and Collaboration (ISC)" an dem öffentlich geförderten Forschungszentrum "Centre de Recherche Public - Gabriel Lippmann" in Luxemburg. Er ist seit Januar 2004 Mitglied des EMISA-Leitungsgremiums und realisiert mit seinen Teams praxisnahe Forschungsprojekte mit Partnern aus dem privaten und öffentlichen Sektor in Luxemburg sowie mit anderen europäischen Forschungszentren und Universitäten.

<http://www.lippmann.lu>

## **Prof. Dr. Ulrich Frank**



„Als Sprecher der Fachgruppe MobIS (Modellierung betrieblicher Informationssysteme) ist mir sehr an einer engen Kooperation mit der EMISA gelegen. Die bisherige Zusammenarbeit verlief sehr erfreulich. Sie umfasst die Publikation der Zeitschrift "Enterprise Modelling and Information Systems Architectures" sowie die gemeinsame Durchführung einer internationalen Tagungsreihe gleichen Namens. Ich werde mich dafür einsetzen, dass beide Fachgruppen auch in Zukunft einen regen Austausch pflegen.“

Ulrich Frank ist seit 1995 Professor für Wirtschaftsinformatik, insbesondere Unternehmensmodellierung. Nachdem er zunächst die Forschungsgruppe Unternehmensmodellierung im Fachbereich Informatik der Universität Koblenz-Landau leitete, wechselte er 2004 an die Universität Duisburg-Essen ins Institut für Informatik und Wirtschaftsinformatik (ICB). Zu seinen Forschungsinteressen gehören die Spezifikation und Evaluation von Modellierungssprachen, die Entwicklung von Referenzmodellen, das Method Engineering sowie die Auseinandersetzung mit Forschungsmethoden der Wirtschaftsinformatik.

<http://www.wi-inf.uni-duisburg-essen.de/FGFrank>

---

## **Prof. Dr. Andreas Gadatsch**



„Die EMISA bietet ein weites Dach für vielfältige Aktivitäten und führt Forscher und wissenschaftlich interessierte Praktiker zusammen. Mein persönliches Interesse gilt hier insbesondere den Modellierungsmethoden, und die Frage der Tauglichkeit in realen Projekten.“

Andreas Gadatsch ist Inhaber der Professur für Betriebswirtschaftslehre, insb. Wirtschaftsinformatik am Fachbereich Wirtschaft der Fachhochschule Bonn-Rhein-Sieg in Sankt Augustin. Er hat langjährige Erfahrung als Berater, Projektleiter und IT-Manager. Seine Hauptarbeitsgebiete sind IT-Controlling, Geschäftsprozessmanagement (insb. Geschäftsprozessmodellierung) und betriebliche Informationssysteme. Er ist seit 2007 Mitglied im Leitungsgremium.

<http://www.wis.fh-brs.de/gadatsch/>

## Horst Kremers



„Die für Komplexanwendungen erforderlichen interdisziplinären Einbindungen von heterogenen, akteur- und prozessmodellierten Informationssystemen finde ich bei EMISA in praxisrelevantem Erfahrungsaustausch. Es war daher für mich sehr interessant, dass wir die Kooperation des deutschen Nationalausschusses für CODATA (Committee on Data for Science and Technology des International Council for Science) <http://www.codata-germany.org> mit EMISA etablieren konnten. Unsere diesjährigen Symposien geben einiges dieser gemeinsamen Fachinteressen wieder.“

Horst Kremers befasst sich mit Methodik und Technik der Informationssysteme schwerpunktmäßig in den Bereichen Geoinformation, Digitale Kartographie, Umweltinformation. Nach semantischen Entwicklungen (Metainformationssysteme seit den Anfängen ca. 1988) und den Erfahrungen aus dem Kommunalen Umweltbereich hat meine Leitung des AK Umweltdatenbanken im GI FA 4.6 über mehrere Jahre mich dann verstärkt auf die Erfordernisse akteur- und damit prozessorientierter Darstellungen / Modelle gelenkt.

<http://www.horst-kremers.de/>

---

## Dr. Dominik Kuropka



„Integration von verschiedenen Forschungsgebieten ist starker Katalysator für innovative Forschungsergebnisse. Die EMISA-Fachgruppe ist eine Plattform die eine derartige Integration in dem weiten Umfeld von Informationssystemen bietet und fördert.“

Dominik Kuropka arbeitet als wissenschaftlicher Assistent im Fachgebiet Business Process Technology am Hasso Plattner Institut an der Universität Potsdam. Er beschäftigt sich mit der semantischen Beschreibung und Modellierung von Anwendungsdomänen und insbesondere Services als auch mit ontologie-basiertem Information-Filtering und Retrieval. Er ist seit 2007 Mitglied im EMISA-Leitungsgremium.

<http://www.kuropka.net/>

## **Prof. Dr. Bernd Müller**



„Die Entwicklung betrieblicher Informationssysteme ist sowohl durch eine sich steigernde, technische Komplexität als auch durch zunehmend fachlich anspruchsvollere Anforderungen geprägt. Die EMISA kann dazu beitragen, zwischen universitären Modellen, technologischen Rahmenbedingungen und praktischen Anforderungen zu vermitteln und die Qualität der Ergebnisse der Modellierung von Informationssystemen zu verbessern.“

Bernd Müller ist seit 2005 Professor für Software-Technik an der Fachhochschule Braunschweig/Wolfenbüttel. Zuvor war er seit 1998 Professor für Wirtschaftsinformatik an der Hochschule Harz in Wernigerode. Seine aktuellen fachlichen Interessen liegen im Bereich der transaktionalen Informationssysteme und deren technische Umsetzung.

[www.pdbm.de/](http://www.pdbm.de/)

---

## **Prof. Dr. Markus Nüttgens**



„Informationssysteme werden im Regelfall interdisziplinär entwickelt und genutzt. Die Fachgruppe EMISA kann hier auf eine lange und breite Erfahrung aufsetzen und verbindet traditionell Vertreter aus der Angewandten Informatik, der Wirtschaftsinformatik, dem Ingenieurwesen und den Sozialwissenschaften. Diese Mischung aus Fachgebieten und Vertretern der Wissenschaft und Praxis fordert und fördert neue Ideen und Konzepte.“

Markus Nüttgens ist Inhaber des Lehrstuhls für Wirtschaftsinformatik an der Universität Hamburg. Seine Forschungsschwerpunkte sind Informations- und Geschäftsprozess-management, Betriebswirtschaftliche Standardsoftware, Informations-systeme in Industrie, Dienstleistung und Verwaltung, Open Source/Access, IT-Governance/-Outsourcing und IT-Entrepreneurship. Er ist Sprecher des GI-Arbeitskreises "Geschäftsprozessmanagement mit Ereignisgesteuerten Prozessketten (WI-EPK)" und Mitglied des Leitungsgremiums der Fachgruppe „Entwicklungsmethoden für Informations-systeme und deren Anwendung (EMISA)" und des Leitungsgremiums des Fachbereichs "Wirtschaftsinformatik (FB-WI)"

<http://www.wiso.uni-hamburg.de/>

## **Prof. Dr. Andreas Oberweis**

„Die EMISA stellt von ihrer inhaltlichen Ausrichtung her die Verbindung zwischen Vorgehensmodellen, Methoden und Werkzeugen von Information Engineering, Prozess Engineering und Software Engineering her. Gerade in Zeiten immer weiter reichender Spezialisierung in der Informatik ist es wichtig, dass auch die Beziehungen zwischen den Teilgebieten gesehen werden und dass interdisziplinäre Ansätze verfolgt werden, um suboptimale Partillösungen bei der Gestaltung soziotechnischer Systeme zu vermeiden. Dies soll auch künftig zentrale Aufgabe der EMISA bleiben.“



Andreas Oberweis ist Professor für Angewandte Informatik an der Universität Karlsruhe (TH). Daneben ist er Direktor im Forschungszentrum Informatik (FZI) Karlsruhe. Seine Forschungs- und Lehrgebiete umfassen Information Engineering, Prozess Engineering und Software Engineering. Seit 22 Jahren Mitglied in der EMISA.

[www.aifb.uni-karlsruhe.de](http://www.aifb.uni-karlsruhe.de)

---

## **Dr.-Ing. Hansjürgen Paul**

„Wissenschaftliches Arbeiten meint nicht nur die Entwicklung von neuen Sichtweisen, Theorien und Methoden, sondern auch deren Anwendung und damit Evaluation und Evolution in der Auseinandersetzung mit der Praxis. EMISA steht für dieses Prinzip in der Wissensdomäne der Informationssysteme - im engeren wie auch weiteren Sinn. Ich möchte mich dafür einsetzen, dass dieser Bezug zur Praxis nicht aus den Augen verloren wird und sich als roter Faden auch durch die Aktivitäten der EMISA in den kommenden Jahren zieht.“



Hansjürgen Paul ist seit 1989 als wissenschaftlicher Mitarbeiter am Institut Arbeit und Technik (IAT) in Gelsenkirchen tätig. Bis 2003 arbeitete er in der Abteilung Produktionssysteme, dann bis Ende 2006 für die Abteilung Bildung und Erziehung im Strukturwandel (BEST). Mit Übernahme des IAT durch die Fachhochschule Gelsenkirchen 2007 im Rahmen der Neuordnung unabhängiger wissenschaftlicher Institute in NRW ist er nun im Forschungsschwerpunkt Wissen und Kompetenz (WISDOM) des IAT tätig. Er hat einen Lehrauftrag am Fachbereich Informatik der Fachhochschule Gelsenkirchen und vertritt dort das Fach Informatik und Gesellschaft. Darüber hinaus unterrichtet er in den Fachgebieten Interaktive Systeme und Software-Technik.

<http://www.iat.eu/personal/paul.html>

## **Dr. Stefanie Rinderle-Ma**

„Neben dem klassischen Aspekt der Modellierung von Informationssystemen stellt die EMISA auch die geeignete Plattform für die Anregung und Diskussion neuer interessanter Fragestellungen dar. Dies bildet die ideale Grundlage für innovative Forschung in diesem Bereich.“



Stefanie Rinderle-Ma arbeitet am Institut für Datenbanken und Informationssysteme an der Universität Ulm. Ihre Forschungsinteressen liegen im Bereich von Prozessmodellierung und -analyse, flexiblen Prozess-Management-Systemen, Integration semantischer Aspekte in Prozess-Management und der Evolution organisatorischer Strukturen. Sie ist seit Oktober 2006 Mitglied des Leitungsgremiums der EMISA.

<http://www.informatik.uni-ulm.de/dbis/01/staff/rinderle/rinderle-t.htm>

---

## **Dr.-Ing. Eike Schallehn**

„Der GI Arbeitskreis Grundlagen von Informationssystemen und sein Leitungsgremium, dessen Sprecher ich bin, sieht sein Arbeitsfeld in den zahlreichen Fragestellungen, die sich aus dem Einsatz von Datenbanken in Informationssystemen ergeben. Deshalb ist eine enge Kooperation mit den Fachgruppen EMISA und Datenbanken für uns von besonderer Bedeutung.“



Eike Schallehn arbeitet als wissenschaftlicher Assistent in der Arbeitsgruppe Datenbanken und Informationssysteme der Otto-von-Guericke-Universität Magdeburg. Seine aktuellen Forschungsinteressen liegen in den Bereichen Self-Tuning von Datenbanken sowie Integration von Datenbanken und Informationssystemen. Er ist seit Juni 2007 Mitglied des Leitungsgremiums.

<http://www.witi.cs.uni-magdeburg.de/~eike>

## **Prof. Dr. Gottfried Vossen**



„Die EMISA hat sich seit jeher als Fachgruppe für Informationssysteme, deren Entwurf und deren Anwendungen im engeren sowie im weiteren Sinne verstanden. Diese Sicht gibt ihr die Möglichkeit, umfassend auf Konzeptions- sowie Realisierungs- und Anwendungsfragen einzugehen und diese auch im Lichte aktueller Technologiediskussionen zu bewerten. Dies sollte auch in Zukunft integraler Bestandteil der Fachgruppenarbeit sein.“

Gottfried Vossen ist seit 1993 Professor für Informatik, insbesondere Datenbanken und Informationssysteme am Institut für Wirtschaftsinformatik der Universität Münster. Seine aktuellen fachlichen Interessen finden sich im elektronischen Lernen, in transaktionalen Informationssystemen, Service-orientierten Architekturen, Web Services sowie in Anwendungen von Web 2.0-Technologien.

<http://dbms.uni-muenster.de>

---

# Ankündigung:

## Erstes Treffen des EMISA-Arbeitskreises Effizientes Service Discovery

### Zeit und Ort der Veranstaltung

Donnerstag, den 6. September 2007 von 10:00 bis 17:00 Uhr.

Hasso Plattner Institut  
Prof.-Dr.-Helmert-Str. 2 – 3  
14482 Potsdam

Raum B-2.1

### Hintergrund

Mit der zunehmenden Umstellung von Unternehmens-IT auf Service-orientierte Architekturen ist langfristig mit einer ansteigenden Anzahl von Funktionalitäten bzw. Dienstleistungen zu rechnen, die über Web-Services (oder über andere Protokolle) angeboten werden. Schon im Kontext größerer Unternehmen kann die Anzahl der vorhandenen und bei Bedarf wiederverwendbaren Funktionalitäten ein unüberschaubares Maß erreichen. Um eine effiziente Wiederverwendung solcher Funktionalitäten zu ermöglichen, kommt einer geeigneten Infrastruktur zentrale Bedeutung zu. Diese muss zum eine das Einstellen von über Services verfügbaren Funktionalitäten unterstützen, zum anderen eine benutzerfreundliche und zielgerichtete Suche (Service Discovery) ermöglichen.

Derzeit werden in Industrie und Wissenschaft zwei grundsätzliche Ansätze für das Service Discovery verfolgt. Der erste Ansatz basiert auf dem Standard UDDI (Universal Description, Discovery and Integration), der die Funktionalität eines Verzeichnisdienstes für Web-Services festlegt. Bei diesem Ansatz können Services anhand von verschiedenen Attributen klassifiziert und auch gesucht werden. Dieser eher technisch orientierte Ansatz hat sich aus Sicht der Anwender bisher als nur bedingt tauglich erwiesen. In der Tat sind die drei großen, von IBM, Microsoft und SAP betriebenen, öffentlichen UDDI-Verzeichnisse vor kurzem geschlossen worden.

Der zweite – zur Zeit vor allem in der Forschung diskutierte – Ansatz basiert auf einer formal-logischen Beschreibung der semantischen Funktionalität von Services. Mit Hilfe von logischen Sprachen (Description Logic, F-Logic, OWL, WSMML) werden präzise die möglichen bzw. erlaubten Eingaben, Vorbedingungen, Ergebnisse und Effekte eines

Dienstaufrufs beschrieben. Somit kann sehr genau nach Diensten mit einer bestimmten Funktionalität gesucht werden. Während dieser Ansatz aus theoretischer Sicht deutliche Vorteile aufweist, ist seine Umsetzung in der Praxis mit zum Teil erheblichen Schwierigkeiten verbunden. Eine semantische Charakterisierung der Funktionalität von Diensten ist nicht nur deshalb mit Schwierigkeiten verbunden, weil logische Beschreibungen meist wenig intuitiv und sehr explizit sind. Vor allem besteht die Notwendigkeit, sich vorab mit allen Beteiligten auf eine genaue und formale Ontologie zu einigen. Dies stellt angesichts heterogener Benutzergruppen mit variablen und veränderlichen Anforderungen zumindest ein immenses praktisches Problem dar. Die beschriebenen Schwierigkeiten betreffen sowohl die Anbieter von Diensten, als auch diejenigen, die nach Diensten suchen. Sie behindern die tatsächliche Umsetzung dieses theoretisch fundierten Ansatzes.

Das Ziel dieses Arbeitskreises ist es, einen dritten Weg für effizientes Service Discovery zu finden. Im Gegensatz zu den anderen beiden Ansätzen sollen die Bedürfnisse der beiden Anwendergruppen – Dienstanbieter und Dienstsucher – von vornherein bei der Konzeption berücksichtigt werden. Bei der Auswahl der Gründungsmitglieder wurde darauf geachtet, dass hinreichend viele Vertreter aus der Industrie teilnehmen, um die Praktikabilität der Ergebnisse zu sichern. Die wissenschaftlichen Partner repräsentieren einen interdisziplinären Hintergrund in den Bereichen Service Infrastructure, IT-Systeme und Unternehmensmodellierung sowie Information Retrieval und Computerlinguistik. Diese breit gefächerte Kombination von Expertise bietet das Potential, neue und benutzerfreundlichere Wege im Bereich des Service Discovery zu beschreiben.

### Kontakt

Dominik Kuropka

Hasso Plattner Institut  
Prof.-Dr.-Helmert-Str. 2 – 3  
14482 Potsdam

E-Mail: [dominik.kuropka@hpi.uni-potsdam.de](mailto:dominik.kuropka@hpi.uni-potsdam.de)

Telefon: 0331 / 5509 – 193

# share.loc – Transparenz und Integration von Lernobjekten an der Universität Münster

Gottfried Vossen<sup>1</sup>, Daniel Dahl  
European Research Center for Information Systems (ERCIS)  
Westfälische Wilhelms-Universität Münster  
Leonardo-Campus 3  
D-48149 Münster  
{vossen, dahl}@wi.uni-muenster.de

März 2007

## Abstract

*Mit der zunehmenden Verbreitung von e-Learning an Hochschulen ist einhergehend eine wachsende Intransparenz bezüglich vorhandener Inhalte und Plattformen festzustellen. Das an der Universität Münster entwickelte e-Learning Repository share.loc zeigt einen Ausweg aus diesem Dilemma auf, indem an zentraler Stelle ein Überblick über die verteilten digitalen Lernressourcen bereitgestellt wird. Durch den Ansatz, Schnittstellen zu unterschiedlichen Schichten der Systemarchitektur bereitzustellen, wird ein hohes Integrationspotential freigesetzt. Externe Systeme wie komplexe Lernplattformen oder Universitätsportale können Funktionalitäten wie die Suche nach Lernobjekten, basierend auf individuellen, jedoch zu jederzeit standardisierten Metadatenrepräsentationen, in den eigenen Kontext integrieren. Auf diese Weise wird der Heterogenität der Systemlandschaft einer Universität wie auch der Heterogenität der potentiellen Systemnutzer in gleichem Maße Rechnung getragen.*

## 1. Einleitung

Das Konzept des Blended Learning, eine Mischform von Präsenzlehre und elektronischem Lernen, ist längst nicht mehr auf Computer-affine Bereiche der Hochschullehre beschränkt. Eine Umfrage, durchgeführt im Herbst 2005 an der Universität Münster, hat eindeutig aufgezeigt, dass über nahezu alle Fachbereiche verteilt e-Learning genutzt wird bzw. die Absicht besteht, e-Learning einzusetzen [1]. Plattformen wie Freestyle Learning [2], OpenUSS [3], Herbie [4] oder xLx [5] unterstützen Dozenten wie Lernende bei unterschiedlichsten Prozessen, angefangen bei der Organisation von Vorlesungen und Seminaren, über die Distribution von Lernmaterial und Selbstlerneinheiten bis hin zur Abnahme von hoch-automatisiert auswertbaren Testaufgaben. In Folge dessen entstanden und entstehen ständig neue Inhalte, die für Lernende wie auch für Lehrende von großem Interesse sind. Beispiele sind die bereits genannten Selbstlerneinheiten, aber auch Videomitschnitte von Vorlesungen oder auf Benutzerinteraktion zugeschnittene Applets. Mit der zunehmenden Anzahl von Content-Erstellern, die verteilt über die gesamte Universität anzutreffen sind, wächst in gleichem Maße die Intransparenz der e-Learning Bestrebungen. e-Learning Inhalte liegen verteilt auf verschiedensten Rechnern, teilweise versteckt hinter Passwörtern oder in monolithischen Systemen. Sucht ein e-Learning Interessierter einen bestimmten Inhalt, so muss er einen Rechner bzw. ein System nach dem anderen ansteuern. Ein Erfolg ist auf diese Weise natürlich nicht garantiert. Das Dilemma ist umso größer, wenn die Strukturen der Universität dem Suchenden nicht im Detail bekannt sind. So

---

<sup>1</sup> Zurzeit an der University of Waikato Management School, Hamilton, New Zealand.

wächst auf der einen Seite die Menge der Lernobjekte ständig an, die Auffindbarkeit und die anschließende Nutzung werden jedoch zunehmend schwieriger.

Zur Lösung der beschriebenen Problematik sind verschiedene Ansätze denkbar. Ein Server könnte z. B. an einer zentralen Stelle alle e-Learning Inhalte der gesamten Universität verwalten. Zu diskutierende Fragen betreffen in diesem Fall insbesondere rechtliche Aspekte, den hohen Speicherbedarf sowie die Ausfallsicherheit. Alternativ könnte die IT-Zentrale (das frühere „Rechenzentrum“) eine Server-Infrastruktur bereitstellen und betreiben, welche diese Probleme abfedert. Eine weitere Alternative besteht in der Bereitstellung einer Art „Gelben Seiten“ für Lernobjekte: An einer zentralen Stelle können e-Learning Interessierte nach Inhalten suchen, die Objekte selbst werden jedoch extern verwaltet (also weiterhin in den einzelnen Fachbereichen). Verweise werden über einfache HTML-Links realisiert. Im Sinne der Lernobjekte nach [6] werden in dem zentralen System die Metadaten (Daten über Daten) der Inhalte gesammelt. Gemäß dem IEEE Metadatenstandard LOM (Learning Objects Metadata) [7] ermöglichen Attribute wie Titel, Beschreibung, Rechte oder Format eine effektive Suche. Das e-Learning Repository share.loc (shared repository for learning object content) [8] realisiert diesen auf Lernobjekt-Metadaten begründeten Ansatz, weshalb von einem e-Learning Metadata Repository gesprochen wird. Es werden lediglich Metadaten über Lernobjekte in einer zentralen Webanwendung (anytime/anywhere-Zugriff) gesammelt und für eine Suche zur Verfügung gestellt. Die Lernobjekte selbst bleiben unberührt und werden an dezentraler Stelle verwaltet. Diskussionen über Fragen bezüglich des Speicherplatzes oder rechtliche Aspekte erübrigen sich an dieser Stelle folglich.

Im Folgenden wird das e-Learning Metadata Repository share.loc vorgestellt, wobei einige innovative Konzepte und Ansätze im Detail betrachtet werden sollen. Zunächst wird in Abschnitt 2 erläutert, wie eine Transparenz von verteilten Lernobjekten erzielt wird. Die Grundfunktionen wie die Suche über Schlüsselwörter, die Navigation in einem an die Struktur der Universität Münster angelehnten Verzeichnis oder das Browsen durch eine aus Schlagwörtern bestehende Tag-Wolke werden beschrieben. Abschnitt 3 zeigt das Integrationskonzept auf, indem die eine vielschichtige Interoperabilität garantierenden Schnittstellen herausgestellt werden. Diese ermöglichen eine problemlose Integration von share.loc-Funktionalitäten in unterschiedlichste Systeme, wie z. B. komplexe Lernplattformen oder über das e-Learning hinausgehende Suchmaschinen. Abschnitt 4 fasst diese Arbeit zusammen und gibt einen Ausblick auf zukünftige Forschungsfragen im Kontext des e-Learning Repositories.

## **2. Anforderungen an ein universitätsweites e-Learning Repository**

Bezug nehmend auf die im Jahr 2005 an der Universität Münster durchgeführte Umfrage konnten individuelle Anforderungen an ein e-Learning Repository, das fächerübergreifend eingesetzt werden soll, abgeleitet werden. Unter dem Aspekt der Heterogenität der potentiellen Nutzer sollen in besonderem Maße Personen unterstützt werden, die bisher wenig bis keinen Kontakt mit elektronischem Lernen aufweisen können. Entsprechend wurde im gesamten Konzeptions- und Designprozess eine als minimalistisch zu bezeichnende Strategie verfolgt. Dies bedeutet im einzelnen, dass das Rollenkonzept auf nicht mehr als zwei Gruppen basiert: Personen, die Metadaten suchen und ansehen können (Gäste, z. B. Dozenten und Studenten) und Personen, die Metadaten und Verzeichnisstrukturen anlegen und ändern können (Autoren, in erster Linie Dozenten). Autoren müssen sich dafür nicht als Benutzer im System registrieren, sondern können den für jeden Mitarbeiter der Universität Münster vorhandenen ZIV-Account (Zentrum für Informationsverarbeitung) nutzen. Die Eintrittsbarriere ist an dieser Stelle also auf ein Minimum reduziert. Im Sinne eines barrierefreien Designs basieren alle im Folgenden beschriebenen Kernfunktionalitäten auf einfachem HTML, das von jedem Webbrowser interpretiert werden kann. Techniken wie Javascript oder Flash werden bewusst ausgeschlossen. Im Umgang mit dem Metadatenstandard LOM wurde darauf verzichtet, dem Endanwender den vollständigen komplexen Metadatenstandard zur Verfügung zu stellen. Von den über 60 Attributen wurde in Anlehnung an [9] eine Kernmenge an relevanten Attributen identifiziert, die auf der einen Seite erfahrenen e-Learning Anbietern die Möglichkeit gibt, Inhalte umfassend zu beschreiben, und auf der anderen Seite e-Learning Interessierte bei der Suche nicht überfordert. Für eine ausführliche Beschreibung aller abgeleiteten Anforderungen sei auf das Pflichtenheft des e-Learning Repositories [10] verwiesen. Die unter den genannten Kriterien realisierten Grundfunktionen werden im Folgenden beschrieben.

### 3. Grundfunktionen des Repositories share.loc

Im Allgemeinen zeigen e-Learning Repositories dem Benutzer zwei Sichten auf die verfügbaren Lernobjekte: eine mehr oder weniger komplexe Suche und die Möglichkeit der Navigation durch ein kontextabhängiges Verzeichnis [11]. Als Beispiele seien hier Merlot [12] und LernAlberta [13] genannt. Beide Repositories bieten dem Benutzer eine einfache und eine erweiterte Suche, die basierend auf Schlüsselwörtern die Metadaten der Lernobjekte durchsucht. Während das hierarchische Verzeichnis in Merlot auf speziellen Disziplinen begründet ist („Computer Science“ als Unterkategorie von „Science and Technology“), wird in LernAlberta zielgruppenspezifisch kategorisiert (Kategorie „Kindergarten“ neben Kategorie „Grade 12“). Das Repository share.loc bietet neben der schlüsselwortbasierten Suche und einem an die Universitätsstrukturen angelehnten Verzeichnis eine zusätzliche innovative Sicht auf die indizierten Lernobjekte: die in Zusammenhang mit so genannten Web 2.0-Anwendungen immer häufiger anzutreffende Tag-Wolke.

#### 3.1 Schlüsselwortbasierte Suche

Die einfache Form der schlüsselwortbasierten Suche erwartet als Eingabe einen oder mehrere Suchbegriffe (Abbildung 1). Durch die Auswahl einer Verknüpfungsform wird festgelegt, wie die einzelnen Begriffe zu behandeln sind. Die Oder-Verknüpfung erlaubt, dass die Begriffe in unterschiedlichen Lernobjekten vorkommen können. Eine Und-Verknüpfung stellt die Bedingung, dass alle genannten Wörter in ein und demselben Lernobjekt vorhanden sind. In einer einfachen Suche werden alle Metadatenattribute der Lernobjekte durchsucht; es ist also unerheblich, ob ein Suchbegriff im Metadatenfeld Titel, Autor oder der Beschreibung der Rechte vorkommt. Um eine flexible Suche zu ermöglichen ist die Eingabe einfacher Patterns sinnvoll: ein ‚?‘ entspricht einem beliebigen Zeichen, ein ‚\*‘ entspricht einer Menge von selbigen. So werden nach Eingabe des Suchbegriffes ‚v\*n gr\*b‘ u.a. Lernobjekte des Autors Vossen gefunden. Eine Suche nach ‚gr\*b‘ liefert Lernobjekte des Autors Grob.

Abbildung 1: Einfache Suche

Abbildung 2: Erweiterte Suche

Die erweiterte Suche (Abbildung 2) bietet die Möglichkeit einer zielgerichteten, schlüsselwortbasierten Suche. Durch die Verknüpfung von Suchbegriffen und Metadatenfeldern kann explizit ausgedrückt werden, in welchem Metadatenfeld ein Schlüsselwort vorkommen muss, damit es dem Suchkriterium genügt. So lässt sich gezielt nach Lernobjekten suchen, die vom Autor Vossen erstellt wurden. Nicht berücksichtigt werden Lernobjekte, die z. B. in ihrer Beschreibung lediglich den Namen erwähnen.



Abbildung 3: Ergebnisanzeige



Abbildung 4: Metadatenanzeige

Sofern eine Suche erfolgreich war, werden die gefunden Lernobjekte in einer Liste angezeigt (Abbildung 3). Durch die Auswahl eines Lernobjektes gelangt der Nutzer zu einer an eine Karteikarte angelehnten Darstellung des gesamten Metadatenatzes (Abbildung 4). Hervorzuheben ist das Metadatenfeld Ablage, das in Form eines HTML-Links implementiert ist. Der Benutzer gelangt über die Auswahl des Verweises direkt zum entsprechenden Lernobjekt.

Eine schlüsselwortbasierte Suche ist in erster Linie dann sinnvoll, wenn eine gezielte Suche anhand von bestimmten Parametern durchgeführt werden soll. Ist z. B. der Autor eines Lernobjektes bekannt, so kann eine effektive Anfrage gestellt werden. Kann jedoch lediglich vermutet werden, an welcher Fakultät oder an welchem Lehrstuhl ein nicht näher zu beschreibendes Lernobjekt veröffentlicht wurde, so sollte eine alternative Form der Suche gewählt werden.

### 3.2. Navigation im Verzeichnis

Das Verzeichnis in share.loc ist an den Strukturen der Universität Münster ausgerichtet. Dies bedeutet, unterhalb des Wurzelknotens „Universität Münster“ sind die einzelnen Fachbereiche anzutreffen (Abbildung 5), auf welche die einzelnen Institute folgen. Den Lehrstuhl für Informatik erreicht man beispielsweise über die „Wirtschaftswissenschaftliche Fakultät“ und das Institut für „Wirtschaftsinformatik“. Auf jeder Ebene der Struktur können Lernobjekte vorhanden sein. Abbildung 6 zeigt den Inhalt des Ordners „e-Learning Plattformen“: alle an diesem Lehrstuhl entwickelten e-Learning Systeme. Nach der Auswahl eines Datensatzes wird dem Benutzer die aus Abbildung 4 bekannte Karteikarte angezeigt.

Auf diese Weise kann, ergänzend zur schlüsselwortbasierten Suche, das e-Learning Verzeichnis der Universität Münster explorativ erkundet werden. So stößt der Benutzer auf Lernobjekte, die sich durch eine zielgerichtete Suche über Schlüsselwörter nicht erschließen würden und einzelne Fachbereiche können intuitiv angesteuert werden.

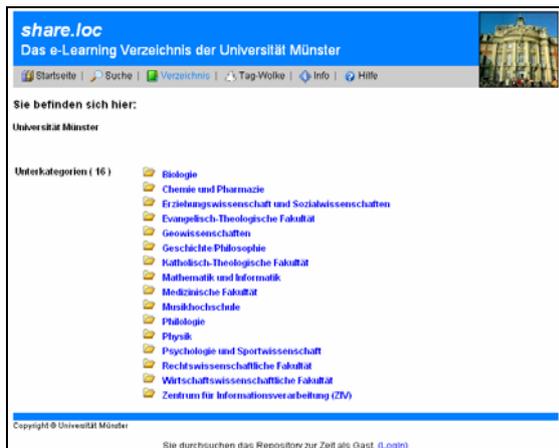


Abbildung 5: Startpunkt Verzeichnis

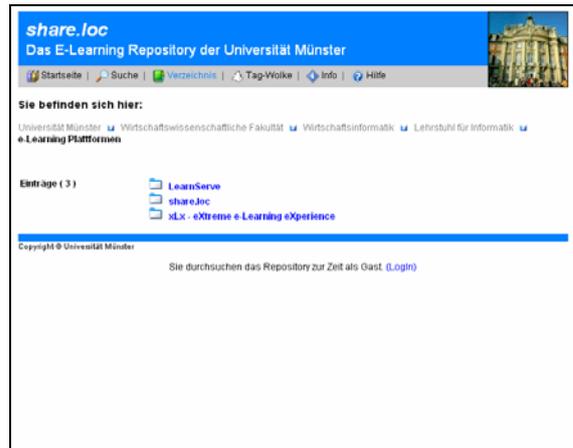


Abbildung 6: Navigation im Verzeichnis

### 3.3 Anlegen / Ändern von Metadatenätzen und Verzeichnisstrukturen

Das Anlegen und Editieren von share.loc-Einträgen ist ausschließlich authentifizierten Benutzern gestattet. Über den individuellen ZIV-Account kann sich ein Benutzer auf der Startseite der Anwendung anmelden, wodurch der Gaststatus verlassen wird. Besucht der Anwender nun das Verzeichnis, so kann er nicht nur selbiges navigieren, sondern hat zusätzlich die Option, die vorhandenen Strukturen zu verändern bzw. zu ergänzen oder einen neuen Metadatenatz anzulegen (Abbildung 7).

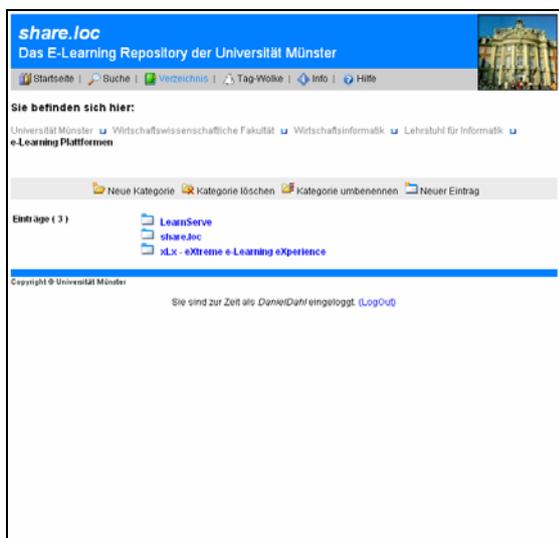


Abbildung 7: Verzeichnisansicht mit Optionen

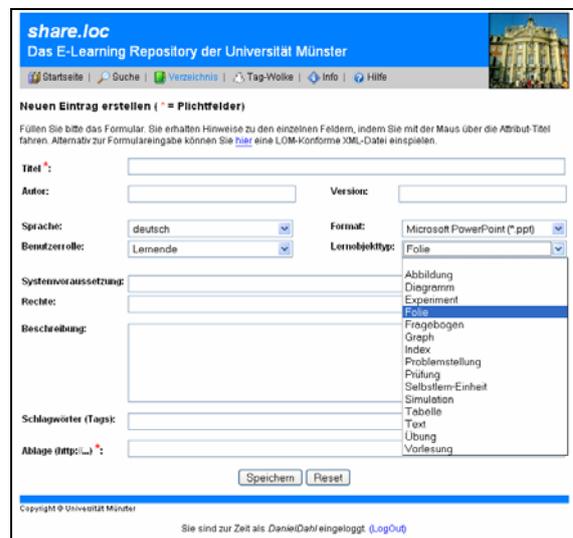


Abbildung 8: Anlegen eines neuen Datensatzes

Das Anlegen eines neuen Datensatzes kann auf zwei Wegen erfolgen (Abbildung 8). Entweder füllt der Anwender das vorliegende, auf relevante LOM-Attribute reduzierte Webformular (bestehend aus Drop-Down-Listen und Freitextfeldern), oder er lädt eine dem LOM XML-Binding [14] genügende Metadatenbeschreibung in das System. Das Ändern eines bereits vorhandenen Datensatzes geschieht analog, wobei die bereits vorhandenen Daten editierbar angezeigt werden.



bislang noch keinen Kontakt mit elektronischem Lernen haben, sollen nun einfach und schnell feststellen können, welche Potentiale e-Learning bietet und wer bereits erfolgreich e-Learning einsetzt, so dass die Kontaktaufnahme problemlos möglich ist. Nehmen wir an, ein Interessierter, mit der Thematik bisher nicht tiefgreifend vertrauter Dozent findet sich auf der share.loc Startseite wieder. Für e-Learning Neulinge bringen die bisher bekannte Sichten auf Repository-Inhalte große Hürden mit sich: Wonach soll der Benutzer suchen? Er kann weder gezielt Autoren, noch Titel oder Formate von Lernobjekte in den Suchdialogen spezifizieren. Die Suche über das Verzeichnis ist ebenso wenig hilfreich: In welchen Teilbaum der Verzeichnisstruktur muss navigiert werden, um interessante Inhalte zu finden? Dem Neuling ist schließlich nicht bekannt, welche Fakultät oder welcher Lehrstuhl sich bereits intensiv mit e-Learning auseinandergesetzt hat. Die neu eingeführte Tag-Wolke bringt an dieser Stelle einen erheblichen Mehrwert: der Benutzer sieht auf einen Blick, zu welchen Thematiken bereits Lernobjekte existieren. Zudem ist auf Anhieb ersichtlich, welche Bereiche bereits eine große Menge an e-Learning Angeboten aufweisen können. Wie beschrieben ist eine Verfeinerung der Auswahl intuitiv möglich. Auch Personen, die sich bereits intensiv mit e-Learning und Lernobjekten in Repositories beschäftigen, bietet die Tag-Wolke einen innovativen Mehrwert: Implizite Zusammenhänge zwischen Lernobjekten werden explizit dargestellt, indem mit gleichen Tags gekennzeichnete Inhalte in einem Feld der Tag-Wolke zusammengefasst werden. Anders als in der Verzeichnis-Sicht werden Kategorien und deren zugehörige Inhalte nicht explizit hierarchisch angelegt. Eine Zusammengehörigkeit ergibt sich implizit, wenn das gleiche Tag für mehrere Lernobjekte benutzt wird. Die Visualisierung dieser impliziten Zusammenhänge geschieht durch die Tag-Wolke. Dem Benutzer ist es nun möglich, Beziehungen zwischen Lernobjekten zu erkennen, die weder eine Gemeinsamkeit bezüglich Autor oder Titel, noch bezüglich des Ablageortes im Verzeichnis haben. So lassen sich z. B. durchaus denkbare Zusammenhänge zwischen fachfremden Lernobjekten explizit aufzeigen.

#### 4. Schnittstellen zu externen Systemen

Neben der Transparenz von Lernobjekten ist deren Integration in bestehende Strukturen eine hervorzuhebende Herausforderung. Nicht nur an der Universität Münster sind in den vergangenen Jahren monolithische Systeme, aufgrund fehlender Kommunikation mit anderen Plattformen oft als „Insellösungen“ bezeichnet, entstanden. Lernangebote existieren innerhalb verschiedenster Systeme, ein automatisierter externer Zugriff ist jedoch nicht möglich. Das aufkommende Paradigma der Service-Orientierung zeigt auf, wie ein hohes Integrationspotential erzielt werden kann. Bei der Entwicklung des Repositories share.loc wurde dieses Konzept durchgängig umgesetzt, indem Schnittstellen zu jeder einzelnen Schicht der Systemarchitektur realisiert wurden [16]. Somit ist eine Integration in etlichen Szenarien denkbar.

Die Systemarchitektur einer Webapplikation ist gewöhnlich in drei Schichten unterteilt. Während die Datenschicht die Persistenz sicherstellt, realisiert die Logikschicht die Kernfunktionalität der Anwendung und die Präsentationsschicht die Visualisierung. Gehen wir davon aus, dass unterschiedliche Anwendungen unterschiedliche Repräsentationsformen von Metadaten benötigen, so genügt es nicht mehr, eine einzige Schnittstelle für den Datenaustausch zur Verfügung zu stellen. Folgende Szenarien sollen diesen Ansatz verdeutlichen:

- i) Ein Meta-Repository nach [17] stellt an einer zentralen Stelle Inhalte mehrerer verteilter Repositories zur Verfügung.
- ii) Die zentrale Suchmaschine der Universität Münster soll neben Internet- und Intranetseiten auch das Repository nach Inhalten durchsuchen. Die Ergebnisse werden in Form eines navigierbaren Touchgraphen dargestellt.
- iii) Ein Portal soll benutzerabhängig Grundfunktionalitäten des Repositories share.loc zur Verfügung stellen.

In Szenario i) ist es wünschenswert, die Metadaten in Form von rohem XML auszutauschen. Die Metadaten sind vollständig und gleichzeitig frei von kontextabhängigen Zusatzinformationen. In Szenario ii) ist der Austausch über reines XML ebenfalls denkbar, erscheint jedoch nicht sinnvoll. Zur Aufbereitung der Metadaten, z. B. in einem Touchgraph, müssen lediglich bestimmte Attribute ausgelesen werden. Hier wäre es wünschenswert, im

Sinne der Objektorientierung gezielt Methoden auf Objekten auszuführen um die nötigen Informationen zu erhalten. Szenario iii) geht noch einen Schritt weiter: auch hier könnte rohes XML ausgetauscht werden, welches jedoch geparkt und in eine geeignete Darstellung überführt werden müsste. Eher sollte der Datenaustausch auf HTML-Fragmenten basieren, die die nötigen Metadatenattribute bereits in visuell aufbereiteter Form enthalten.

Es ist also festzuhalten, dass in unterschiedlichen Szenarien unterschiedliche Repräsentationen der Metadaten erwünscht sind. Wird diese Feststellung missachtet, so müssen bestimmte Prozesse in jedem integrierenden System implementiert werden. Würde z. B. lediglich eine Schnittstelle auf XML-Basis realisiert, so müssten integrierende Systeme die rohen Daten parsen und in eine weiterverarbeitbare Form überführen. Würde eine Schnittstelle lediglich auf HTML-Fragmenten basieren, so müssten Systeme, die ausschließlich an Inhalten interessiert sind, diese aus dem HTML-Code extrahieren. In beiden Fällen sind eine unnötige Fehleranfälligkeit und eine erhöhter Verarbeitungsaufwand gegeben. Als Ausweg wird der Ansatz verfolgt, individuelle Schnittstellen, die auf die unterschiedlichen Schichten der Systemarchitektur zugreifen, zur Verfügung zu stellen.

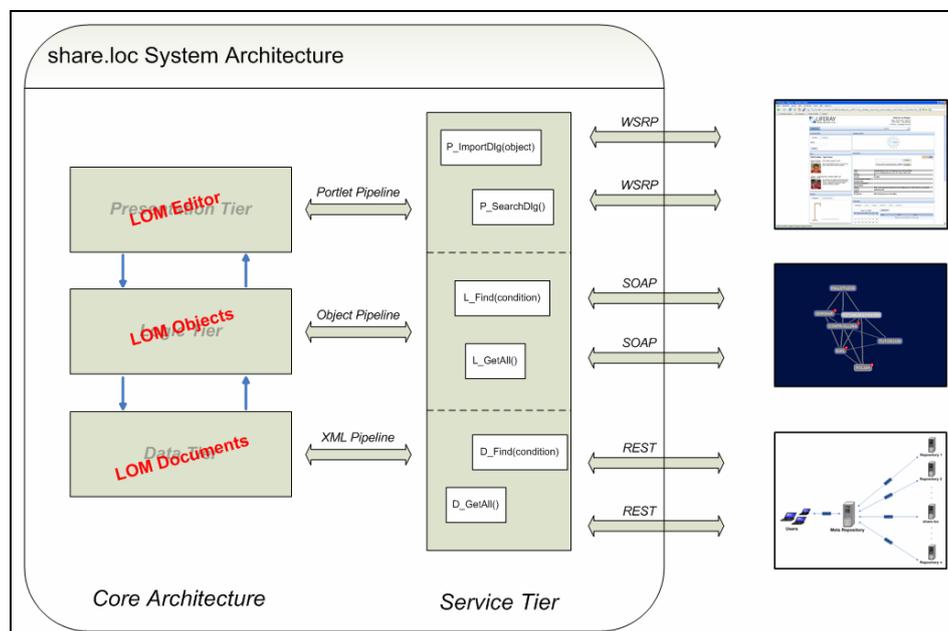


Abbildung 11: share.loc Systemarchitektur und Schnittstellen

Die in Abbildung 11 abgebildete dreischichtige Kernarchitektur ist als Webapplikation realisiert. Benutzer können über den Internetbrowser alle Funktionalitäten des Systems nutzen. Das Konzept der Kernarchitektur besteht darin, in jeder einzelnen Schicht mit standardisierten Repräsentationen der Lernobjektmetadaten zu arbeiten. Im Detail bedeutet dies, dass in der Datenschicht die Metadaten gemäß dem LOM XML-Binding in einer XML-Datenbank abgelegt werden. Die Logik-Schicht greift auf die Datenbank zu und überführt die rohen XML-Daten in Java-Objekte. Dies geschieht automatisiert durch den Einsatz des LOM Java-Bindings [18], wodurch auch hier von einer standardisierten Repräsentationsform gesprochen werden kann. Die Logikschicht liest durch Objektzugriffe relevante Metadaten aus und reichert diese mit HTML-Auszeichnungen an. Die dem Benutzer präsentierte GUI erlaubt dem Benutzer das Anlegen neuer Metadatenätze, wobei ausschließlich LOM-konforme Daten akzeptiert werden. Über DropDown-Listen und Validierungsmechanismen wird sichergestellt, dass sowohl die Attributtypen als auch das verwendete Vokabular dem Standard entspricht. Folglich wird dem Anwender ein LOM-Editor (vgl. [19], [20]) angeboten, dessen Attributanzahl unter Berücksichtigung des in Abschnitt 2 angesprochenen Einsatzumfelds auf 15 Attribute reduziert wird (siehe auch Abbildung 8). Zusammenfassend werden die Lernobjektmetadaten in der gesamten Kernarchitektur standardkonform verarbeitet, wodurch ein großes Potential für externe Schnittstellen vorhanden ist.

Das vorhandene Potential wird durch die Einführung einer vertikal ausgerichteten Service-Schicht nutzbar. Diese bietet externen Systemen Dienste, die über so genannte Pipelines auf die einzelnen Schichten der Kernarchitektur zugreifen. So werden z. B. insgesamt drei verschiedene Suchdienste angeboten, die mit externen Systemen über standardisierte Protokolle kommunizieren. Der über das REST-Protokoll (REpresentational State Transfer) angesprochene Suchdienst greift über die XML-Pipeline auf die Logikschicht der Kernarchitektur zu und liefert die Suchergebnisse in Form von rohem XML. Der über SOAP (Simple Object Access Protocol) erreichbare Suchdienst greift auf die Logikschicht zu und liefert Java-Objekte, die eine besonders komfortable Weiterverarbeitung von individuell relevanten Daten ermöglichen. Mittels WSRP (Web Services for Remote Portlets) kann sogar ein auf die Suchfunktion reduzierter Dialog über die Portlet-Pipeline angefordert werden. Hier ist kein weiterer Verarbeitungsschritt auf der anfragenden Seite mehr nötig, da bereits eine Anreicherung um Visualisierungskomponenten in der Präsentationsschicht der Kernarchitektur stattgefunden hat.

Zurückkommend auf die eingangs angesprochenen Szenarios bieten sich externen Systemen individuelle Wege, Grundfunktionen des e-Learning Repositories zu integrieren:

- i) Das Meta-Repository fordert reine XML-Daten über das REST-Protokoll an. Die Daten genügen dem LOM-XML Standard, wodurch eine automatisierte Verarbeitung möglich ist.
- ii) Die zentrale Suchmaschine führt über die SOAP-Schnittstelle eine Suche im Datenbestand des Repositories durch. Da in der Visualisierung (Touchgraph) zunächst nur der Titel der Ergebnisse relevant ist, werden die entsprechenden Informationen durch einfache Objektaufrufe extrahiert. Es ist nicht nötig, beispielsweise einen XML-Parser zu implementieren.
- iii) Eine besonders komfortable Form der Integration bietet sich dem Portal, das über WSRP einen Suchdialog einbinden kann. Es ist lediglich nötig, die Zieladresse des entsprechenden Dienstes in der share.loc-Architektur anzugeben. Weitere Verarbeitungsschritte sind somit obsolet.

## 5. Zusammenfassung

In diesem Papier wurden zunächst die Anforderungen an ein universitätsweit einzusetzendes e-Learning Repository beschrieben. Aus diesen wurden Designprinzipien für das an der Universität Münster entwickelte Verzeichnis share.loc abgeleitet. In Form der Grundfunktionen des entwickelten Systems wurde aufgezeigt, wie eine Transparenz von verteilten Lernobjekten erzielt werden kann. Hervorzuheben ist der innovative Ansatz, neben bekannten Sichten wie der schlüsselwortbasierten Suche und der Navigation durch ein kontextabhängiges Verzeichnis das Konzept der aus dem Web 2.0 bekannten Verschlagwortung und der einhergehenden Tag-Wolke auf Lernobjekte zu übertragen. Des Weiteren wurde aufgezeigt, wie der Ansatz der mehrschichtigen Interoperabilität ein hohes Integrationspotential für bestehende Systeme und Strukturen freisetzt. Externe Applikationen erhalten über standardisierte Protokolle auf die jeweiligen Anforderungen zugeschnittene Metadatenrepräsentationen, die selbst wiederum konform zum anerkannten LOM-Standard sind. Eine Einbindung in ein Universitätsportal oder eine übergreifende Suchmaschine ist ohne aufwändige Verarbeitungsschritte möglich.

Durch die Einführung des Tagging in den Anwendungsbereich der e-Learning Repositories ergibt sich eine Vielzahl interessanter Fragestellungen. Zunächst lässt sich das Tagverhalten der Systembenutzer analysieren: Wieviele Tags werden pro Lernobjekt angelegt? Auf wieviele Lernobjekte verweist ein Tag? Durch Beantwortung dieser Fragen lässt sich unter Anderem ermitteln, welche Arten von Tags gesetzt werden: Handelt es sich eher um Schlüsseltags, also generelle Tags, die an sehr vielen Lernobjekten angebracht wurden, oder eher um spezielle Tags, die ein Lernobjekt sehr detailliert kategorisieren und somit lediglich an diesem sinnvoll anzubringen sind? Wie sehen die Beziehungen zwischen Lernobjekten aus, die sich implizit durch das Setzen von gemeinsamen Tags ergeben? Die Beantwortung dieser Frage ist insbesondere unter dem Aspekt, dass erwartungsgemäß die Taganzahl schneller ansteigt als die Anzahl der referenzierten Lernobjekte, von großer Bedeutung. Sind vor diesem Hintergrund die Beziehungen, die in der Tag-Wolke explizit dargestellt werden, von anhaltender praktischer Relevanz?

Eine weitere Fragestellung behandelt das der share.loc-Architektur zugrunde liegende LOM-Datenmodell. Aktuell zeichnet sich die Entwicklung eines produktiv einsetzbaren LOM RDF-Bindings ab. Im Sinne der mehrschichtigen Systemarchitektur ist es durchaus denkbar, in der Datenschicht eine Umstellung von XML auf RDF vorzunehmen. Es ist abzuwägen, ob die resultierenden Vorteile wie z. B. das aus dem Semantic Web bekannte Reasoning im share.loc-Kontext eine Neuausrichtung rechtfertigen.

## 6. Literatur

- [1] G. Vossen, J. Schwieren, "Umfrageergebnisse e-Learning an der Westfälischen Wilhelms-Universität Münster", Working Paper 2006
- [2] Freestyle Learning, <http://www.freestyle-learning.de>, *zuletzt besucht am 24.03.2007*
- [3] OpenUSS, <http://openuss.uni-muenster.de>, *zuletzt besucht am 24.03.2007*
- [4] Herbie, <http://herbie-group.de>, *zuletzt besucht am 24.03.2007*
- [5] xLx – eXtreme e-Learning eXperience, <http://dbms.uni-muenster.de/xLx>, *zuletzt besucht am 24.03.2007*
- [6] D.A. Wiley, "The Instructional Use of Learning Objects", Agency for Instructional Technology, 2002
- [7] IEEE Learning Object Metadata Standard, <http://ltsc.ieee.org/wg12/20020612-Final-LOM-Draft.html>, *zuletzt besucht am 24.03.2007*
- [8] share.loc - Shared Repository For Learning Object Content, <http://shareloc.uni-muenster.de>, *zuletzt besucht am 24.03.2007*
- [9] J. Brase, W. Nejdl, "Ontologies and Metadata for eLearning", *Handbook on Ontologies*, Springer 2003
- [10] Pflichtenheft e-Learning Repository der Universität Münster, ERCIS (European Research Center for Information Systems): Kontaktperson Daniel Dahl, <http://www.ercis.de>, *zuletzt besucht am 24.03.2007*
- [12] E. Duval, F. Neven, "Reusable Learning Objects: a Survey of LOM-Based Repositories", Proceedings of the 10th ACM International Conference on Multimedia, Juan les Pins (France), 2002
- [12] G. P. Schell, M. Burns, "Merlot: A Repository of e-Learning Objects for Higher Education", e-Service Journal 1-2 2002, Indiana University Press
- [13] LearnAlberta, <http://www.learnalberta.ca>, *zuletzt besucht am 24.03.2007*
- [14] IEEE Learning Object Metadata XML-Schema, <http://ltsc.ieee.org/xsd/lomv1.0/>, *zuletzt besucht am 24.03.2007*
- [15] T. O'Reilly, "What is Web 2.0", <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>, 2005, *zuletzt besucht am 24.03.2007*
- [16] D. Dahl, G. Vossen, P. Westerkamp, "share.loc – A Multi-Tiered Interoperable E-Learning Metadata Repository", Proceeding of the 6th IEEE International Conference on Advanced Learning Technologies (ICALT), Kerkrade, The Netherlands, 2006
- [17] D. R. Rehak, P. Dodds, L. Lannom, "A Model and Infrastructure for Federated Learning Content Repositories", 14<sup>th</sup> International World Wide Web Conference, Chiba 2005
- [18] LOM Java-Binding & Implementation, <http://www.cancore.ca/swcomponents.html>, *zuletzt besucht am 24.03.2007*
- [19] LOM-Editor v1.0, <http://www-i5.informatik.rwth-aachen.de/i5new/staff/chatti/LOMEditor/index.html>, *zuletzt besucht am 24.03.2007*
- [20] Z. Cebeci, Y. Erdogan, "Tree View Editing Learning Object Metadata", Interdisciplinary Journal of Knowledge and Learning Objects Volume 1, 2005

# DB2 goes SQL:2003 : Native SQL/XML-Unterstützung der IBM DB2-Datenbank

Gunnar Thies, Gottfried Vossen

Institut für Wirtschaftsinformatik

Universität Münster

Leonardo-Campus 3, 48161 Münster

Abstract

Die Verarbeitung von XML-Daten in objektrelationalen Datenbanken war lange Zeit nur unzureichend umgesetzt, obwohl die SQL/XML-Erweiterung bereits im SQL:2003-Standard definiert wurde. Mittlerweile ist in den aktuellen Releases vieler Datenbankhersteller diese Erweiterung umgesetzt. In diesem Bericht setzen wir uns mit der IBM DB2 Version 9 auseinander und zeigen, wie IBM's SQL/XML-Erweiterung den SQL:2003-Standard umsetzt. Dabei werden nicht nur XML-Inhalte am Beispiel eingelesen, bearbeitet und ausgegeben, sondern auch XML-Schemata zum Validieren eingesetzt und Indizes für XML-Daten erzeugt.

## 1 Einführung

Die *Extensible Markup Language* (XML) erlangt immer größere Beliebtheit in Wissenschaft und Wirtschaft. Dieses semi-strukturierte Format zur Speicherung von Daten findet sich vor allem in großen Teilen des World Wide Web: Viele Standards (beispielsweise des W3C) werden durch XML-Strukturen definiert und beschrieben. XML basiert auf der *Standard Generalized Markup Language* (SGML), die unter anderem von Charles Goldfarb entwickelt wurde und sich seit 1986 als ISO-8879 durchsetzte. XML 1.0 wurde 1996/1997<sup>1</sup> vom W3C vorgestellt.<sup>2</sup> XML basiert auf einer wohlgeformten Baumstruktur, die zunächst nur folgenden Regeln folgt:

- Es gibt genau ein Wurzelement;
- jedes Tag beginnt mit einem Starttag und endet mit einem End-Tag (Ausnahmen sind leere Tags, diese enden mit />);
- Verschachtelungen innerhalb des Dokuments müssen streng geordnet sein und in umgekehrter Reihenfolge ihres Öffnens wieder geschlossen werden.

---

<sup>1</sup> Das genaue Datum ist schwer feststellbar, vgl. [www.w3c.org](http://www.w3c.org) oder [MI02], S. 26.

<sup>2</sup> Nähere Informationen zum 10-jährigen Bestehen des W3C unter: <http://www.w3.org/2004/11/15-presskit-single.html>

Die Struktur eines XML-Dokuments lässt sich über eine Grammatik beschreiben. Die beiden zur Verfügung stehenden Grammatik-Formalismen sind *Document Type Definition* (DTD) oder *XML-Schema*, anhand derer die Gültigkeit eines XML-Dokuments – zur im Grammatik-Dokument vorgeschriebenen Struktur – geprüft werden kann. Die aktuellste Empfehlung zu XML ist in der W3C Recommendation des 16. August 2006 zu XML 1.0 (Fourth Edition)<sup>3</sup> nachzulesen.

XML wird beispielsweise in der Prozessmodellierung für die *Business Process Execution Language* (BPEL), bei Web Services für die *Web Service Description Language* (WSDL, Version 2.0 seit Juni 2007), bei Vektorgrafiken für *Scalable Vector Graphics* (SVG), beim Aufbau von Webseiten für die *Extensible Hypertext Markup Language* (XHTML) sowie in vielen anderen Anwendungen verwendet. Um spezifische Daten aus einem XML-Dokument zu extrahieren, gibt es mehrere Möglichkeiten. Zum Einen ist hierbei die Sprache *XSLT* zu nennen, die es ermöglicht, aus einem XML-Dokument durch spezifische Mechanismen ein Subset (einen Teilbaum) als neue XML-Datei zu erhalten. Zum Anderen gibt es die Sprachen *XPath* und *XQuery*, die es ermöglichen, unterschiedliche Formen von Pfadausdrücken sowie Daten aus einem XML-Dokument nach komplexen Anfragen zu extrahieren.

Traditionell werden Daten in strukturierter Form heute überwiegend in relationalen Datenbanken gespeichert. Gängige Systeme sind IBM DB2, Oracle 10g, MySQL und PostgreSQL um nur einige zu nennen. Die Daten werden darin in Tabellen angeordnet und in Tupeln (Zeilen) abgespeichert. Um Daten aus einer Datenbank zu extrahieren, editieren und zu verarbeiten verwendet man SQL (und einige systemspezifische Dialekte bzw. Erweiterungen). Damit können weitgehend komplexe Anfragen an eine Datenbank formuliert werden.

Oft haben XML-Daten und relationale Datenbanken bisher nebeneinander existiert.<sup>4</sup> Das Abspeichern von XML-Dateien in relationalen Datenbanken als Datentyp Binary (oder BLOB) ist schon seit längerer Zeit möglich. Der große Nachteil daran war bisher aber, dass XML-Inhalte nur als Ganzes ausgelesen und bearbeitet werden konnten: bei DB2 beispielsweise mit dem XML Extender, obwohl die als SQL/XML bezeichnete SQL-XML-Erweiterung für relationale Datenbankmanagementsysteme bereits in dem SQL:2003 Standard definiert wurde. Diese native XML-Unterstützung (bei DB2 als „pureXML“ bezeichnet), die einen direkten Bezug zwischen dem Inhalt einer XML-Datei und den Inhalten einer relationalen Datenbank herstellen kann, wird mittlerweile in einigen aktuellen Versionen relationaler oder objektrelationaler Datenbanksysteme geboten.

Um einen generellen, herstellerunabhängigen Überblick über den in SQL:2003 definierten Befehlsumfang von SQL/XML zu erlangen, bietet sich Standardlektüre wie [TUE03] oder [ST05] an. Wir wollen in diesem Bericht nicht die Gesamtheit der SQL/XML Syntax von SQL:2003 vorstellen, vielmehr soll deren konkrete Umsetzung innerhalb der DB2 Version 9 exemplarisch gezeigt werden. Dazu gehen wir im Folgenden auf verschiedene Arten des Einfügens und des Bearbeitens von XML-Daten und darüber hinaus auf das Validieren anhand von XML-Schema-Definitionen und das Erstellen von Indizes für XML-Daten ein.

In Abschnitt 2 wird auf die Unterstützung von XML in der neuesten Version der relationalen Datenbank DB2 (Version 9.0, zuvor als Viper bezeichnet) von IBM eingegangen und untersucht,

---

<sup>3</sup> <http://www.w3.org/TR/2006/REC-xml-20060816/>

<sup>4</sup> Es existieren auch XML-Datenbanken, wie bspw. das Open Source Projekt *eXist* (<http://exist.sourceforge.net>), die XML-Daten in Datenbankmanier abspeichern und die Daten per XQuery und XUpdate (einer Möglichkeit der einfachen XML-Editierung) bearbeiten können.

welche Einsatzszenarien sich damit ergeben.<sup>5</sup> Abschnitt 3 widmet sich dann einer kurzen Vorstellung zweier möglicher Einsatzszenarien der Verknüpfung von XML und relationaler Datenbanken bzw. der Verwendung der nativen XML-Unterstützung. In Abschnitt 4 wird abschließend ein Fazit gezogen.

## 2 Nutzungsszenarien

Man kann hierbei zunächst zwischen zwei grundlegend Fälle unterscheiden:

- Speichern von XML-Daten in der Datenbank
- Anfragen oder Bearbeiten von gespeicherten XML-Daten aus der Datenbank

Der erste Fall lässt lediglich zwei Unterscheidungen zu: die ungeprüfte Speicherung von XML und die Speicherung nach Validierung gegen ein XML-Schema (bzw. einer DTD). Der zweite Fall lässt sich in einige verschiedene Unterarten einteilen. So lassen sich bspw. Teile aus den XML-Daten extrahieren und editieren (per XQuery oder SQL), relationale Daten und XML-Daten in einer SQL-Abfrage mischen und sogar Views über XML-Dateien erstellen. Im Folgenden wird ein gängiges Beispielszenario aus den Tutorials<sup>6</sup> von IBM kurz vorgestellt, und es werden anhand derer die Möglichkeiten der DB2 erläutert. Im Anschluss daran wird ein kurzer Blick auf die Indizierungsmöglichkeiten der XML-Daten geworfen.

Für unser Beispiel werden zwei Tabellen in der Datenbank angelegt (siehe Tabelle 1):

- Tabelle Items (Produkte)
- Tabelle Client (Kunden)

Wichtig ist hierbei jeweils das letzte Attribut, das den Datentyp XML aufweist. Die Datei wird hier binär gespeichert und kann dann durch XQuery oder SQL/XML (SQL angereichert mit XML-spezifischen Teilen) durchsucht werden.

Tabelle: Items		Tabelle: Clients	
Name	Datentyp	Name	Datentyp
id	int, PK	id	int, PK
brandname	varchar(30)	name	varchar(50)
itemname	varchar(30)	status	varchar(10)
sku	int	contactinfo	Xml
srp	decimal(7,2)		
comments	xml		

**Tabelle 1: Datentabellen für das Beispiel.**

<sup>5</sup> Dabei wird die Funktionsweise von XQuery und XPath nicht näher erklärt.

<sup>6</sup> Es existiert mittlerweile ein IBM Redbook, das sich mit der Thematik pureXML beschäftigt: <http://www.redbooks.ibm.com/abstracts/SG247298.html?Open>

Hat man nun XML-Dateien für Kundeninformationen und Produktkommentare vorliegen, werden diese entsprechend der schematischen Zeichnung in Abbildung 1 zu einem jeweiligen Datensatz gespeichert. Pro Datensatz kann hier jeweils eine komplette XML-Datei gespeichert werden.

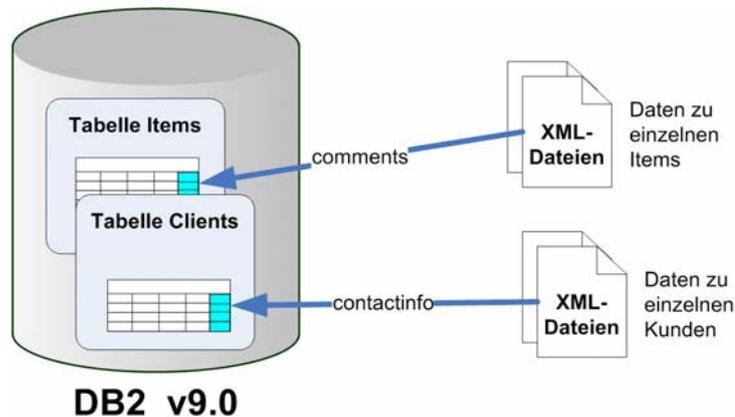


Abbildung 1: XML-Daten-Speicherung in der DB2.

Auf die Darstellung der SQL-Syntax zur Erstellung der Tabellen wird hier verzichtet.

## 2.1 Speichern von XML-Dateien

Nachdem die Datenbankstruktur angelegt ist, sollen die ersten Datensätze in den Tabellen gespeichert werden. Um schnelle Ergebnisse zu erlangen, schicken wir folgendes SQL-Statement direkt zur Datenbank:

```
insert into clients
values (1, 'Max Müller', 'Gold',
'<addr>Musterstr. 1, Münster, Germany</addr>');
```

Listing 1: Einfaches Insert-Statement.

Somit wird ein Datensatz angelegt, der im Attribut *contactinfo* die Adressdaten in der angegebenen XML-Form speichert.

Es gibt eine weitere Möglichkeit, XML-Strukturen in die Datenbank zu speichern, die ein wenig komfortabler ist. Hierfür benötigt man die XML-Quelldateien und deren Namen in einem Verzeichnis (hier zur Vereinfachung C:/IBM/XMLdemo) und erstellt eine ASCII-Datei, die die einzutragenden relationalen Daten und einen Verweis auf die dazugehörigen XML-Dateien enthält (Listing 2). Diese Datei speichern wir unter dem Namen *clients.del*.

```
2, Ute Mayer, Gold, <XDS FIL='Client2.xml' />
3, Max Frisch, Gold, <XDS FIL='Client3.xml' />
4, Peter Müller, Silver, <XDS FIL='Client4.xml' />
```

**Listing 2: ASCII-Datei zum Datenimport.**

Der Inhalt einer ClientX.xml-Datei sieht hier beispielsweise aus wie in Listing 3:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<client>
  <address>
    <street>Musterstr. 4</street>
    <city>Münster</city>
    <state>Germany</state>
    <zip>48155</zip>
  </address>
  <phone>
    <work>0251/1234567890</work>
    <home>0251/1212121212</home>
    <cell></cell>
  </phone>
  <email>ute.mayer@uni-muenster.de</email>
</client>
```

**Listing 3: XML-Datei eines Kunden.**

Es muss nun für jede in Listing 2 aufgeführte Zeile eine entsprechende XML-Datei vorhanden sein; dann kann mit folgendem Statement der Datenimport gestartet werden.

```
import from C:/IBM/XMLdemo/clients.del of del xml from C:/IBM/XMLdemo
insert into clients;
```

**Listing 4: SQL-Befehl zum Import.**

Sollten die XML-Dateien gefunden werden, so sind anschließend alle drei Personen mit ihren zugehörigen XML-Daten in der Datenbanktabelle *clients* zu finden.

Der Vorteil von XML-Dateien gegenüber Plaintext oder CSV-Dateien ist die durch Grammatiken ermöglichte Gültigkeitsprüfung. So kann eine XML-Datei – bevor sie in der Datenbank gespeichert wird – auf Validität überprüft werden. Daher kann man im XML-Schemarepository (XSR) der DB2 XML-Schemas registrieren, die diese Prüfung übernehmen. Zunächst gehen wir von folgendem XML-Schema (Listing 5) aus.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="address">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="street" type="xsd:string"/>
        <xsd:element name="city" type="xsd:string" />
        <xsd:element name="state" type="xsd:string" />
        <xsd:element name="zip" type="xsd:integer"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="phone">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="work" type="xsd:string"/>
        <xsd:element name="home" type="xsd:string"/>
        <xsd:element name="cell" type="xsd:string"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="client">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element ref="address"/>
        <xsd:element ref="phone"/>
        <xsd:element name="email" type="xsd:string"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>

```

**Listing 5: XML-Schema für Kunden.**

Dieses (rudimentäre) Schema entspricht exakt der XML-Struktur unsere ClientX.xml-Dateien und soll nun im XSR der DB2 registriert werden. Es ist unter C:/IBM/ClientSchema.xsd gespeichert und kann mit folgendem Statement als Schema „mySchema“ registriert werden.

```

register xmlschema 'http://placeholder.for.namespaces' from
  C:/IBM/XMLdemo/ClientSchema.xsd as mySchema complete;

```

**Listing 6: Registrierung eines XML-Schemas.**

Nach erfolgreicher Registrierung kann man XML-Daten beim Importieren per ASCII-Datei von diesem Schema validieren lassen. Wir löschen zunächst alle Daten aus der Tabelle *Clients* und fügen die Daten aus der ASCII-Datei (Listing 2) mit Hilfe des nachfolgenden Befehls wieder zur

Tabelle hinzu; die Besonderheit: diesmal wird das Schema „mySchema“ als Validierungsschema angegeben. Die XML-Dateien, die nicht gültig sind, werden auch nicht in die Tabelle gespeichert.

```
import from C:/IBM/XMLdemo/clients.del of del xml from C:/IBM/XMLdemo
xmlvalidate using xds default mySchema
insert into clients;
```

### Listing 7: Hinzufügen von Daten mit XML-Schema-Validierung.

Auftretende Fehler beim Importieren der Daten werden nun von DB2 zurückgegeben und man hat die Möglichkeit, die XML-Dateien dementsprechend zu verbessern. DB2 bietet auch die Möglichkeit eine DTD zu hinterlegen und dann darauf die XML-Validitätsprüfung vorzunehmen, dies wird hier aber nicht weiter betrachtet.

Wir haben nun die wichtigsten Möglichkeiten besprochen, wie relationale Daten und XML-Dateien in eine Tabelle gespeichert werden. Es folgt eine Betrachtung der vielfältigen Ausgabemöglichkeiten per XQuery und SQL (bzw. SQL/XML).

## 2.2 Verarbeitung und Anfrage von XML-Dateien

Beim Anfragen von Daten gibt es im Falle der DB2 (Viper bzw. Version 9) zwei Möglichkeiten. Die erste realisiert Abfragen mit SQL/XML, die zweite stellt Anfragen in XQuery-Notation. Anhand der folgenden Beispiele soll exemplarisch gezeigt werden, wie gängige Anfragen an die Datenbank bzw. deren XML-Inhalte gestellt werden können. Die Ausgangsbasis für die hier besprochenen Anfragen bilden die oben angelegten Tabellen und deren Inhalte.

### 2.2.1 Anfragen mit SQL/XML

Kenntnisse über normale „Select From Where“-Statements und deren Aufbau werden hier vorausgesetzt. Um die XML-Daten aus der Datenbank anzusprechen, verwenden wir hier die spezielle XML-Syntax von IBM als Ergänzung zum Standard-SQL.

#### SQL/XML

Wir gehen von obiger XML-Kundenkontaktinformation aus (Auszug einer ClientX.xml-Datei):

```
<?xml version="1.0"? encoding="ISO-8859-1">
<client>
  <address>
    <street>Musterstr. 4</street>
    <city>Muenster</city>
    <state>Germany</state>
    <zip>48155</zip>
  </address>
  ...
```

### Listing 8: Auszug aus der ClientX.xml.

Betrachten wir nun folgende Aufgabe: Wir suchen alle Zeilen der Tabelle *Clients*, die in Münster wohnen. Dazu benutzen wir folgendes SQL-Statement (Listing 9).

```
SELECT name FROM clients
WHERE xmlexists('$c/client/address[city="Münster"]'
               passing clients.contactinfo as "c")
```

**Listing 9: SQL-XML-Statement.**

Man sieht im Statement, dass das Attribut *contactinfo*, in dem die XML-Daten der Kunden liegen, durch den Befehl *passing clients.contactinfo as „c“*, an die Variable *c* gebunden wird. Anschließend kann durch die Traversierung am XML-Baum durch den XPath *\$c/client/adress* ein Attribut auf der Stufe *address* – hier *city* – als WHERE-Bedingung angegeben werden. Somit erhalten wir alle Namen der Kunden aus der Tabelle *client*, in denen die XML-Datei den Ort der Person als „Münster“ angibt. Wichtig ist hier die Notation mit den eckigen Klammern (*...address[city="Münster"]*), die nicht mit einer SQL-typischen Notation verwechselt werden darf (*...address/city = "Münster"*).

In obigem Fall bekamen wir „relational gespeicherte Werte“ zurück: alle Namen, die der WHERE-Klausel entsprachen. Nun wollen wir XML-Informationen erhalten (z.B. die Straße, in der der Kunde wohnt) und verwenden dazu den Befehl *xmlquery* in der SELECT-Klausel (siehe Listing 10).

```
SELECT xmlquery('$c/client/address/street' passing cli-
               ents.contactinfo as "c")
FROM clients
WHERE xmlexists('$c/client/address[city="Münster"]'
               passing clients.contactinfo as "c")
```

**Listing 10: XML-Rückgabewerte per SQL/XML-Klausel.**

Die Rückgabe ist in diesem Fall auf XML-Basis, d.h. es wird jeweils ein Teil der XML-Datei zurückgegeben.

```
<?xml version="1.0" encoding="UTF-16" ?><street>Musterstr. 4</street>
<?xml version="1.0" encoding="UTF-16" ?><street>Autorenstr.
1</street>
<?xml version="1.0" encoding="UTF-16" ?><street>Müllerstr.
40</street>
```

**Listing 11: XML-Ausgabe.**

Dies ist nur ein erstes Beispiel für die Ausgabemöglichkeiten von XML-Daten über die DB2. Es lassen sich auch in SQL-typischer Art Bedingungen und Ausgaben kombinieren, Views über XML-Daten erstellen etc.

Im folgenden Beispiel (Listing 12) verwenden wir den Befehl *xmltable*, um XML-Daten auf gleicher Ebene mit relationalen Daten auszugeben.

```
SELECT c.name, ct.city, ct.workphone
FROM clients as c,
xmltable('$c/client' passing c.contactinfo as "c"
         columns city varchar(100) path 'address/city',
                workphone varchar(100) path 'phone/work') as ct
```

**Listing 12: XML und SQL auf selber Ebene ausgeben.**

Durch die Verwendung des Befehls *xmltable* wird aus der Telefonnummer am Arbeitsplatz (phone/work) und dem Wohnort (address/city) eine temporäre Tabelle erstellt, deren Attribute dann mit den „relationale“ vorliegenden Attributen einheitlich ausgegeben werden kann. Als Ergebnis des obigen Statements erhält man eine Tabelle mit folgenden Informationen:

NAME	CITY	WORKPHONE
Ute Mayer	Münster	0251/1234567890
Max Frisch	Münster	0251/9878889
Peter Müller	Münster	0251/100200300

**Tabelle 2: Ausgabe relationaler und XML-Daten.**

Auf die gleiche Weise lassen sich auch View, Joins etc. über XML-Daten erstellen. Weitere Beispiele sind in [SA06a] zu finden.

## 2.2.2 Anfragen mit XQuery

Eine weitere Möglichkeit der Anfrage (neben den XPath-Ausdrücken) besteht in der Formulierung von XQuery-Statements. XQuery und die Syntax der FLWOR-Ausdrücke werden an dieser Stelle nicht weiter vertieft und es wird davon ausgegangen, dass die Syntax und die Funktionsweise bekannt sind.<sup>7</sup> Um mit einem einfachen Beispiel zu beginnen, wollen wir alle Straßennamen (genau wie oben mit dem SQL/XML-Statement) der Kunden ausgegeben bekommen.

```
xquery
for $y
in db2-fn:xmlcolumn('CLIENTS.CONTACTINFO')/client/address/street
return $y
```

**Listing 13: Einfacher XQuery-Ausdruck.**

Das Ergebnis dieser Anfrage ist eine XML-Struktur der extrahierten Daten und sieht folgendermaßen aus:

---

<sup>7</sup> Nähere Informationen zu XQuery unter <http://www.w3.org/TR/xquery>

```
<street>Musterstr. 4</street>
<street>Autorenstr. 1</street>
<street>Muellerstr. 40</street>
```

**Listing 14: Ergebnis des einfachen X-Query-Ausdrucks.**

Man kann durch die Verwendung des Befehls *text()* (also: *return \$y/text()* ) auch nur den Text ohne die XML-Tags erhalten; dies ist aber hier nicht gewünscht.

Als zweites Beispiel soll hier ein kompletter FLWOR-Ausdruck vorgestellt werden, der die Telefonnummer (der Arbeitsstelle) von jedem Kunden ausgibt, der entweder in Münster wohnt oder überhaupt eine Emailadresse angegeben hat (also irgendeine Form der Kundenkommunikation möglich ist).

```
xquery
for $y in db2-fn:xmlcolumn('CLIENTS.CONTACTINFO')/client
where $y/address/city="Münster" or $y/email!=" "
order by $y/zip
return $y/phone/work
```

**Listing 15: Kompletter FLWOR-Ausdruck.**

Es lassen sich so auch alle weiteren Formen von XQuery-Ausdrücken auf der DB2-Datenbank ausführen: beispielsweise if-then-else-Statements, HTML-Ausgaben, let-Klauseln, etc.

Ein wichtiger Aspekt ist die Verknüpfung von XQuery mit SQL-Statements; Dies kann mit dem Befehl *xquery*, wie auch schon im vorherigen Abschnitt beschrieben, geschehen. Ein Beispiel für die Verwendung ist in Listing 16 zu sehen.

```
SELECT name, xmlquery('for $e in $c/client/email[1] return $e'
                      passing contactinfo as "c")
FROM clients
WHERE status = 'Gold'
```

**Listing 16: Verknüpfung von XQuery und SQL.**

Die Where-Klausel muss auf jeden Fall mit einfachen Anführungszeichen versehen werden!

### 2.2.3 Generierung von XML- oder HTML-Strukturen

Es ist ebenfalls möglich, direkt aus der Datenbank HTML oder XML-Ausgaben zu erstellen, auch aus relationalen Datenbeständen. Dies soll in zwei kurzen Beispielen verdeutlicht werden.

Wenn wir direkt XML-Strukturen erstellen wollen, so müssen wir den Befehl *xmlelement* verwenden; dieser erlaubt das „Erstellen“ von XML-Tags. Dies ist beispielsweise sinnvoll, um vorhandene XML-Strukturen umstrukturiert auszugeben oder vorhandene relationale Daten in XML-

Struktur auszugeben (oder auch relationale und XML-strukturierte Daten in ein XML-Dokument zu schreiben). Im folgenden Listing ist ein Beispiel wiedergegeben, wie Daten aus der Tabelle *Clients* (in gemischter Form) als XML-Dokument ausgegeben werden können.

```
SELECT xmlelement(name "Kunde",
    xmlelement(name "id", c.id),
    xmlelement(name "name", c.name),
    xmlelement(name "stadt",
        xmlquery('$c/client/address/city/text()'
            passing c.contactinfo as "c"))
FROM clients as c
WHERE c.id = 2
```

**Listing 17: XML-Dokument erstellt aus relationalen und XML-Daten.**

Aus dem obigen Listing ergibt sich dann folgendes XML-Dokument:

```
<Kunde>
  <id>2</id>
  <name>Ute Mayer</name>
  <stadt>Muenster</name>
</Kunde>
```

**Listing 18: Aus DB2 erstelltes XML-Dokument.**

Weitere Möglichkeiten ergeben sich hier durch die Verwendung des Befehls *xmlagg*, der es ermöglicht, ganze Sequenzen von XML-Daten auszugeben, Näheres dazu findet man in [CS06]).

HTML-Strukturen sind von Vorteil, wenn wir Daten aus der Datenbank z. B. direkt auf einer Webseite in „realtime“ anzeigen wollen. Dazu verwenden wir die im vorigen Abschnitt gezeigten FLWOR-Ausdrücke.

```
SELECT xmlquery('for $e in $c/client/email[1]/text()
    return <p>{$e}</p>'
    passing client.contactinfo as "c")
FROM clients
WHERE status = 'Gold'
```

**Listing 19: FLWOR-Ausdruck direkt in HTML.**

Mit diesem Konstrukt erhalten wir alle Emailadressen der Kunden, die den Status „Gold“ haben, jedoch als HTML-Ausgabe.

## 2.2.4 Aktualisieren, Editieren und Löschen von XML-Daten

Um XML-Daten zu editieren, werden analog zu der Bearbeitung von relationalen Daten UPDATE- oder DELETE-Statements benutzt. Der einzige Unterschied in der Verwendung der SQL-Befehle besteht beim Aktualisieren von XML-Daten in der Datenbank. Es können hier **keine** partiellen Änderungen an den XML-Daten vorgenommen werden, es kann lediglich der **gesamte** Inhalt geändert werden.

## 2.3 Indexierung von XML-Daten

Hier soll ein kurzer Blick auf die Indexierungsmöglichkeiten der DB2-Datenbank geworfen werden. Die Indexierung erfolgt im Wesentlichen analog zu der von relationalen Daten. Auch dort kann man einen speziellen Index für ein Attribut einer Tabelle anlegen: so ist es also auch hier möglich einen Index auf einen speziellen Teil der XML-Struktur zu erstellen. Man kann hierbei zwei Arten unterscheiden: den Index auf einen speziellen Wert oder den Volltextindex.

### 2.3.1 Wertindex

Eine weitere Neuerung bei XML-verarbeitenden Datenbanken ist die Möglichkeit, einen Index auf ein beliebiges Tag einer XML-Struktur zu erstellen. So lassen sich performante Suchoperationen durchführen. Hier wollen wir am Beispiel der Emailadresse der Kunden einen UNIQUE-Index auf ein XML-Tag erstellen.

```
CREATE UNIQUE INDEX idx1 ON clients(contactinfo)
GENERATE KEY USING XMLPATTERN '/client/email' AS SQL varchar(100)
```

**Listing 20: Indexerstellung auf die Emailadresse der Kunden.**

Durch diese einfache Syntax wird ein Index erstellt, der gewährleistet, dass keine doppelten Emailadressen in den XML-Daten der Tabelle *Clients* auftreten. Im XMLPATTERN-Teil wird der Pfad zum zu indizierenden Teil der XML-Struktur angegeben und durch AS SQL Type lässt sich der Datentyp des Index für das spezielle Datum angeben. Folgende Datentypen können verwendet werden: VARCHAR(n), VARCHAR HASHED,DOUBLE, DATE und TIMESTAMP.

### 2.3.2 Volltextindex

Der Volltextindex ermöglicht bessere Suchergebnisse auf dem kompletten Bestand der XML-Daten und sollte bei größeren Textsammlungen innerhalb der XML-Dateien verwendet werden. Ein Volltextindex macht am ehesten Sinn, wenn er das gesamte XML-Dokument erfasst: dies sollte mit folgendem Code (Listing 21) möglich sein<sup>8</sup>.

---

<sup>8</sup> Leider konnte der Volltextindex nicht mit dieser Syntax – obwohl von zwei Autoren so erklärt – in der Datenbank angelegt werden. Möglicherweise ist die Syntax für die Endversion (DB2 9) gegenüber der Beta (DB2 Viper) doch noch geändert worden.

```
CREATE INDEX idx2 FOR TEXT ON clients(contactinfo) FORMAT XML
```

**Listing 21: Volltextindex auf die XML-Daten erstellen.**

### 3 Mögliche Nutzung von „pureXML“ am Beispiel

Die XML-Features der DB2 werden in der Presse und von IBM selbst mit dem Schlagwort „pureXML“ betitelt. Dies trifft auch durchaus zu, da man alle relevanten XPath und XQuery-Funktionalitäten verwenden kann. Im Folgenden werden zwei Beispielszenarien vorgestellt, in denen der Nutzen einer XML-verarbeitenden Datenbank deutlich zu sehen ist.

Die derzeitige Standardtechnik zur Umsetzung von Service-orientierten Architekturen sind Web Services, die eine lose Kopplung von Diensten in einem heterogenen Umfeld ermöglichen. Die technische Beschreibung eines solchen Dienstes wird dabei in einem XML-Dokument – dem *Web Service Description Language*-Dokument (WSDL) – vorgehalten. Darin werden einzelne Methoden, Input- und Outputparameter und weitere Daten beschrieben. Um einen passenden Web Service zu finden, werden z. B. UDDI-Verzeichnisse (Universal Description, Discovery and Integration) verwendet: Darin sind Web Services, deren Anbieter und Beschreibungen der Dienste hinterlegt. Einige große Unternehmen nutzen solche Verzeichnisse als eine Art „Gelbe Seiten“, um mögliche IT-Dienstleistungen zu finden.

Wir streben an, eine Art UDDI-Verzeichnis von Dienstleistungen zu konzipieren (siehe Abbildung 2), welches (vorerst nur interne) Web Services verwaltet. Natürlich kann hier auf UDDI direkt zurückgegriffen werden; möglicherweise ist aber die Komplexität der UDDI-Spezifikation (aktuell in Version 3.0.2) ein technischer Overkill, der vermieden werden kann. In Abbildung 2 ist ein solches Szenario für den Anwendungsfall einer Universität dargestellt. Hierbei ist eine Datenbank (DB2 9) mit einer Tabelle hinterlegt, die eine ID, eine Beschreibung sowie eine Metadaten- und WSDL-Datei im XML-Format speichern kann. Ein (hier nicht näher spezifiziertes) Web-Frontend lässt die einzelnen Benutzer nach passenden Services suchen. Mit einer Suche werden so leicht alle Daten (eben auch technische Details aus der WSDL-Datei) eines Services angesprochen.

Ein weiteres Szenario ist beispielsweise ein E-Learning-System (wie xLx<sup>9</sup>), bei dem XML-Kenntnisse gefördert werden und gefordert sind. Hier kann man die Validierung der zu speichernden XML-Daten anhand eines XML-Schemas so verwenden, dass gültige Lösungen zu einer Aufgabe direkt gespeichert (und damit auch automatisch korrigiert) werden. Ungültige Lösungen werden erkannt, da sie nicht validiert werden können. Ein möglicher Ansatz ist in Abbildung 3 zu sehen.

---

<sup>9</sup> *eXtreme eLearning eXperience*: E-Learning-System am Institut für Wirtschaftsinformatik (Database Group) der Universität Münster.

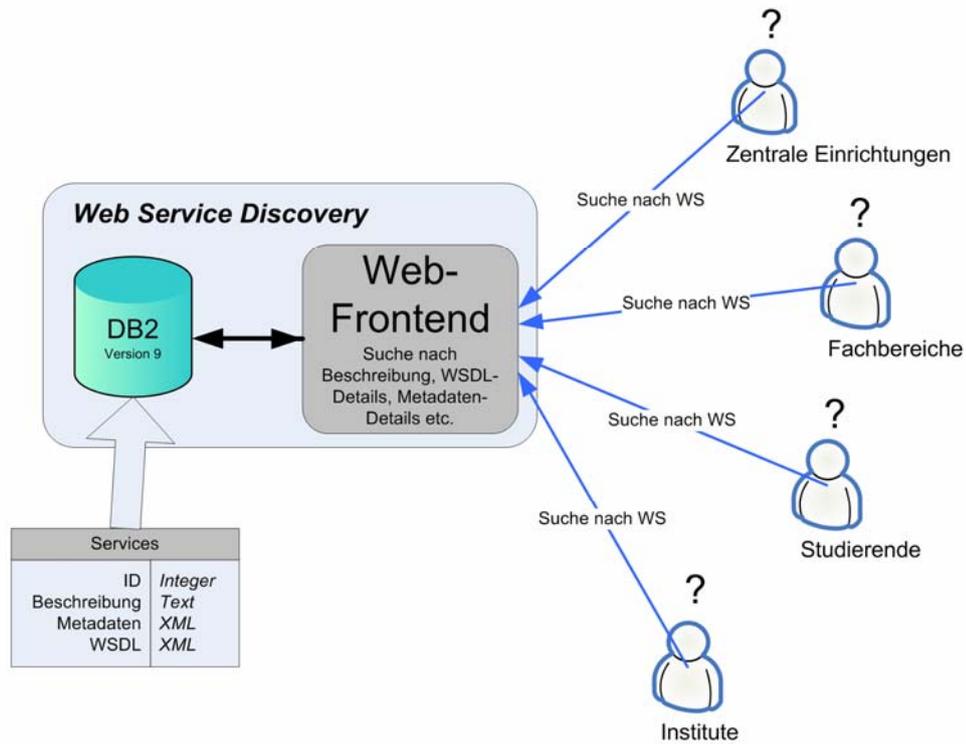


Abbildung 2: Einsatzszenario der DB2 als UDDI-Ersatz.

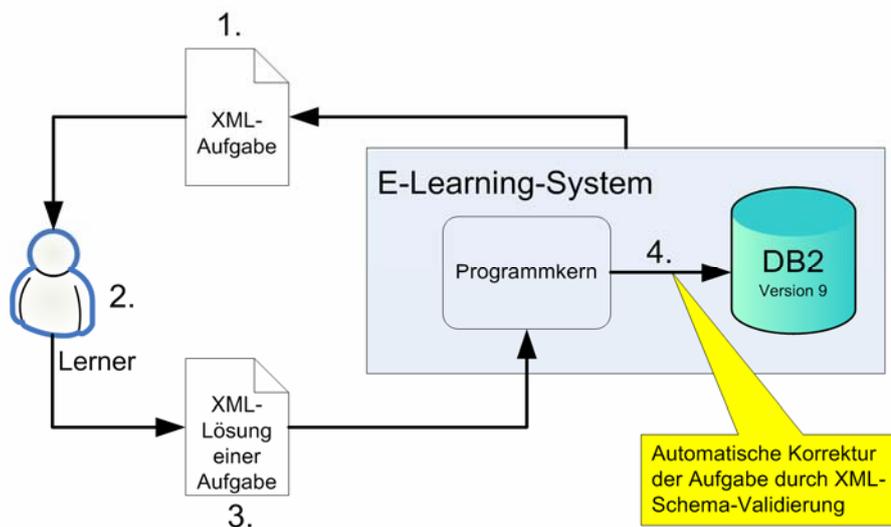


Abbildung 3: Einsatzszenario für automatische Korrektur von XML-Aufgaben.

Im 1. Schritt wird dem Lernenden eine Aufgabe (XML-Aufgabe) gestellt. Daraufhin erfolgt im zweiten Schritt die Bearbeitung der Aufgabe. Die Lösung des Lernenden wird in Teilschritt 3 an das E-Learning-System weitergegeben. Hier wird es zwischengespeichert werden, bis der Bearbeitungszeitraum der Aufgabe abläuft. Schließlich wird die abgegebene Lösung im 4. Schritt in

der DB2 (bspw. in der Tabelle Lösungen) unter Verwendung eines XML-Schemas gespeichert. Lösungen, die nicht validiert werden können, fallen automatisch heraus, da sie nicht gespeichert werden. So hat man die einfachste Form der automatischen Kontrolle.

## 4 Fazit

Der Einsatz neuer Datenbankfeatures für operationale Systeme in Unternehmen läuft meist sehr langsam an.<sup>10</sup> Der alte Wahlspruch der IT: „Never change a running system“ gilt wohl zu Recht auch hier. Es ist sehr schwer, verlässliche Informationen zur Verwendung von „pureXML“-Features in der Industrie und der Wirtschaft zu bekommen, um die Verbreitung dieser Technologie einzuschätzen. Dies liegt wohl auch daran, dass die Unterstützung von nativen XML-Funktionen erst seit kurzem in den aktuellen Releases der Datenbankgrößen zu finden sind. Klar ist, dass neben IBM auch Oracle in der neuesten Version ihrer Datenbank Oracle 10g die native XML-Verarbeitung eingebaut hat. Wahrscheinlich ist es nun nur eine Frage der Zeit, bis die ersten Unternehmen ernsthaft mit der Funktionalität arbeiten werden, die die neuen Versionen der größten kommerziellen Datenbanken bieten.

Die in Kapitel 3 angesprochenen Szenarien zeigen einige wenige Möglichkeiten der neuen XML-Features der DB2-Datenbank und insbesondere für den Bereich der Service Orientierten Architekturen und Web Services sollten diese Erweiterung mit Sicherheit eine Erleichterung für die Verarbeitung von XML-Daten schaffen, die es zu Nutzen gilt.

## Literaturverzeichnis

[CS06] Chamberlin, D.; Saracco, C. M.: *Query DB2 XML data with XQuery*, IBM Developer Works (Website), IBM, April 2006.

[KM03] Klettke, M.; Meyer, H.: *XML & Datenbanken*, dpunkt.verlag, Heidelberg, 2003.

[MI02] Mintert, S.: *XML & Co*, Addison-Wesley, München, 2002.

[NL05] Nicola, M.; van der Linden, B.: *Native XML Support in DB2 Universal Database*, 31st International Conference on Very Large Data Bases 2005 (VLDB2005), Trondheim (Norway), September 2005.

[NO06] North, K.: *XML in DB2 9*, DB2 Magazine, 2006 (3rd Quarter).

[SA06a] Saracco, C. M.: *Query DB2 XML data with SQL*, IBM Developer Works (Website), IBM, März 2006.

[SA06b] Saracco, C. M.: *Get off to a fast start with DB2 Viper*, IBM Developer Works (Website), IBM, March 2006.

[SA06c] Saracco, C. M.: *Opening the door to XML*, DB2 Magazine, 2006 (3rd Quarter).

[SA06] Sanders, R. E.: *Distributed DBA*, DB2 Magazine, 2006 (3rd Quarter).

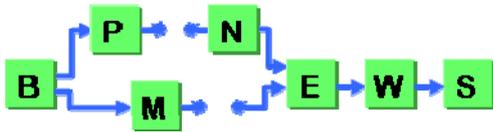
[ST05] Saake, G.; Türker C.: *Objektrelationale Datenbanken*. Ein Lehrbuch, dpunkt Verlag, 2005

[TUE03] Türker, C.: *SQL:1999 & SQL:2003*, dpunkt Verlag, 2003.

[WH06] Whitney, J.: *Eye on XQuery*, DB2 Magazine, 2006 (3rd Quarter).

---

<sup>10</sup> Dies ergaben u.a. Nachfragen bei ehemaligen Institutsmitarbeitern, die mittlerweile in Beratungsunternehmen der IT-Branche arbeiten.



## Business Process Management – Kolumne

### Folge 5

Manfred Reichert (Universität Twente)

Stefanie Rinderle (Universität Ulm)

Barbara Weber (Universität Innsbruck)

Liebe Leser,

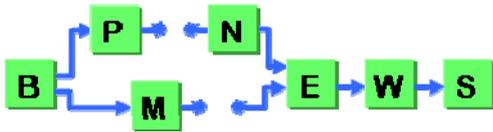
mit der aktuellen Ausgabe des EMISA Forums erhalten Sie Folge 5 unserer Kolumne zum Thema *Business Process Management* (BPM). Diese berichtet über aktuelle Themen, Entwicklungen, wissenschaftliche Arbeiten und Veranstaltungen aus dem BPM-Umfeld.

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 Reichert (m.u.reichert@utwente.nl), Stefanie Rinderle (stefanie.rinderle@uni-ulm.de) oder Barbara Weber (Barbara.Weber@uibk.ac.at).



## Veranstungsbericht

### 2<sup>nd</sup> IEEE Workshop on Agile Cooperative Process-Aware Information Systems (ProGility 2007)

Paris, 19. Juni 2007

Workshop im Rahmen der WETICE'07

Workshop-Bericht von

*Barbara Weber*

Universität Innsbruck

*Manfred Reichert*

University of Twente, NL

*Rik Eshuis*

TU Eindhoven, NL

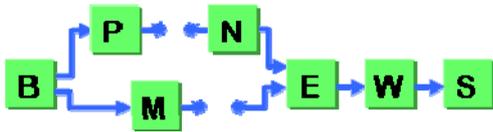
**Summary.** The goal of the ProGility workshop is to bring together practitioners and researchers from different communities who share an interest in agile cooperative process-aware information systems. The vision is to be able to rapidly implement new processes and to quickly adapt existing process implementations to environmental changes. This workshop report gives an overview of the presented papers that address various flexibility issues of process-aware information systems.

**Introduction.** The economic success of an enterprise more and more depends on its ability to react to changes in its environment in a quick and flexible way. Business trends such as increasing product and service variability, quicker time-to-market, and increasing division of labor along a global supply chain of goods and services force enterprises to collaborate with each other in networks that are dynamic, flexible, ad-hoc, and adaptive. Thus enterprises are developing a growing interest in new concepts, systems, and solutions which help them to flexibly align their organizational structures, business processes, and supporting information systems to these new requirements and to optimize interactions with customers and business partners. While there has been major progress in disciplines that are interested in structured and unstructured intra-organizational business processes, the agile enterprise is still a vision. Agility in this context refers to the ability of an enterprise to rapidly set up new business processes and projects in order to quickly adapt to changes in the environment. To support these changes, traditional enterprises have to align their existing information systems while virtual enterprises need to customize and integrate the individual partner processes. To meet its business objectives, the agile enterprise continuously re-aligns its business processes as well as the interactions with its partners and customers to meet the current requirements.

The ProGility workshop is a merger of two previous WETICE workshops on process flexibility (ProFlex) and process integration (PINCET). Building upon the success of these workshops, the goal of the ProGility workshop is to bring together practitioners and researchers from different communities such as BPM, software engineering, service-oriented computing, artificial intelligence, and Computer Supported Cooperative Work (CSCW) who share an interest in flexibility of cooperative process-aware information systems and team support in both an intra- and inter-organizational setting. The workshop aims at discussing the current state of ongoing research and at sharing practical experiences. Submitted papers have been evaluated on the basis of significance, originality and technical quality.

#### Overview of the Papers.

The paper by Bobrik and Bauer presents ongoing work done in the Proviado project on the visualization of large business processes. The presented approach provides for flexible and highly configurable business process visualizations, which can be personalized and adapted to users' needs. In detail, Proviado distinguishes three dimensions for process visualization: First, it allows to reduce complexity by discarding or aggregating process information not relevant in the given context. Second, the notation and graphical appearance of process elements (e.g., activities, data objects, control connectors) is customizable. Third, different presentation forms (e.g., process graph, swim lane, calendar, table) are supported. Regarding the first dimension, Proviado provides a powerful and customizable view mechanism. The supported view



concept allows to flexibly adapt the presented process information through customizable graph aggregation and graph reduction techniques.

The paper by Minor, Schmalen, Koldehoff and Bergmann addresses run-time process flexibility. In the proposed approach flexibility is achieved through structural adaptations of process instances, late planning, and breakpoint mechanisms. Structural adaptations of running process instances (e.g., to add or delete process steps) are supported by an authoring component. This component also enables the case-based retrieval of similar past workflows. Similarity calculation is based on the structure and the status of the ongoing process instance as well as on its context. The late modeling feature allows to deal with uncertainty by leaving parts of the model unspecified during build-time and by deferring the decision how the respective parts should look like to the run-time. Finally, the suggested breakpoint mechanism ensures that parts of a process can be further executed while other parts have to be blocked in order to decide about a change.

Müller, Reichert, Herbst and Poppa present a tool which addresses one of the biggest challenges for cooperative processes in the engineering domain, namely the modeling, coordination and maintenance of the many interdependencies existing between the different sub-processes in engineering. For dealing with this challenge, the COREPRO project follows a model-driven approach. It allows to semi-automatically derive process structures from large data structures. Basic to this is the observation that in the engineering domain respective process structures are characterized by a strong relationship with the assembly of the product; i.e., the sub-processes to be coordinated can be related to the different product components. COREPRO utilizes this information about the product and its structure for deriving, coordinating and maintaining such data-driven process structures. This will be of particular benefit if hundreds up to thousands of sub-processes have to be coordinated (as it is the case, for example, in the automotive domain).

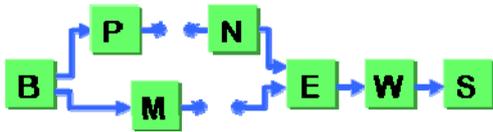
Mehandjiev, Stalker and Carpenter present an approach which aims at more flexible supply chains. In particular, companies and large organizations should be enabled to build *Instant Virtual Enterprises* in order to react to emerging market opportunities. For this, the authors suggest an approach, which is based on coordination theory and virtual organizations. By systematically delaying the recruitment of partners more flexible supply chains can be built.

Vonk, Wang and Grefen aim at more agility by facilitating the exchange of services or partners. The authors argue that bridging the gap between business-level SLAs and IT-level SLAs is fundamental. The paper focuses on transactional QoS and suggests a mapping between process and services.

Finally, the paper of Ziemann, Werth, Matheis and Kahl present ongoing work towards tool support for integrated modeling of private and public business processes. The paper presents a view concept for modeling private and public processes. The tool aims at supporting the transformation of public to private processes as well as the transformation of business level representations to executable process specifications (i.e., technical level).

**Conclusion and Future Research Topics.** The different papers illustrate various flexibility needs of process-aware information systems (PAIS) from different domains, and they make evident that any PAIS will not be accepted by its users if rigidity or inflexibility comes with it. Flexibility issues do not only arise at the level of intra-organizational business processes, but increasingly emerge at the level of business collaborations (i.e., inter-organizational business processes) as well. Many different aspects have to be considered when realizing flexible process support including the correctness of (dynamic) process changes, ease of change, cost of change, and security.

Often, tool vendors promise flexible technologies and paradigms for realizing process-aware information systems, but are often unable to cope with fundamental issues related to process flexibility and process change (e.g., correctness and robustness). The existence of different process support paradigms and the lack of methods for comparing existing change approaches makes it even more challenging for PAIS engineers to



choose the adequate technology. This necessitates a framework which fosters the systematic comparison of existing process management technologies with respect to their flexibility and change support. The recent emergence of change as well as exception handling patterns constitute an important step towards such a comparison framework. Furthermore, more experimental and empirical research is needed in future.

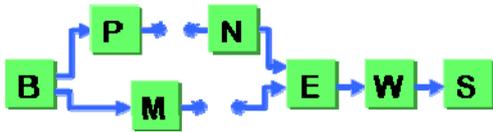
Though considerable progress has been achieved with respect to the realization of flexible PAIS, a lot of challenges remain. While there exist many approaches and paradigms for the flexible support of intra-organizational processes, only little flexibility support has been provided for inter-organizational processes so far. One major challenge existing in this context concerns the evolution of process choreographies, i.e., the change of interactions and message exchanges between the partner processes in a cross-organizational setting. If respective modifications are conducted in an uncontrolled manner, severe inconsistencies or errors might occur in the sequel. In particular, modifications of private processes performed by a single party may indirectly affect the implementation of the private processes of partners as well. What is needed is a powerful change framework for choreographies that allows process engineers to reason about changes of private processes and their effects on public processes as well as on partner processes.

### Abgeschlossene Dissertationen

- Jan Mendling: On the detection and prediction of errors in EPC business process models. Wirtschaftsuniversität Wien, 2007

Eine Zusammenfassung der Dissertationsschrift von Jan Mendling findet sich in dieser Kolumne.

BPM Veranstaltungskalender	
<b>September 2007</b>	
03.09. – 07.09.	DEXA 2007 - 18th International Conference on Database and Expert Systems Applications, Regensburg [ <a href="http://www.dexa.org/">http://www.dexa.org/</a> ]  Begleitende Konferenz: - EC-Web 2007 - 8th International Conference on Electronic Commerce and Web Technologies [ <a href="http://www.dexa.org/drupal/?q=ecweb">http://www.dexa.org/drupal/?q=ecweb</a> ]
17.09. – 20.09.	ICSOC 2007 - International Conference on Service-Oriented Computing, Wien, Österreich [ <a href="http://www.icsoc.org/">http://www.icsoc.org/</a> ]
23.09 – 28.09.	VLDB 2007 - 33 <sup>rd</sup> International Conference on Very Large Data Bases, Wien, Österreich [ <a href="http://www.vldb2007.org/">http://www.vldb2007.org/</a> ]
24.09. – 27.09.	BPM 2007 - 5th International Conference on Business Process Management, Brisbane, Australia (ein ausführliches Programm findet sich im letzten Teil dieses Heftes)  [ <a href="http://bpm07.fit.qut.edu.au/">http://bpm07.fit.qut.edu.au/</a> ] Begleitende Workshops: - BPI 2007 - 3rd International Workshop on Business Process Intelligence [ <a href="http://is.tm.tue.nl/bpi07/">http://is.tm.tue.nl/bpi07/</a> ] - BPD 2007 - 3rd International Workshop on Business Process Design [ <a href="http://is.tm.tue.nl/staff/hreijers/BPD07/">http://is.tm.tue.nl/staff/hreijers/BPD07/</a> ] - CBP 2007 - 1st International Workshop on Collaborative Business Processes [ <a href="http://www.ict.swin.edu.au/conferences/cbp2007">http://www.ict.swin.edu.au/conferences/cbp2007</a> ] - GPWW - 3 <sup>rd</sup> International Workshop on Grid and Peer-to-Peer Based Workflows [ <a href="http://www.ict.swin.edu.au/conferences/gpww/">http://www.ict.swin.edu.au/conferences/gpww/</a> ] - ProHealth - 1st International Workshop on Process-oriented Systems in Healthcare [ <a href="http://wwwhome.cs.utwente.nl/~reichertm/ProHealth07.htm">http://wwwhome.cs.utwente.nl/~reichertm/ProHealth07.htm</a> ]



	<ul style="list-style-type: none"> <li>- <i>ERCIS 2007 - 10th International Workshop on Reference Modeling</i> [<a href="http://www.ercis.de/ERCIS/news/10threfmod.html">http://www.ercis.de/ERCIS/news/10threfmod.html</a>]</li> <li>- <i>semantics4ws - 2nd International Workshop on Advances in Semantics for Web Services</i> [<a href="http://events.deri.at/semantics4ws2007/">http://events.deri.at/semantics4ws2007/</a>]</li> </ul>
<b>Oktober 2007</b>	
15.10. – 19.10.	<i>EDOC 2007 - 11th IEEE International EDOC Conference</i> , Annapolis, Maryland, USA [ <a href="http://edoc.mitre.org/">http://edoc.mitre.org/</a> ]
25.10. – 26.10.	<i>Geschäftsprozesse explorieren, modellieren und optimieren</i> , Seminar der Deutschen Informatik Akademie (DIA), Mannheim [ <a href="http://www.dia-bonn.de/gpo1_2007/gpo_main.html">http://www.dia-bonn.de/gpo1_2007/gpo_main.html</a> ]
<b>November 2007</b>	
05.11. – 09.11.	<i>ER 2007 – 27<sup>th</sup> International Conference on Conceptual Modeling</i> , Auckland, New Zealand [ <a href="http://er2007.massey.ac.nz/">http://er2007.massey.ac.nz/</a> ]
22.11. – 23.11.	<i>Service-orientierte Implementierung von Prozessen: Paradigmen, Konzepte, Technologien</i> Seminar, Deutsche Informatik Akademie (DIA), Heidelberg [ <a href="http://www.dia-bonn.de">http://www.dia-bonn.de</a> ]
25.11. – 30.11.	<p><i>OTM 2007 – On The Move Federated Conferences and Workshops</i>, Vilamoura, Algarve, Portugal [<a href="http://www.cs.rmit.edu.au/fedconf/index.html?page=home">http://www.cs.rmit.edu.au/fedconf/index.html?page=home</a>]</p> <p>Konferenzen:</p> <ul style="list-style-type: none"> <li>- <i>CoopIS 2007 - 15th International Conference on Cooperative Information Systems</i> [<a href="http://www.cs.rmit.edu.au/fedconf/index.html?page=coopis2007cfp">http://www.cs.rmit.edu.au/fedconf/index.html?page=coopis2007cfp</a>]</li> <li>- <i>DOA 2007 - The 9th International Symposium on Distributed Objects, Middleware, and Applications (DOA)</i> [<a href="http://www.cs.rmit.edu.au/fedconf">http://www.cs.rmit.edu.au/fedconf</a>]</li> <li>- <i>ODBASE 2007 - The 6th International Conference on Ontologies, DataBases, and Applications of Semantics</i> [<a href="http://www.cs.rmit.edu.au/fedconf">http://www.cs.rmit.edu.au/fedconf</a>]</li> <li>- <i>GADA 2007 - International Conference on Grid computing, high-performance and Distributed Applications</i> [<a href="http://www.cs.rmit.edu.au/fedconf/gada2007cfp.html">http://www.cs.rmit.edu.au/fedconf/gada2007cfp.html</a>]</li> <li>- <i>IS 2007 - The 2nd International Symposium on Information Security</i> [<a href="http://www.cs.rmit.edu.au/fedconf">http://www.cs.rmit.edu.au/fedconf</a>]</li> </ul>
<b>Dezember 2007</b>	
03.12. – 07.12.	<p><i>WISE 2007 – The 8th International Conference on Web Information Systems Engineering</i>, Nancy, Frankreich [<a href="http://wise2007.loria.fr/pmwiki/pmwiki.php">http://wise2007.loria.fr/pmwiki/pmwiki.php</a>]</p> <p>Begleitende Workshops (Auswahl):</p> <ul style="list-style-type: none"> <li>- <i>Coordination of Inter-Organization Workflows: Agent and Semantic Web Based Models</i></li> <li>- <i>Governance, Risk and Compliance in Web Information Systems</i></li> <li>- <i>Human-friendly (Web) Service Description, Discovery and Matchmaking</i></li> </ul>

# Aktuelles Schlagwort: Process Change Patterns

Barbara Weber<sup>1</sup>, Stefanie Rinderle<sup>2</sup>, and Manfred Reichert<sup>3</sup>

<sup>1</sup>Quality Engineering Research Group, University of Innsbruck, Austria  
Barbara.Weber@uibk.ac.at

<sup>2</sup>Inst. Databases and Information Systems, Ulm University, Germany  
stefanie.rinderle@uni-ulm.de

<sup>3</sup>Information Systems Group, University of Twente, The Netherlands  
m.u.reichert@cs.utwente.nl

## 1 Introduction

Process-aware information systems (PAIS) allow for process changes at different levels. In general, respective changes must be enabled at the process type as well as the process instance level [1]. While the former become necessary to deal with the evolving nature of real-world processes (e.g., to adapt the process to legal changes), the latter are needed to cope with exceptional situations. Vendors often promise flexible solutions for realizing adaptive PAIS, but force users to realize process adaptations at a rather low level of abstraction. Apart from this, several competing paradigms exist, all trying to tackle the need for process flexibility (e.g., adaptive process management vs. case handling). So far, there has been no method for systematically comparing the different change frameworks, making it difficult for PAIS engineers to choose the right technology.

We have studied a variety of processes and process change scenarios from different domains (e.g. [2,3]). We have further elaborated the change support features provided by existing tools. Taking these experiences, we have designed a set of *process changes patterns* to foster the comparison of existing approaches with respect to process change definition. More precisely, change patterns allow for high-level process adaptations at both the process type and the process instance level. In this paper we sketch 17 characteristic patterns we identified as relevant for *control flow changes* (cf. Fig. 1). Adaptations of other process aspects (e.g., data or resources) are outside the scope of this paper. Change patterns reduce the complexity of process change (like design patterns in software engineering reduce system complexity) and raise the level for expressing changes by providing abstractions which are above the level of single node and edge operations.

As illustrated in Fig. 1, we divide change patterns into *adaptation patterns* and *patterns for predefined changes*. Adaptation patterns allow modifying a process schema based on high-level change operations. Generally, adaptation patterns can be applied to the whole process schema; i.e., the region to which they are applied may be chosen dynamically. By contrast, for predefined changes,

at build-time, the process engineer defines regions in the process schema where potential changes may be performed during run-time.

For each pattern we provide a name, a brief description, an illustrating example, a description of the problem it addresses, a couple of design choices, remarks regarding its implementation, and a reference to related patterns. *Design Choices* allow for parametrization of patterns keeping the number of distinct patterns manageable. Design choices which are not only relevant for particular patterns, but for a whole pattern category, are described only once at the category level. Typically, existing approaches only support a subset of the design choices in the context of a particular pattern. We denote the combination of design choices supported by a particular approach as a *pattern variant*.

CHANGE PATTERNS			
ADAPTATION PATTERNS (AP)			
Pattern Name	Scope	Pattern Name	Scope
AP1: Insert Process Fragment <sup>(*)</sup>	I / T	AP8: Embed Process Fragment in Loop	I / T
AP2: Delete Process Fragment	I / T	AP9: Parallelize Process Fragment	I / T
AP3: Move Process Fragment	I / T	AP10: Embed Process Fragment in Conditional Branch	I / T
AP4: Replace Process Fragment	I / T	AP11: Add Control Dependency	I / T
AP5: Swap Process Fragment	I / T	AP12: Remove Control Dependency	I / T
AP6: Extract Sub Process	I / T	AP13: Update Condition	I / T
AP7: Inline Sub Process	I / T		
PATTERNS FOR PREDEFINED CHANGES (PP)			
Pattern Name	Scope	Pattern Name	Scope
PP1: Late Selection of Process Fragments	I / T	PP3: Late Composition of Process Fragments	I / T
PP2: Late Modeling of Process Fragments	I / T	PP4: Multi-Instance Activity	I / T

I... Instance Level, T ... Type Level

<sup>(\*)</sup>A process fragment can either be an atomic activity, an encapsulated sub process or a process (sub) graph

Fig. 1. Change Patterns Overview

## 2 Adaptation Patterns

Adaptation patterns allow to structurally change process schemes. Examples include the insertion, deletion and re-ordering of activities (cf. Fig. 1). Fig. 2 describes general design choices valid for all adaptation patterns. First, each adaptation pattern can be applied at the process type or process instance level. Second, adaptation patterns can operate on an atomic activity, an encapsulated sub process or a process (sub-)graph (cf. Fig. 2). We abstract from this distinction and use the generic concept *process fragment* instead. Third, the effects resulting from the use of adaptation patterns at the instance level can be permanent or temporary. A *permanent instance change* remains valid until instance completion (unless it is undone by a user). By contrast, a *temporary instance change* is only valid for a certain period of time (e.g. one loop iteration).

We describe four selected adaptation patterns in more detail. These four patterns allow for the insertion, deletion, movement, and replacement of process fragments in a given process schema. The *Insert Process Fragment* pattern (cf. Fig. 3a) can be used to add process fragments to a process schema. In addition to

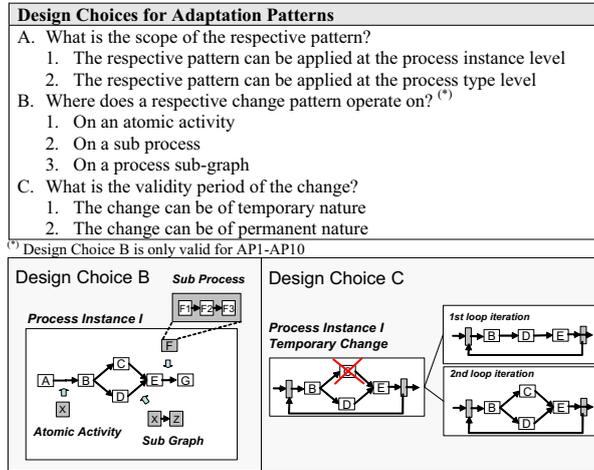


Fig. 2. Design Choices for Adaptation Patterns

the general options described in Fig. 2, one major design choice for this pattern (Design Choice D) describes the way the new process fragment is embedded in the respective schema. There are systems which only allow to serially insert a fragment between two directly succeeding activities. By contrast, other systems follow a more general approach allowing the user to insert new fragments between two arbitrary sets of activities. Special cases of the latter variant include the insertion of a fragment in parallel to another one or the association of the newly added fragment with an execution condition (*conditional insert*). The *Delete Process Fragment* pattern, in turn, can be used to remove a process fragment (cf. Fig 3b). No additional design choices exist for this pattern. Fig. 3b depicts alternative ways in which this pattern can be implemented.

The *Move Process Fragment* pattern (cf. Fig. 4a) allows to shift a process fragment from its current position to another one. Like for the *Insert Process Fragment* pattern, an additional design choice specifies the way the fragment can be embedded in the process schema afterwards. Though the *Move* pattern could be realized by the combined use of patterns AP1 and AP2, we introduce it as separate pattern as it provides a higher level of abstraction to users. The latter also applies when a fragment has to be replaced by another one. This change is captured by the *Replace Process Fragment* pattern (cf. Fig. 4b).

We have only described the most relevant adaptation patterns. Additional patterns we identified are: swapping of activities (AP5), extraction of a sub process from a process schema (AP6), inclusion of a sub process into a process schema (AP7), embedding of an existing process fragment in a loop (AP8), parallelization of process fragments (AP9), embedding of a process fragment in a conditional branch (AP10), addition of control dependencies (AP11), removal of control dependencies (AP12), and update of transition conditions (AP13). A detailed description of these patterns can be found in [4, 5].

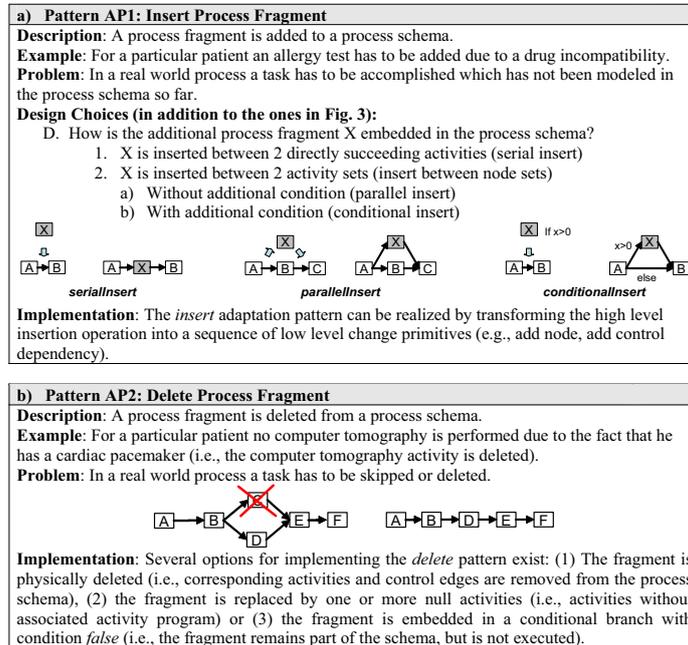


Fig. 3. Insert (AP1) and Delete (AP2) Process Fragment patterns

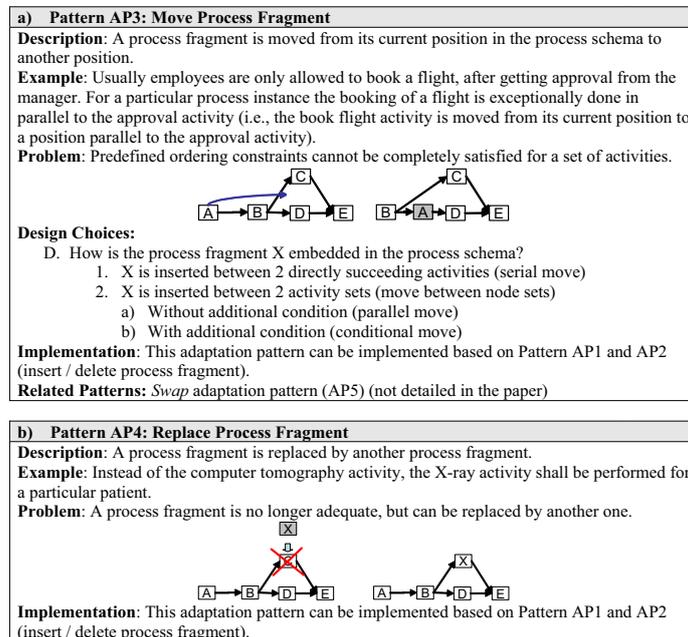


Fig. 4. Move (AP3) and Replace (AP4) Process Fragment patterns

### 3 Patterns for Predefined Changes

The applicability of adaptation patterns is not restricted to a particular process part a priori. By contrast, the following patterns predefine constraints concerning the parts that can be changed; i.e., at run-time changes are only permitted within these parts. In this category we have identified 4 patterns, *Late Selection* (PP1), *Late Modeling* (PP2), *Late Composition* (PP3) and *Multi-Instance Activity* (PP4) (cf. Fig. 5). The *Late Selection* pattern (cf. Fig. 6) allows to select the implementation for a particular process step at run-time either based on predefined rules or user decisions. The *Late Modeling* pattern (cf. Fig. 7a) offers more freedom and allows to model selected parts of the process schema at run-time. The *Late Composition* pattern (cf. Fig. 7b) enables the on-the fly composition of process fragments (e.g., by dynamically introducing control dependencies between them). Finally, the *Multi-Instance Activity* pattern allows to determine the number of instances created for a particular process activity at run-time.

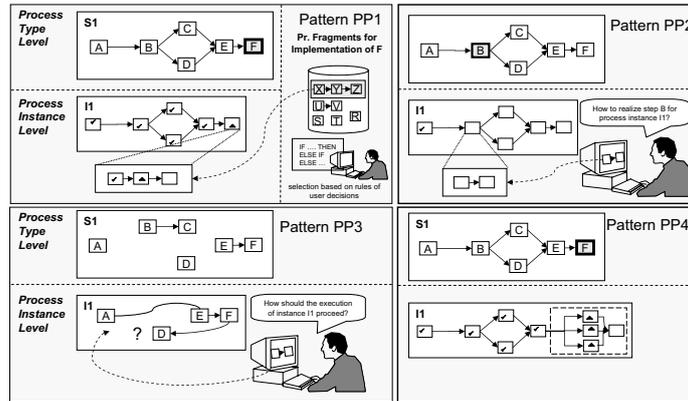


Fig. 5. Patterns for Predefined Changes (Overview)

### 4 Related Work

Patterns were first used to describe solutions to recurring problems by Alexander, who applied them to describe best practices in architecture [6]. Patterns also have a long tradition in computer science. Gamma et al. applied the same concepts to software engineering [7].

In the area of workflow management, patterns have been introduced for analyzing the expressiveness of process modeling languages (i.e., control flow patterns [8]). In addition, data patterns [9] describe different ways for modeling the data aspect in PAIS. The introduction of these patterns has had significant impact on the design of PAIS and has contributed to their systematic evaluation. However, to evaluate the powerfulness of a PAIS regarding its ability to deal with changes, the existing patterns are important, but not sufficient. In addition, a set of patterns for the aspect of workflow change is needed. Further, the

<p><b>Pattern PP1: Late Selection of Process Fragments</b></p> <p><b>Description:</b> For particular activities the corresponding implementation (activity program or sub process model) can be selected during run-time. At build time only a placeholder is provided, which is substituted by a concrete implementation during run-time (cf. Fig. 6).</p> <p><b>Example:</b> For the treatment of a particular patient one of several different sub-processes can be selected depending on the patient's disease.</p> <p><b>Problem:</b> There exist different implementations for an activity (including sub-processes), but for the selection of the respective implementation run-time information is required.</p> <p><b>Design Choices:</b></p> <ul style="list-style-type: none"> <li>A. How is the selection process done? <ul style="list-style-type: none"> <li>1. Automatically based on predefined rules</li> <li>2. Manually by an authorized user</li> <li>3. Semi-automatically: options are reduced by applying some predefined rules; user can select among the remaining options</li> </ul> </li> <li>B. What object can be selected? <ul style="list-style-type: none"> <li>1. Atomic activity</li> <li>2. Sub process</li> </ul> </li> <li>C. When does late selection take place? <ul style="list-style-type: none"> <li>1. Before the placeholder activity is enabled</li> <li>2. When enabling the placeholder activity</li> </ul> </li> </ul> <p><b>Implementation:</b> By selecting the respective sub process or activity program, a reference to it is dynamically set and the selected sub-process or activity program is invoked.</p> <p><b>Related Patterns:</b> Prerequisite for Pattern <i>Late Modeling of Process Fragment (PP2)</i></p>
---

**Fig. 6.** Late Selection of Process Fragments (PP1)

degree to which control flow patterns are supported provides an indication of how complex the change framework under evaluation is.

In [10] exception handling patterns are proposed. In contrast to change patterns, exception handling patterns like *Rollback* only change the state of a process instance (i.e., its behavior), but not its schema. The patterns described in this paper do not only change the observable behavior of a process instance, but additionally adapt the process structure. For a complete evaluation of flexibility, both change patterns and exception handling patterns must be considered.

## 5 Summary and Outlook

We designed 17 change patterns which allow to assess the power of a particular change definition framework. Future work will include change patterns for aspects other than control flow (e.g., data or resources) and patterns for more advanced adaptation policies (e.g., the accompanying adaptation of the data flow when introducing control flow changes).

## References

1. Rinderle, S., Reichert, M., Dadam, P.: Correctness criteria for dynamic changes in workflow systems – a survey. *Data and Knowledge Engineering* **50** (2004) 9–34
2. Lenz, R., Reichert, M.: IT support for healthcare processes - premises, challenges, perspectives. *Data and Knowledge Engineering* (2007) 39–58
3. Mueller, D., Herbst, J., Hammori, M., Reichert, M.: IT support for release management processes in the automotive industry. In: *Proc. 4th Int'l Conf. Business Process Management (BPM'06)*, Vienna (2006) 368–377

<p><b>a) Pattern PP2: Late Modeling of Process Fragments</b></p> <p><b>Description:</b> Parts of the process schema have not been defined at build-time, but are modeled during run-time for each process instance (cf. Fig. 6). For this purpose, placeholder activities are provided, which are modeled and executed during run-time. The modeling of the placeholder activity must be completed before the modeled process fragment can be executed.</p> <p><b>Example:</b> The exact treatment process of a particular patient is composed out of existing process fragments at run-time.</p> <p><b>Problem:</b> Not all parts of the process schema can be completely specified at build time.</p> <p><b>Design Choices:</b></p> <p>A. What are the basic building blocks for late modeling?</p> <ol style="list-style-type: none"> <li>1. All process fragments (including activities) from the repository can be chosen</li> <li>2. A constraint-based subset of the process fragments from the repository can be chosen</li> <li>3. New activities or process fragments can be defined</li> </ol> <p>B. What is the degree of freedom regarding late modeling?</p> <ol style="list-style-type: none"> <li>1. Same modeling constructs and change patterns can be applied as for modeling at the process type level <sup>(*)</sup></li> <li>2. More restrictions apply for late modeling than for modeling at the process type level</li> </ol> <p>C. When does late modeling take place?</p> <ol style="list-style-type: none"> <li>1. When a new process instance is created</li> <li>2. When the placeholder activity is instantiated</li> <li>3. When a particular state in the process is reached (which must precede the instantiation of the placeholder activity)</li> </ol> <p>D. Does the modeling start from scratch?</p> <ol style="list-style-type: none"> <li>1. Late modeling may start with an empty template</li> <li>2. Late modeling may start with a predefined template which can then be adapted</li> </ol> <p><b>Implementation:</b> After having modeled the placeholder activity with the editor, the fragment is stored in the repository and deployed. Finally, the process fragment is dynamically invoked as an encapsulated sub-process. The assignment of the respective process fragment to the placeholder activity is done through late binding.</p> <p><b>Related Patterns:</b> necessitates <i>Late Selection of Process Fragments</i> (PP1) of the dynamically modified fragment</p> <p><sup>(*)</sup> Which of the adaptation patterns are supported within the placeholder activity is determined by the expressiveness of the used modeling language.</p>
<p><b>b) Pattern PP3: Late Composition of Process Fragments</b></p> <p><b>Description:</b> At build time a set of process fragments is defined out of which a concrete process instance can be composed at run time. This can be achieved by dynamically selecting fragments and adding control dependencies on the fly (cf. Fig. 6).</p> <p><b>Example:</b> Several medical examinations can be applied for a particular patient. The exact examinations and the order in which they are performed are defined for each patient individually.</p> <p><b>Problem:</b> There exist several variants of how process fragments can be composed. In order to reduce the number of process variants to be specified by the process engineer during build time, process instances are dynamically composed out of fragments.</p>

**Fig. 7.** Late Modeling (PP2) and Late Composition of Process Fragments (PP3)

4. Weber, B., Rinderle, S., Reichert, M.: Change patterns and change support features in process-aware information systems. In: Proc. 19th Int'l Conf. on Advanced Information Systems Engineering (CAiSE'07). (2007) 574–588
5. Weber, B., Rinderle, S., Reichert, M.: Identifying and evaluating change patterns and change support features in process-aware information systems. Technical Report Report No. TR-CTIT-07-22, CTIT, Univ. of Twente, The Netherlands (2007)
6. Alexander, C., Ishikawa, S., Silverstein, M.: A Pattern Language. Oxford University Press, New York (1977)
7. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley (1995)
8. van der Aalst, W.M.P., ter Hofstede, A.H.M., Kiepuszewski, B., Barros, A.P.: Workflow patterns. Distributed and Parallel Databases **14** (2003) 5–51
9. Russell, N., ter Hofstede, A., Edmond, D., van der Aalst, W.: Workflow data patterns. Technical Report FIT-TR-2004-01, Queensland Univ. of Techn. (2004)
10. Russell, N., van der Aalst, W.M., ter Hofstede, A.H.: Exception handling patterns in process-aware information systems. In: CAiSE'06. (2006)

# On the detection and prediction of errors in EPC business process models - extended abstract -

Jan Mendling<sup>1,2</sup>

<sup>1</sup>Faculty of Information Technology  
Queensland University of Technology  
Level 5, 126 Margaret Street, Brisbane QLD 4000, Australia  
[j.mendling@qut.edu.au](mailto:j.mendling@qut.edu.au)

<sup>2</sup> Institute of Information Systems and New Media  
Vienna University of Economics and Business Administration  
Augasse 2-6, A-1090 Wien, Austria  
[jan.mendling@wu-wien.ac.at](mailto:jan.mendling@wu-wien.ac.at)

**Abstract:** This paper gives an extended abstract of my doctoral thesis on detection and prediction of errors in EPC business process models (see [Men07]).<sup>1</sup> This research is motivated by the observation that business process models are valuable design artifacts in the business process management life cycle, and therefore, they are subject to quality considerations. In this context, the absence of formal errors, such as deadlocks, is of paramount importance for the subsequent implementation of the process. This doctoral thesis provides a fourfold contribution to the understanding of such errors in business process models with a particular focus on Event-driven Process Chains (EPCs), a business process modeling language that is frequently used in practice. Firstly, we formalize the operational semantics of EPCs in a novel way so that matching OR-splits and OR-joins never deadlock. Secondly, and based on these semantics, we introduce a soundness criterion for EPCs that offers a precise identification of those models which have errors. For the verification of this soundness notion in practice, we present two analysis tools, a ProM plug-in for a verification based on the reachability graph, and a batch program called *xoEPC* for a verification based on reduction rules. Thirdly, we define a set of business process model metrics that are supposed to serve as predictors for error probability of an individual EPC. Fourthly, we use statistical methods and a sample of about 2000 EPCs from practice to derive a regression function for the prediction of error probability. This function is validated against a holdout set of 113 EPCs from textbooks showing that 90% of the EPCs could be classified correctly as having errors or not. These results emphasize the importance of quality issues in business process modeling and provides the foundations for innovations in tool support.

---

<sup>1</sup>A PDF of the thesis can be downloaded at  
<http://wi.wu-wien.ac.at/home/mendling/publications/Mendling%20Doctoral%20thesis.pdf>

## 1 Introduction

It is a fundamental insight of software engineering that design errors should be detected as early as possible (see e.g. [Boe81, Moo05]). The later errors are detected, the more rework has to be done, and the more design effort has been at least partially useless. This also holds for the consecutive steps of analysis, design, and implementation in the business process management life cycle (cf. e.g. [Ros06]). In the design phase, process models are typically created with semi-formal business process modeling languages while formal executable models are needed for the implementation. This problem is often referred to as the gap between business process design and implementation phase (see e.g. [MR04]). Therefore, quality frameworks, such as the Guidelines of Process Modeling [BRU00], stress correctness as the most important quality attribute of business process models.

In the following sections we briefly sketch the results of this doctoral thesis. In Section 2 we present the research contribution of the thesis. Section 3 discusses the findings and its implications for business process modeling. Finally, Section 4 gives an outlook on future research.

## 2 Research Contribution

The research objective of this doctoral thesis is the development of a framework for the detection of formal errors in business process models, and the prediction of error probability based on quality attributes of these models. We will focus on Event-driven Process Chains (EPCs), a business process modeling language that is heavily used in practice. The advantage of this focus is, firstly, that the results of this thesis are likely to have an impact on current modeling practices. Secondly, there is a large empirical basis for analysis. By tapping the extensive stock of EPC model collections, we aim to bring forth general insights into the connection between process model attributes and error probability. In order to validate such a connection, we first need to establish an understanding of model attributes that are likely connected with error probability. Furthermore, we must formally define an appropriate notion of correctness, which gives an answer to the question whether a model has a formal error or not. It is a prerequisite to answering this question that we define the operational semantics of the process modeling language, i.e. EPCs, in a formal way. Against the state of the art, this thesis provides the following technical contributions.

**Formalization of the OR-join:** The semantics of the OR-join have been debated for more than 10 years now. Existing formalizations suffer either from a restriction of the EPC syntax (see e.g. [CS94, LA94, LSW98, Aal99, DR01]) or from non-intuitive behavior (see e.g. [NR02, Kin06, AH05, WEAH05]). In Chapter 2 of this thesis we formalize the EPC semantics concept as proposed in [MA06]. In comparison to other approaches, this novel formalization has the advantage that it is not restricted to a subset of EPCs, and that it provides intuitive semantics for blocks of matching OR-splits and joins since they cannot deadlock. The calculation of the reachability graph was implemented as a plug-in for ProM as a proof of concept. In this way,

this novel semantics definition contributes to research on the specification of business process modeling languages.

**Verification of process models with OR-joins and multiple start and end events:** Verification techniques for process models with OR-joins and multiple start and end events suffer from one of two problems. Firstly, they build on an approximation of the actual behavior and, therefore, do not provide a precise answer to the verification problem, e.g. by considering a relaxed notion of soundness [DR01], by involving user decisions [DAV05], or by approximating relaxed soundness with invariants [VA06]. Secondly, there are verification approaches for semantics definitions (see [CFK05, WAHE06]) that suffer from the previously mentioned non-intuitive behavior. While this is not a problem of the verification itself, all these approaches are not tailored to cope with multiple start and end events. In Chapter 4 of this thesis, we specify a dedicated soundness criterion for EPC business process models with OR-joins and multiple start and end events. Furthermore, we define two verification approaches for EPC soundness, one as an explicit analysis of the reachability graph, and a second based on reduction rules to provide a better verification performance. Both approaches were implemented as a proof of concept. In this way, we contribute to the verification of process models with OR-joins and multiple start and end events, and in particular, we extend the set of reduction rules for business process models.

**Metrics for business process models:** Metrics play an important role in the operationalization of various quality-related aspects in software engineering, network analysis, and business process modeling. Several authors use metrics to capture different aspects of business process models that are presumably related to quality (see e.g. [LY92, Nis98, Mor99, RV04, Car05, BG05, CGP<sup>+</sup>05, CMNR06, LG06, ARG06, MMN<sup>+</sup>06a, MMN<sup>+</sup>06b]). A problem of these works is that business process-specific concepts like sequentiality, decision points, concurrency, or repetition are hardly considered, and too often simple count metrics are defined. Furthermore, there appears to be little awareness of related research, maybe because process model measurement is conducted in separate disciplines including software process management, network analysis, Petri nets theory, and conceptual modeling. In Chapter 5 of this thesis, we will provide an extensive list of metrics for business process models and relate it to previously isolated research. Beyond that, we provide a detailed discussion of the rationale and the limitations of each metric, which is meant to serve as a predictor for error probability. We formulate a hypothesis for each metric based on whether it is positively or negatively correlated with error probability.

**Validation of metrics as error predictors:** Up to now, there is little empirical evidence for the validity of business process model metrics as predictors for error probability. Some empirical work was conducted, but with a different focus. *Lee and Yoon* investigate the empirical relationship between parameters of Petri nets and their state space [LY92]. *Canfora et al.* empirically evaluate the suitability of metrics to serve as predictors for maintainability of the process model [CGP<sup>+</sup>05]. *Cardoso* analyzes the correlation between the control flow complexity metric with the perceived

complexity of process models [Car06]. Most related to this thesis is an analysis of the SAP Reference Model where *Mendling et al.* test a set of simple count metrics as error predictors [MMN<sup>+</sup>06a, MMN<sup>+</sup>06b]. In Chapter 6 of this thesis, we use logistic regression for the test which is similar to the analysis of the SAP Reference Model. Still, we consider both the broader set of metrics from Chapter 5, a precise notion of EPC soundness as defined in Chapter 4, and a much broader sample of EPC models from practice. The results do not only show that certain metrics are indeed a good predictor for error probability, but also that simple count metrics fail to capture important aspects of a process model.

Little research in information systems tries to combine design science and behavioral science research paradigms (see e.g. [BH05]). Since the previously listed contributions cover both design and behavioral aspects, we consider the main contribution of this thesis to be the innovative and holistic combination of both these research paradigms in order to deliver a deeper understanding of errors in business process modeling.

### 3 Discussion

The results of this thesis have some more general implications for business process modeling. On the following pages we discuss the implications for 1) the importance of verification in practice, 2) for improvements of the business process modeling process, 3) for future business process modeling tools, and 4) for teaching of business process modeling languages.

1. *Importance of Verification:* The amount of errors in the different EPC model collections from practice that we used in this thesis emphasizes the importance of verification. We showed that an error ratio of about 10% is the average over the samples, with 3.3% being the minimum. While verification has been discussed for some time, this thesis demonstrates that the different techniques have matured to handle large sets of several thousand business process models on a common desktop computer. This observation relates to the gap between business analysis and information systems engineering in business process modeling (see [STA05, p.141] or [HR07, pp.424]), i.e. the refusal of engineers to reuse process documentations for systems implementation. While this gap is frequently accepted as a natural breach, this thesis tells a different story: the considerable amount of formal errors in documentation models makes it hardly possible to directly reuse them in the implementation. Therefore, the utilization of verification techniques in practice might be the key to eventually close this gap in the future.
2. *Business Process Modeling Process:* In this thesis, we gathered substantial theoretical and statistical evidence that formal errors are connected with several characteristics of the process model. This finding provides the opportunity to use process model metrics for the management of the design process and of process model quality. This is in particular the case if different design options have to be evaluated, and one

of multiple models might be considered to be superior regarding error-probability based on some metric. Furthermore, the strong connection between size and errors offers objective input for decisions regarding the question when a model should be split up or when model parts should be put in a sub-process. There is clearly a need for further research in this area. Nevertheless, our findings represent a major step towards establishing business process modeling as an engineering discipline beyond the intuition of the modeler.

3. *Business Process Modeling Tools:* Both items 1) and 2) call for respective tool support. While the verification techniques are apparently mature enough to deal with large models from practice, there seems to be too little attention paid to this issue by tool vendors. Indeed, tool vendors should have an interest in these topics since the lack of respective features has a negative impact on the productivity of the business process modeling exercise: models cannot be reused for system development, business users cannot interpret the models properly, and conclusions can hardly be drawn from the models regarding process performance. Beyond verification support, modeling tools could easily calculate process metrics to assist the modeler in the design task. Building on such features, the tool vendors could easily provide a greater benefit to their customers and help to improve the process of designing business process models.
4. *Teaching Process Modeling:* Apparently, there is a weakness in the way business process modeling is taught, if practitioners introduce a formal error in every tenth model (at least in our sample). Even worse, the four textbooks on business process modeling that we used to build the holdout sample had an even higher error ratio. While these rates might be partially attributed to missing verification support in the tools, there seems to be a problem for many modelers to understand the behavioral implications of their design. This has two consequences. Firstly, teaching of business process modeling should pay less attention to teaching a specific business process modeling language, but instead focus on conveying the general principles behind it, i.e., concurrency, synchronization, repetition, and other aspects captured, as by the workflow patterns. Secondly, formal errors seem to get too little attention. Concepts like deadlocks should not only be taught as a technical property of a business process model, but also an erroneous business rule that also leads to problems in real-world business processes that are not supported by information systems. Furthermore, the metrics are a good starting point for teaching patterns that are unlikely to result in errors, such as a high degree of structuredness appears to be less prone to cause errors. Such an approach might eventually deliver a better awareness and attention of formal errors in business process modeling practice.

## 4 Future Research

There are several open questions that could not be addressed in detail in this thesis. In particular, we focused on business process model metrics and their capability to predict

errors in business process models. We found strong evidence that our set of metrics can indeed explain a great share of the variation in error probability. Still, there are other factors we did not investigate in detail including personal factors, modeling purpose, domain knowledge, modeling language, or graphical layout that all might be connected with error probability. Furthermore, they might also be related to other important quality aspects that we did not analyze, like maintainability or understandability. In particular, we strongly agree with *Moody* [Moo05] who calls for more empirical research in business process modeling. This thesis and its findings gives an idea of the benefits we might gain from this research, and therefore may be regarded as an encouragement to follow *Moody's* call.

## References

- [Aal99] W.M.P. van der Aalst. Formalization and Verification of Event-driven Process Chains. *Information and Software Technology*, 41(10):639–650, 1999.
- [AH05] W.M.P. van der Aalst and A.H.M. ter Hofstede. YAWL: Yet Another Workflow Language. *Information Systems*, 30(4):245–275, 2005.
- [ARGP06] E. Rolón Aguilar, F. Ruiz, F. García and M. Piattini. Towards a Suite of Metrics for Business Process Models in BPMN. In Y. Manolopoulos, J. Filipe, P. Constantopoulos and J. Cordeiro, editors, *ICEIS 2006 - Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration (III)*, Paphos, Cyprus, May 23-27, 2006, pages 440–443, 2006.
- [BG05] S. Balasubramanian and M. Gupta. Structural metrics for goal based business process design and evaluation. *Business Process Management Journal*, 11(6):680–694, 2005.
- [BH05] H.U. Buhl and B. Heinrich. Meinung/Dialog: Empirical Research Strategies in Conceptual Modeling - Silver Bullet or Academic Toys? *Wirtschaftsinformatik*, 47(2):152–162, 2005.
- [Boe81] B.W. Boehm. *Software Engineering Economics*. Prentice-Hall, Englewood Cliffs, 1981.
- [BRU00] J. Becker, M. Rosemann and C. von Uthmann. Guidelines of Business Process Modeling. In W.M.P. van der Aalst, J. Desel and A. Oberweis, editors, *Business Process Management. Models, Techniques, and Empirical Studies*, pages 30–49. Springer, Berlin et al., 2000.
- [Car05] J. Cardoso. *Workflow Handbook 2005*, chapter Evaluating Workflows and Web Process Complexity, pages 284–290. Future Strategies, Inc., Lighthouse Point, FL, USA, 2005.
- [Car06] J. Cardoso. Process control-flow complexity metric: An empirical validation. In *Proceedings of IEEE International Conference on Services Computing (IEEE SCC 06)*, Chicago, USA, September 18-22, pages 167–173. IEEE Computer Society, 2006.
- [CFK05] N. Cuntz, J. Freiheit and E. Kindler. On the semantics of EPCs: Faster calculation for EPCs with small state spaces. In M. Nüttgens and F.J. Rump, editor, *Proceedings of the 4th GI Workshop on Business Process Management with Event-Driven Process Chains (EPK 2005)*, pages 7–23, Hamburg, Germany, December 2005. German Informatics Society.

- [CGP<sup>+</sup>05] G. Canfora, F. García, M. Piattini, F. Ruiz and C.A. Visaggio. A family of experiments to validate metrics for software process models. *Journal of Systems and Software*, 77(2):113–129, 2005.
- [CMNR06] J. Cardoso, J. Mendling, G. Neumann and H. Reijers. A Discourse on Complexity of Process Models. In Johann Eder and Schahram Dustdar, editors, *Proceedings of BPM Workshops 2006*, volume 4103 of *Lecture Notes in Computer Science*, pages 115–126, Vienna, Austria, 2006. Springer-Verlag.
- [CS94] R. Chen and A. W. Scheer. Modellierung von Prozessketten mittels Petri-Netz-Theorie. Heft 107, Institut für Wirtschaftsinformatik, Saarbrücken, Germany, 1994.
- [DAV05] B.F. van Dongen, W.M.P. van der Aalst and H.M.W. Verbeek. Verification of EPCs: Using Reduction Rules and Petri Nets. In O. Pastor and J. Falcão e Cunha, editors, *Advanced Information Systems Engineering, 17th International Conference, CAiSE 2005, Porto, Portugal, June 13-17, 2005, Proceedings*, volume 3520 of *Lecture Notes in Computer Science*, pages 372–386. Springer, 2005.
- [DR01] J. Dehnert and P. Rittgen. Relaxed Soundness of Business Processes. In K.R. Dittrick, A. Geppert and M.C. Norrie, editors, *Proceedings of the 13th International Conference on Advanced Information Systems Engineering*, volume 2068 of *Lecture Notes in Computer Science*, pages 151–170, Interlaken, 2001. Springer.
- [HR07] M. Hepp and D. Roman. An Ontology Framework for Semantic Business Process Management. In A. Oberweis, C. Weinhardt, H. Gimpel, A. Koschmider, V. Pankratius and B. Schnizler, editors, *eOrganisation: Service-, Prozess-, Market Engineering*, pages 423–440. Universitätsverlag Karlsruhe, 2007. Tagungsband der 8. Internationalen Tagung Wirtschaftsinformatik. Band 1.
- [Kin06] E. Kindler. On the semantics of EPCs: Resolving the vicious circle. *Data & Knowledge Engineering*, 56(1):23–40, 2006.
- [LA94] F. Leymann and W. Altenhuber. Managing Business Processes as an Information Resource. *IBM Systems Journal*, 33(2):326–348, 1994.
- [LG06] R. Laue and V. Gruhn. Complexity Metrics for Business Process Models. In Witold Abramowicz and Heinrich C. Mayr, editors, *9th International Conference on Business Information Systems (BIS 2006)*, volume 85 of *Lecture Notes in Informatics*, pages 1–12, 2006.
- [LSW98] P. Langner, C. Schneider and J. Wehler. Petri Net Based Certification of Event driven Process Chains. In J. Desel and M. Silva, editor, *Application and Theory of Petri Nets*, volume 1420 of *Lecture Notes in Computer Science*, pages 286–305, 1998.
- [LY92] G.S. Lee and J.-M. Yoon. An empirical study on the complexity metrics of Petri nets. *Microelectronics and Reliability*, 32(3):323–329, 1992.
- [MA06] J. Mendling and W.M.P. van der Aalst. Towards EPC Semantics based on State and Context. In M. Nüttgens and F.J. Rump and J. Mendling, editor, *Proceedings of the 5th GI Workshop on Business Process Management with Event-Driven Process Chains (EPK 2006)*, pages 25–48, Vienna, Austria, December 2006. German Informatics Society.
- [Men07] J. Mendling. *Detection and Prediction of Errors in EPC Business Process Models*. PhD thesis, Vienna University of Economics and Business Administration, 2007.

- [MMN<sup>+</sup>06a] J. Mendling, M. Moser, G. Neumann, H.M.W. Verbeek, B.F. van Dongen and W.M.P. van der Aalst. A Quantitative Analysis of Faulty EPCs in the SAP Reference Model. BPM Center Report BPM-06-08, BPMCenter.org, 2006.
- [MMN<sup>+</sup>06b] J. Mendling, M. Moser, G. Neumann, H.M.W. Verbeek, B.F. van Dongen and W.M.P. van der Aalst. Faulty EPCs in the SAP Reference Model. In J.L. Fiadeiro S. Dustdar and A. Sheth, editors, *Proceedings of BPM 2006*, volume 4102 of *Lecture Notes in Computer Science*, page 451457, Vienna, Austria, 2006. Springer-Verlag.
- [Moo05] D.L. Moody. Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. *Data & Knowledge Engineering*, 55(3):243–276, 2005.
- [Mor99] S. Morasca. Measuring Attributes of Concurrent Software Specifications in Petri Nets. In *METRICS '99: Proceedings of the 6th International Symposium on Software Metrics*, pages 100–110, Washington, DC, USA, 1999. IEEE Computer Society.
- [MR04] M. zur Muehlen and M. Rosemann. Multi-Paradigm Process Management. In *Proc. of the Fifth Workshop on Business Process Modeling, Development, and Support - CAiSE Workshops*, 2004.
- [Nis98] M.E. Nissen. Redesigning reengineering through measurement-driven inference. *MIS Quarterly*, 22(4):509–534, 1998.
- [NR02] M. Nüttgens and F.J. Rump. Syntax und Semantik Ereignisgesteuerter Prozessketten (EPK). In J. Desel and M. Weske, editor, *Proceedings of Promise 2002, Potsdam, Germany*, volume 21 of *Lecture Notes in Informatics*, pages 64–77, 2002.
- [Ros06] M. Rosemann. Potential pitfalls of process modeling: part A. *Business Process Management Journal*, 12(2):249–254, 2006.
- [RV04] H.A. Reijers and I.T.P. Vanderfeesten. Cohesion and Coupling Metrics for Workflow Process Design. In J. Desel, B. Pernici and M. Weske, editors, *Business Process Management: Second International Conference, BPM 2004, Potsdam, Germany, June 17-18, 2004. Proceedings*, volume 3080 of *Lecture Notes in Computer Science*, pages 290–305. Springer, 2004.
- [STA05] A.-W. Scheer, O. Thomas and O. Adam. *Process Aware Information Systems: Bridging People and Software Through Process Technology*, chapter Process Modeling Using Event-Driven Process Chains, pages 119–146. Wiley Publishing, 2005.
- [VA06] H. M. W. Verbeek and W.M.P. van der Aalst. On the verification of EPCs using T-invariants. BPMCenter Report BPM-06-05, BPMcenter.org, 2006.
- [WAHE06] M.T. Wynn, W.M.P. van der Aalst, A.H.M. ter Hofstede and D. Edmond. Verifying Workflows with Cancellation Regions and OR-joins: An Approach Based on Reset Nets and Reachability Analysis. In S. Dustdar, J.L. Fiadeiro, and A. Sheth, editors, *Proceedings of BPM 2006*, volume 4102 of *Lecture Notes in Computer Science*, pages 389–394, Vienna, Austria, 2006. Springer-Verlag.
- [WEAH05] M.T. Wynn, D. Edmond, W.M.P. van der Aalst and A.H.M. ter Hofstede. Achieving a General, Formal and Decidable Approach to the OR-join in Workflow using Reset nets. In G. Ciardo and P. Darondeau, editors, *Applications and Theory of Petri Nets 2005*, volume 3536 of *Lecture Notes in Computer Science*, pages 423–443, 2005.

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

## EMISA-Edition, Folge 16

Gottfried Vossen, Universität Münster

In dieser Rubrik möchte ich Ihnen wieder ein paar Websites vorstellen, die mir in letzter Zeit selbst oder mit fremder Hilfe aufgefallen sind. Ich weise vorab wieder darauf hin, dass die Inhalte sämtlicher hier beschriebenen Webseiten urheberrechtlich geschützt sind, allerdings ist nach einschlägiger Meinung das Copyright nur relevant für die Verwendung in anderen Webseiten bzw. wenn Gestaltungselemente für andere Designs übernommen würden, was beides nicht der Fall ist. Insofern betrachte ich das Beschreiben von Seiten in der hier vorgenommenen Form weiterhin als Werbung für diese.

Ich beginne diesmal mit einem Thema, mit dem sich die EMISA schon seit Längerem beschäftigt: Service-Orientierung. Im Rahmen aktueller Web 2.0-Entwicklungen erlebt das Thema *Application Service Provisioning* (ASP) ja eine neue Blüte, wobei ein wesentlicher Unterschied zu früheren ASP-Ansätzen die Vielfalt der Dienste ist, die inzwischen im Web gehostet und über das Web angeboten werden, sowie die Einfachheit und geringen Kosten, mit denen man die Dienste beanspruchen kann. Ein typisches Beispiel ist 37signals, denn dort findet man gleich mehrere interessante Dienste: das Projektmanagementwerkzeug Basecamp, die Kalenderapplikation Backpack, das Gruppenchatsystem Campfire und das Kontaktverwaltungssystem Highrise. Über diese Dienste kann man dort folgendes nachlesen:

The screenshot shows the 37signals website with a navigation bar at the top containing links for 'Our products', 'Our book', 'Company blog', and 'Product blog'. The main content area features a large heading 'Work well.' followed by the text 'Over 1 million people and businesses use our web-based applications to get things done the simple way.' Below this, it states 'We aim for the software sweet spot: Elegant, thoughtful products that do just what you need and nothing you don't.' A quote from TIME magazine is also present: 'One of the Net's rising stars.' The products listed are: Basecamp (Project management and collaboration), Highrise (Track leads, clients, vendors, simple CRM), Campfire (Real-time group chat), and Backpack (Information organizer and calendar).

Join the about-once-a-month 37signals newsletter

Be the first to hear about new products, new features, special limited time promotions, discounts, and more. No spam.

Email:

- a. Basecamp takes a fresh, novel approach to project collaboration. Projects don't fail from a lack of charts, graphs, stats, or reports, they fail from a lack of clear

- communication. Basecamp solves this problem by providing tools tailored to improve the communication between people working together on a project.
- You can use Highrise to review a colleague's notes before calling her contact at the printer, see all the follow-ups scheduled for this week, set a reminder to write your client a thank-you note next Friday, keep all important emails from a customer together on one page, schedule a follow-up sales call with a lead in 30 days, review all communications with your investors.
  - Campfire is a web-based group chat tool that lets you set up password-protected chat rooms in just seconds. Invite a client, colleague, or vendor to chat, collaborate, and make decisions. Link to a room on your intranet for internal communications. There are so many ways to use Campfire. Chatting, file sharing, image previewing, decision making, etc. Up to 60 people can chat at once. You can even browse previous chats by person, room, or date.
  - Backpack is a simple web-based service that allows you to make pages with to-do lists, notes, files, and images. Backpack also features a Calendar and Reminders that can be sent via email or to your cell phone at predefined times. Tagging pages makes it easy to group related pages together.

Insbesondere Basecamp bekommt sehr gute Kritiken für Projektmanagement. Seine Hauptseite zeigt ein Dashboard mit anstehenden Milestones und letzten Aktivitäten. Diese Sichten sind auf einzelne Projektmitarbeiter personalisierbar, so dass jeder seine anstehenden Aktivitäten sehen kann. Letztere sind als To-Do-Listen organisiert und können Milestones zugeordnet werden. Weitere Funktionalitäten sind Zeitnachverfolgung, File Sharing sowie ein Message Board mit Kommentarfunktion.

The screenshot shows the 'eNormicom Extranet' dashboard. At the top, it indicates the user is logged in as 'Larry Appleton'. The main navigation includes 'Dashboard', 'To-Dos', 'Milestones', and 'People'. A sidebar on the right contains 'Templates' and 'Settings'. The dashboard is divided into several sections:

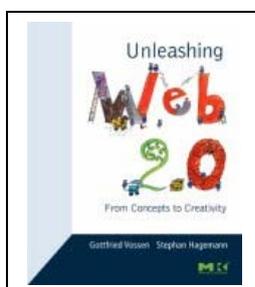
- 1 Late Milestones:** A red box highlights a milestone that is 7 days late: 'Awaiting payment so we can release the document' (Smithco | Intranet experience review | Matt Linderman).
- 2 e-NORMICOM .COM:** The company logo.
- 3 Due in the next 14 days:** A calendar view showing tasks for the week of Oct 27-31. Tasks include 'Review Personal/Basic' on Oct 27 and 'Kick off meeting' on Oct 29.
- 4 Your projects:** A list of projects with a 'Create a new project' button. Projects include 'Office relocation', 'Site redesign', 'Backoffice screens', 'Intranet experience review', 'Snowglobes', and 'Web Site Redesign'.
- 5 Latest activity:** A feed of recent activity categorized by project:
  - Martin — Backoffice screens:** Includes messages and to-dos about 'New Order Process: Revised "Find Item..."', 'Initial Screens', 'Search/Browser Questions', and 'Order Screen Flow'.
  - Tilted — Snowglobes:** Includes comments about 'IP allocation' and 'Cluster Network Performance Analysis'.
  - Smithco — Intranet experience review:** Includes milestones for 'Get sign-off from Marcus' and 'Add new categories inline on the ...'.

Es gibt eine Reihe von vergleichbaren Entwicklungen im Web, die ich in letzter Zeit im Zusammenhang mit Arbeiten an meinem neuen Buch recherchiert habe. Als kleinen „Teaser“

darf ich meinen Leser eine Tabelle zeigen, die auf S. 259 zu finden ist und die eine Übersicht über aktuelle „Software as a Service“ (SaaS) Anwendungen im Web gibt:

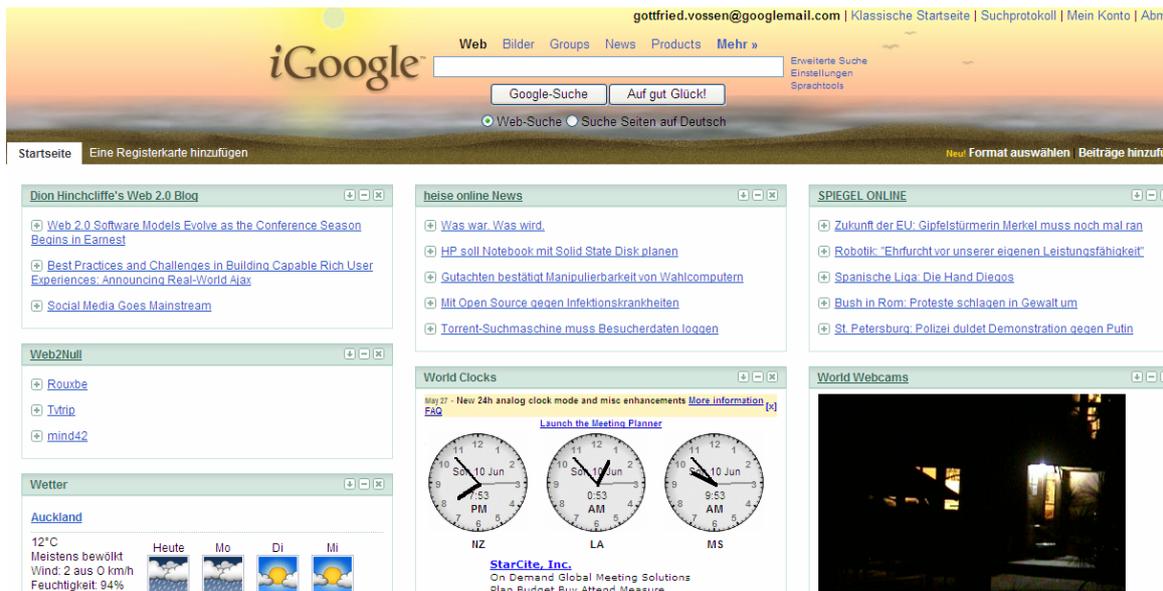
**Table 5.1 SaaS applications.**

Application	Company	URL
Organizer and Calendar	Kiko	<a href="http://www.kiko.com">www.kiko.com</a>
	37signals Backpack	<a href="http://www.backpackit.com">www.backpackit.com</a>
	30Boxes	<a href="http://www.30boxes.com">www.30boxes.com</a>
Planning	PlanHQ	<a href="http://www.planhq.com">www.planhq.com</a>
Word Processing and Spreadsheets	EditGrid	<a href="http://www.editgrid.com">www.editgrid.com</a>
	ThinkFree	<a href="http://www.thinkfree.com">www.thinkfree.com</a>
	Google Docs & Spreadsheets	<a href="http://docs.google.com">docs.google.com</a>
	Ajax 13 ajaxWrite, ajaxXLS	<a href="http://www.ajaxwrite.com">www.ajaxwrite.com</a>
	Num Sum	<a href="http://www.numsum.com">www.numsum.com</a>
	Zoho	<a href="http://www.zoho.com">www.zoho.com</a>
	Peepel	<a href="http://www.peepel.com">www.peepel.com</a>
Distributed Database Applications	ShareOffice	<a href="http://www.sharemethods.com/products/shareoffice.html">www.sharemethods.com/products/shareoffice.html</a>
	Smallthought Systems Dabble DB	<a href="http://www.dabledb.com">www.dabledb.com</a>
Invoicing and Time Tracking	FreshBooks	<a href="http://www.freshbooks.com">www.freshbooks.com</a>
Project Management, Collaboration, Shared Workspaces, Conferencing	37signals Basecamp	<a href="http://www.basecamphq.com">www.basecamphq.com</a>
	CentralDesktop	<a href="http://www.centraldesktop.com">www.centraldesktop.com</a>
	WebEx	<a href="http://www.webex.com">www.webex.com</a>
	GoToMeeting	<a href="http://www.gotomeeting.com">www.gotomeeting.com</a>
	activeCollab	<a href="http://www.activecollab.com">www.activecollab.com</a>
	TeamWork Live	<a href="http://www.teamworklive.com">www.teamworklive.com</a>
	LiveOnTheNet	<a href="http://www.liveonthenet.com">www.liveonthenet.com</a>
Group Chat	Zimbra Collaboration Suite	<a href="http://www.zimbra.com">www.zimbra.com</a>
	37signals Campfire	<a href="http://www.campfirenow.com">www.campfirenow.com</a>
Accounting and Payroll	Xero	<a href="http://www.xero.com">www.xero.com</a>
	Intacct	<a href="http://www.intacct.com">www.intacct.com</a>
	KeepMore	<a href="http://www.keepmore.com">www.keepmore.com</a>
Customer Relationship Management	Etelos CRMforGoogle	<a href="http://www.crmforgoogle.com">www.crmforgoogle.com</a>
	salesforce.com	<a href="http://www.salesforce.com">www.salesforce.com</a>
	RightNow	<a href="http://www.rightnow.com">www.rightnow.com</a>

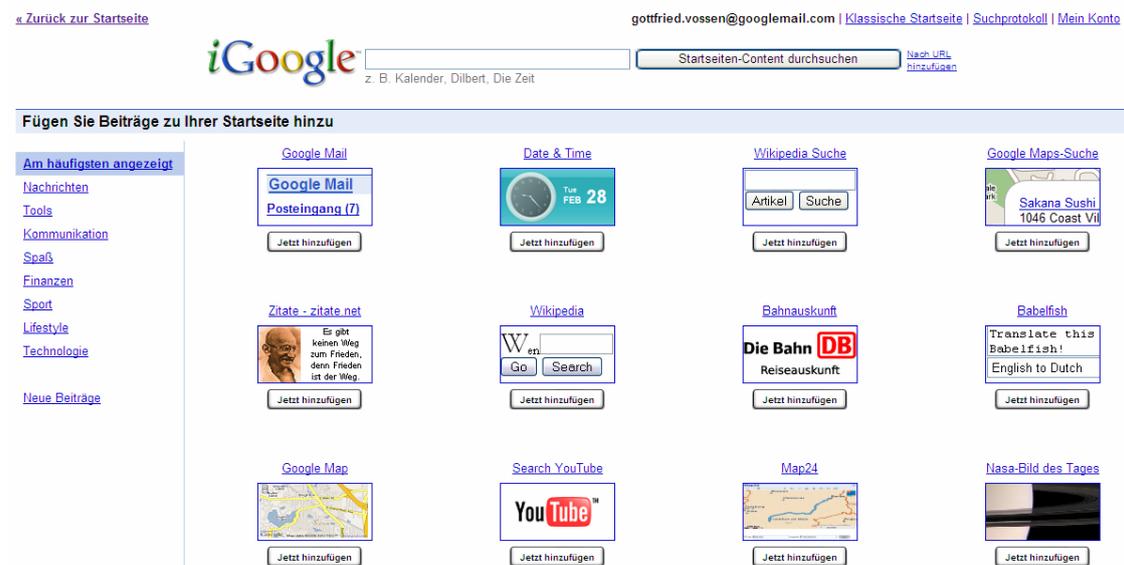


Das Buch selbst wird unter dem Titel *Unleashing Web 2.0* Ende Juli 2007 erscheinen; weitere Informationen finde meine Leserinnen und Leser unter [www.unleashingweb20.com](http://www.unleashingweb20.com).

Eine weitere interessante Anwendung, die immer häufiger anzutreffen ist, ist das Personalisieren von Homepages im Web. Auch Google lässt dies jetzt zu mit seinem iGoogle-Ansatz. Man muss registrierter Google-Nutzer sein und kann sich dann seine Startseite unter [www.google.com/ig](http://www.google.com/ig) nach eigenen Wünschen einrichten. Nachfolgend gezeigt ist meine eigene Startseite, die derzeit noch stark von meiner Sicht auf die Welt aus dem Südpazifik geprägt ist. Das Suchfeld der klassischen Google-Homepage bleibt erhalten, rückt allerdings weit nach oben (und kann etwas ansprechender als in weiß gestaltet werden).



Das Interessante ist daran ist die Fülle an Gadgets, aus denen man auswählen kann. Google bietet natürlich eine Suchfunktion über den Gadget-Katalog, und das Konfigurieren ausgewählter Komponenten geschieht dann per Drag and Drop.



Vergleichbares bietet auch zahlreiche andere Webseiten, einige auch schon länger als Google.

Ich bleibe noch bei Google, um deren neues Feature der „universellen Suche“ zu beleuchten. Dazu ist im GoogleWatchBlog unter [www.googlewatchblog.de/2007/05/16/google20-neues-design-experimental-search-universal-search/](http://www.googlewatchblog.de/2007/05/16/google20-neues-design-experimental-search-universal-search/) nachzulesen: „Bereits Anfang Februar haben Marissa Mayer und Larry Page angekündigt die Suchergebnisse der einzelnen Suchtypen zusammenzufassen. Konkret heißt das, dass die Suchmaschine erkennt, wonach der User sucht, und ihm die jeweiligen Suchergebnisse anbieten wird. Da kann das erste Ergebnis aus der Patentsuche kommen, das zweite aus der Bildersuche, das dritte aus der Buchsuche und danach erst werden ein paar Webseiten gelistet. .... Als weitere Neuerung wird Google demnächst eine Suchmaschine ankündigen die an keine Sprachbarrieren mehr gebunden ist.

Der Suchstring wird automatisch in 11 verschiedene Sprachen übersetzt und die jeweiligen Suchergebnisse wieder in die eigene Sprache zurückübersetzt. Somit findet man jetzt z. B. auch mit einem deutschen Suchbegriff interessante Informationen auf einer französischen Webseite und kann sich diese gleich übersetzen lassen.“ Ich habe das universelle Suchen ausprobiert für den Suchbegriff „hydrofoil“; das Ergebnis ist nachfolgend gezeigt:

[Anmelden](#)


[Web](#) [Bilder](#) [Groups](#) [News](#) [Produkte](#) [Mehr »](#)

[Erweiterte Suche](#)  
[Einstellungen](#)

Suche:  Das Web  Seiten auf Deutsch  Seiten aus Deutschland

---

**Web** Ergebnisse 1 - 10 von ungefähr 832.000 für **hydrofoil**. (0,06 Sekunden)

Tipp: [Suchen nur nach Ergebnissen auf Deutsch](#). Sie können Ihre bevorzugten Spracheinstellungen in [Einstellungen](#) angeben.

[Ergebnisse](#) [Bildersuche nach hydrofoil](#)



**Hydrofoil** - [Wikipedia, the free encyclopedia](#) - [ [Diese Seite übersetzen](#) ]  
A **hydrofoil** is a boat with wing-like foils mounted on struts below the hull. As the craft increases its speed the **hydrofoils** develop enough lift for the ...  
[en.wikipedia.org/wiki/Hydrofoil](http://en.wikipedia.org/wiki/Hydrofoil) - 43k - [Im Cache](#) - [Ähnliche Seiten](#)

[International Hydrofoil Society](#) - [ [Diese Seite übersetzen](#) ]  
The Premier Website for Hydrofoils and Hydrofoilers.  
[www.foils.org/](http://www.foils.org/) - 21k - [Im Cache](#) - [Ähnliche Seiten](#)

[Hydrofoils Incorporated Home Page](#) - [ [Diese Seite übersetzen](#) ]  
Hydrofoils Incorporated is proud to announce a new working partnership with Turbine Marine, the premium marine company specializing in Turbine installations ...  
[www.hydrofoil.com/](http://www.hydrofoil.com/) - 11k - [Im Cache](#) - [Ähnliche Seiten](#)

Anzeigen

**Hydrofoil**  
Günstige Angebote gibt's bei eBay.  
Mitbieten oder Sofort-Kaufen!  
[www.ebay.de](http://www.ebay.de)

[www.DES-Boots-Service.de](http://www.DES-Boots-Service.de)  
Onlineshop - Bootteile und Zubehör  
für Wassersportler / Bootsbesitzer  
[www.DES-Boots-Service.de](http://www.DES-Boots-Service.de)

**Hydrofoil**  
Alles, was Ihr Herz begehrt!  
Wir haben's - bis 75% billiger.  
[www.Preisvergleich.de](http://www.Preisvergleich.de)

Wie man sieht, werden jetzt neben „gewöhnlichen“ Suchergebnissen auch z. B. Bilder mit ausgegeben, wobei das Neue darin besteht, dass diese nicht mehr zu Beginn der betreffenden Quell-Seite erscheinen müssen, sondern irgendwo auf der Seite. Im Laborstadium befindet sich ferner die *Experimental Search* (siehe <http://www.google.com/experimental/>), ein Ausblick auf die Zukunft der Suchmaschinen, bei der man sich Suchergebnisse u. a. zeitlich sortieren lassen kann, wie hier für den Suchbegriff „thomas jefferson“ gezeigt:




[Advanced Search](#)  
[Preferences](#)

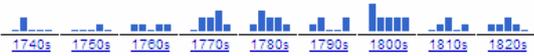
[List view](#) [Timeline view](#) [Map view](#)

[« Return to experiment descriptions](#) | [Send feedback](#) | Example searches: [earthquakes](#) [george washington](#) [disney](#)

---

**Timeline View**

1740s - 1830s



1740s [April 13, 1743](#)  
Thomas Jefferson (13 April 1743 N.S.–4 July 1826) was the third President of the United States (1801–1809), the principal author of the Declaration of ...  
[Thomas Jefferson - Wikipedia, the free encyclopedia ... - en.wikipedia.org](#)  
[1 more result in 1740 - 1750 »](#)

1750s [1 result in 1750 - 1760 »](#)

1760s [4 results in 1760 - 1770 »](#)

1770s [8 results in 1770 - 1780 »](#)

1780s [1785](#)  
Jefferson succeeded Benjamin Franklin as minister to France in 1785. His sympathy for the French Revolution led him into conflict with Alexander Hamilton when ...  
[Biography of Thomas Jefferson - www.whitehouse.gov](#)

[1786](#)  
In years following he labored to make its words a reality in Virginia. Most notably, he wrote a bill establishing religious freedom, enacted in 1786.

Universelle Suche, bei der man sich nicht auf einen Typ von Ergebnis beschränkt, findet sich auch bei zahlreichen anderen Suchmaschinen. Als Beispiel zeige ich nachfolgend das Ergebnis einer Suche nach „hydrofoil“ bei Searchmash.com:

The screenshot shows the Searchmash search engine interface. The search bar contains the word "hydrofoil". Below the search bar, there are two main sections: "Web Pages - about 832,000" and "Images - about 15,800".

**Web Pages - about 832,000**

- Hydrofoil - Wikipedia, the free encyclopedia**  
en.wikipedia.org - A hydrofoil is a boat with wing-like foils mounted on struts below the hull. As the craft increases its speed the hydrofoils develop enough lift for the ...
- Sky Ski - Comptech Billet Hydrofoils & Accessories**  
skyski.com - Manufacturer of the Sky Ski hydrofoil for above-the-water skiing. On-line sales.
- Hydrofoils Incorporated Home Page**  
hydrofoil.com - Hydrofoils Incorporated is proud to announce a new working partnership with Turbine Marine, the premium marine company specializing in Turbine installations ...
- International Hydrofoil Society**  
foils.org - The Premier Website for Hydrofoils and Hydrofoilers.
- The Hydrofoil Water Scooter at Hammacher Schlemmer**  
hammacher.com - Lightweight hydrofoil frame allows you to skim across water by hopping gently up and down, propelling you as fast as 17 mph.
- Welcome to USHA**  
hydrofoil.org - Why should I join USA Waterski for USHA? How can I find or locate other hydrofoil skiers in my area? Then take a look at our membership information page.
- Russian Hydrofoil Page**  
hydrofoils.org - Provides information about hydrofoils, methods of calculations, design service, and photo galleries.
- American Hydrofoil**  
americanhydrofoil.com - American Hydrofoil is a boat manufacturer of pleasure, commercial, and military hydrofoil watercraft.
- Decavitator**  
lancet.mit.edu - History, specifications, manufacturing methods and other information related to a human-powered hydrofoil.

**Images - about 15,800**

See image results »

Below the images, there are sections for "Blogs", "Videos", and "Wikipedia".

**Feedback**

Were these results useful to you?

Web Pages	Yes / No
Images	Yes / No
Blogs	Yes / No
Videos	Yes / No
Wikipedia	Yes / No

Auch hier werden verschiedene Ergebnistypen angeboten (siehe rechts im Screenshot: Images, Blogs, Videos, Wikipedia), wobei ich per Feedback angeben kann, wie wichtig mir diese Quellen (bei zukünftigen Suchen) sind.

Eine weitere universelle Suchmaschine ist [www.hakia.com](http://www.hakia.com), für die ich hier Suchergebnisse zum Begriff „toyota“ zeige:

The screenshot shows the hakia search engine interface. The search bar contains the word "toyota". Below the search bar, there are several sections:

**Headlines**

- News: Isuzu To Build Diesels For Toyota**  
...would be ready to announce an outline when it unveils a new mid-term business plan in July. The press reports added that Isuzu would invest some \$240 million to build a new plant in Hokkaido, Japan to supply about 200,000 diesel engines a year exclusively to Toyota starting about 2012. June 15, 2007...  
Diesel Progress - 6 hours ago
- News: Custom 2000 Toyota Hi-Lux - Taco-Lux**  
... himself cruising the West Coast in California. To accomplish this dream, however, was a long and hard road. To start, he picked up a Toyota Hi-Lux, which is the Japanese version of a Tacoma. They look very similar but are completely different, and Kousuke wanted it to look more like a Taco than it...  
Mini Truckin Magazine - 11 hours ago

**Profile on Car Manufacturer**

**TOYOTA: Company > Company Profile**  
The year 2005 was another significant year for Toyota Motor Corporation (TMC). It was the year that Toyota played part in the EXPO 2005 Aichi, Japan, which attracted 22 million visitors, stimulated the local economy, and entertained all of

**Table of Contents**

- Headlines
- Profile on Car Manufacturer
- Popular Models
- Parts and Accessories
- Find a Dealer
- Used Cars
- History
- News and Interviews
- Photographs and Pictures
- Awards and Accomplishments
- Myths and Controversies
- Criticism, Commentary, and Interpretations

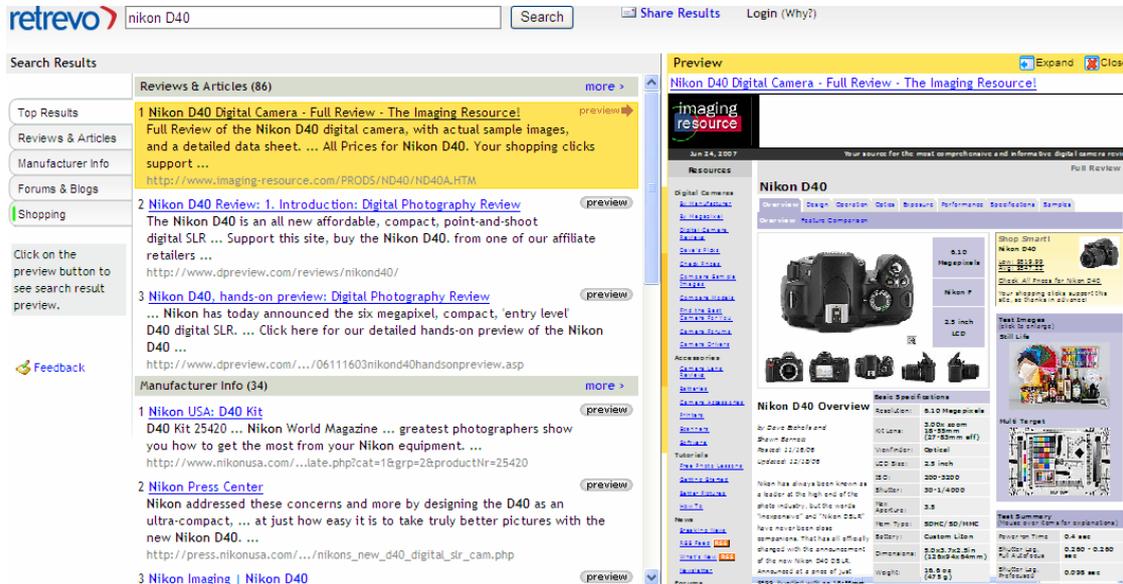
**Sponsored Results**

**Toyota**  
Spitzen-Angebote zu Toyota. Toyota hier!  
www.billiger.de/Toyota

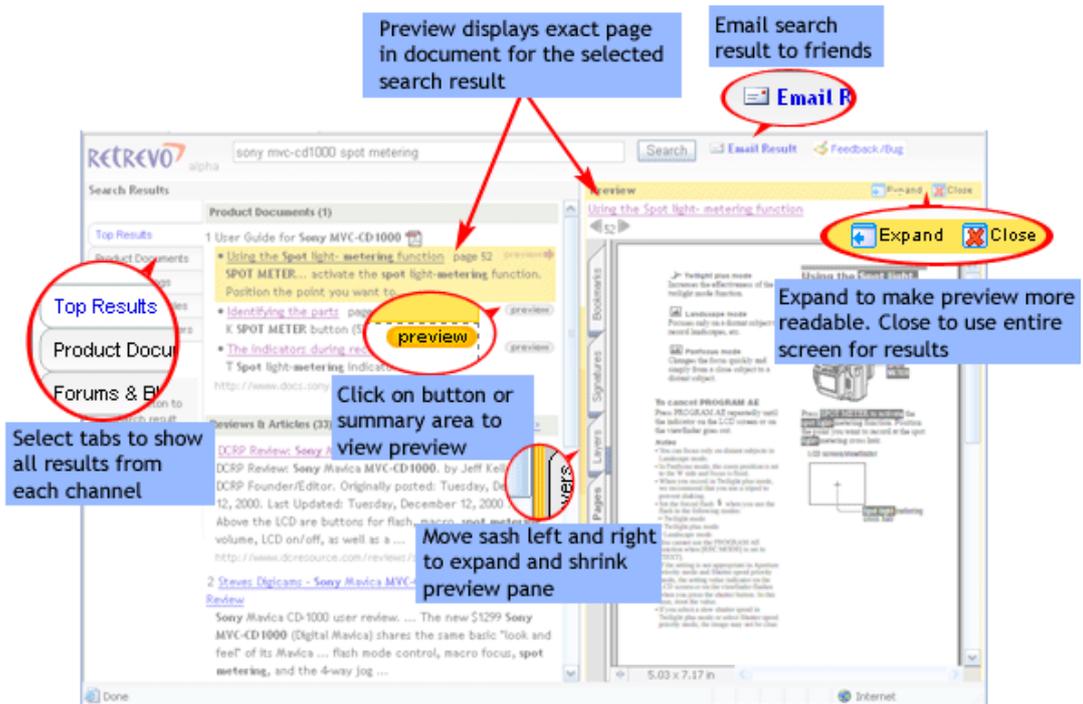
**Toyota**  
Toyota neu & gebraucht - Hier findet jeder das passende Auto  
www.AutoScout24.de/Toyota

Trotz Beta-Stadium werden hier bereits Anzeigen aus Deutschland einbezogen; ansonsten gibt es Headlines, ein Company Profile, Bilder und einen „Table of Contents“, der sämtlich Ergebnisse in der Übersicht darstellt.

Ebenfalls hochinteressant ist [www.retrevo.com](http://www.retrevo.com), eine Suchmaschine, die sich auf Gadgets und Consumer Electronics konzentriert. Suchergebnisse beinhalten hier Dokumente zum gewünschten Produkt, Herstellerinformationen sowie eine Vorschau auf den User Guide zum Produkt als pdf-Datei. Der nachfolgende Screenshot zeigt das Ergebnis einer Suche nach „nikon D40“:



Im Unterschied zu Hakia gibt Retrevo zumindest Erläuterungen, wie die Ergebnisseite zu interpretieren bzw. zu lesen ist:



Auch bei [www.clusty.com](http://www.clusty.com) habe ich nach „hydrofoil“ gesucht, diesmal mit folgendem Ergebnis:

The screenshot shows the Clusty search engine interface. At the top, there are navigation links for 'web', 'news', 'images', 'wikipedia', 'blogs', 'jobs', and 'more'. A search bar contains the query 'hydrofoil'. Below the search bar, there are tabs for 'clusters', 'sources', and 'sites'. The 'clusters' tab is active, showing a list of categories with result counts: All Results (192), Design (22), Picture (17), Collision, Sicily (13), Human Powered Hydrofoil (11), Sailing (13), Speed The Hydrofoils (9), Air (11), History (9), Hydrofoil Water (9), and Flight (9). A search bar with a 'Find' button is located below the clusters. The main search results area shows 'Top 191 results of at least 737,058 retrieved for the query hydrofoil'. The first result is 'Hydrofoil Water Craft' with a sponsored link to 'www.aquaskipper.com'. The second result is 'International Hydrofoil Society' with a description and a link to 'www.foils.org'. The third result is 'Hydrofoil' with a small image and a description: 'A hydrofoil is a boat with wing-like foils mounted on struts below the hull. As the craft increases its speed the hydrofoils develop enough lift for the boat to become foilborne - i.e. to raise the hull up and out of the water. This results in a great reduction in drag and a corresponding increase in speed.' The fourth result is 'Decavitator' and the fifth is 'American Hydrofoil.com - Home Page'.

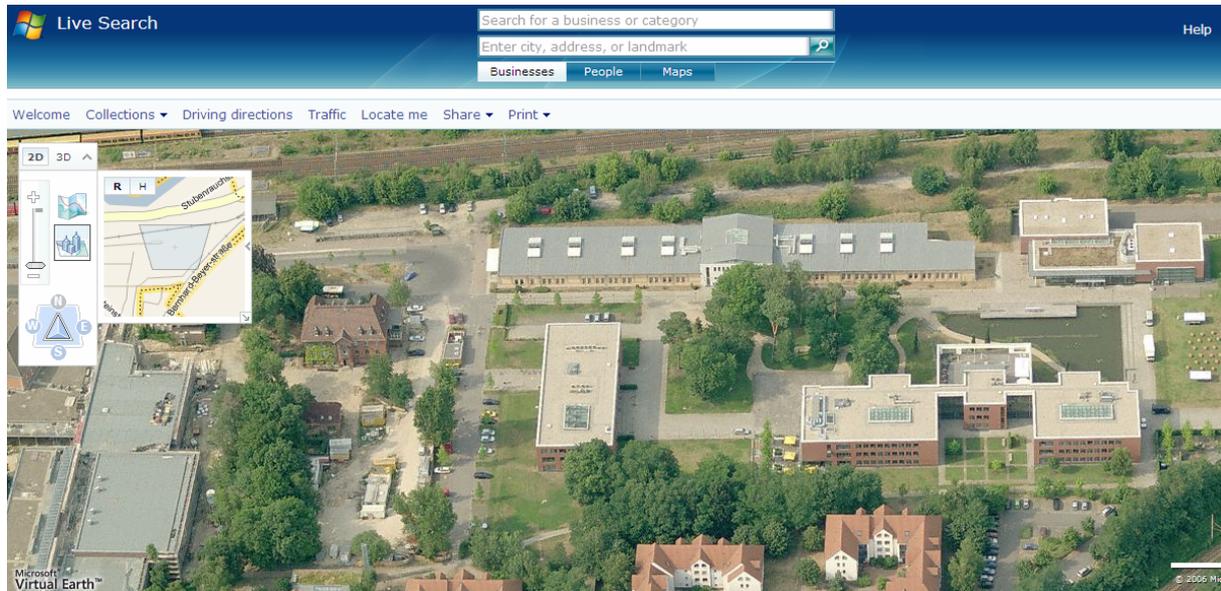
Auch hier erkennt man auf der linken Seite ein Inhaltsverzeichnis mit Clustern, in welche die Suchergebnisse eingeteilt sind. Die angezeigte Liste lässt sich erweitern („more“ bzw. „all clusters“); man kann sich aber auch anzeigen lassen, aus welchen Quellen die Ergebnisse stammen (in meinem Beispiel u. a. von Ask, Gigablast, MSN oder Wisenut) oder („sites“) wie viele .com-Sites, .org-Sites usw. sich unter den Ergebnissen befinden.

Wenn man die bisher betrachteten Suchbegriffe „hydrofoil“, „toyota“, „thomas jefferson“ sowie „nikon D40“ dagegen bei Microsoft’s live.com eingibt, sieht das Ergebnis eher traditionell aus:

The screenshot shows the Live Search interface for the query 'nikon D40'. The search bar is at the top, and there are tabs for 'Web', 'Bilder', 'News', 'Weitere', 'Wissenschaftlich', and 'Feeds'. The search results are displayed in a list format. The first result is 'Nikon Digital-Spiegelreflexkameras D40 & D40x' with a link to 'www.nikon.de'. The second result is 'Nikon D40 - Riesenauswahl an Mode- und Trendmarken' with a link to 'www.ebay.de'. The third result is 'Nikon Deutschland - D40 Kit schwarz' with a link to 'www.nikon.de/product/de\_DE/products/broad/1111/overview.html'. The fourth result is 'Nikon Deutschland - Home' with a link to 'www.nikon.de'. The fifth result is 'Nikon D40 Test - Preisvergleich für Nikon D40' with a link to 'www.ciao.de'. The sixth result is 'Nikon D40 SLR nur € 549,00' with a link to 'www.camerahouse.de'. The seventh result is 'Nikon D40 bei Quelle.de bestellen' with a link to 'www.quelle.de'. The eighth result is 'Nikon D40 - im Preisvergleich' with a link to 'www.preisvergleich.de'. The ninth result is 'Nikon D40 - Spiegelreflexkameras im Vergleich' with a link to 'www.ciao.de'. The tenth result is 'Nikon D40 Kit' with a link to 'www.ciao.de'.

Allerdings kann man auch hier unmittelbar und ohne erneutes Suchen zwischen Webadressen, Bildern, News etc. als Ergebnissen wählen. Live.com hat allerdings etwas anderes zu bieten, was Google durchaus Kopfschmerzen bereiten dürfte, nämlich <http://maps.live.com>. „Live Search Maps is a free global mapping and search service that lets you search for things in

specific locations, see bird's eye or 3D views of many major cities, find traffic reports, and explore your world in new ways.“ Hier kann man sich also Teile der Welt als Karte, aus der Vogelperspektive oder in dreidimensionaler Darstellung anzeigen lassen. Als Beispiel zeige ich hier das HPI in Potsdam; man kann ja jetzt mal versuchen herauszufinden, wo unser Fachgruppensprecher sein Auto parkt (und ob er überhaupt im Büro ist – allerdings sehen wir hier keine Live-Ansicht, auch wenn der Name der Site das suggeriert).



Das Suchen wird uns also noch lange beschäftigen und immer wieder interessante Ergebnisse liefern. Allerdings ist die Intelligenz von Suchmaschinen und deren Gründlichkeit allmählich auch Besorgnis erregend.

## Weitere Ankündigungen:

- Tagungsbericht zur *BIS 2007 – 10<sup>th</sup> Int'l Conference on Business Information Systems*, Posen, 2007
- Call for Participation: *BPM'07 – 5th Int'l Conference on Business Process Management*, Brisbane, 24.–27.09.2007
- Call for Papers: *Modellierung'08*, Berlin, 12. – 14.03.2008
- Call for Papers: *XML4BPM'08 - 4th GI-Workshop on XML Integration and Transformation for Business Process Management*, München, 28.02. – 02.03. 2008
- Call for Papers: *ISTA'08 - 7th International Conference on Information System Technology and its Applications*, Klagenfurt, 22. – 25.04.2008



## **BIS 2007 – 10th International Conference on Business Information Systems**

On April 25-25, 2007, 10th International Conference on Business Information Systems - BIS 2007 took place in Poznań, Poland, organized by the Department of Information Systems at Poznań University of Economics. Since its origins the Conference has focused on design, implementation and application of information systems supporting organizations. The main theme of this year's conference was "Semantic integration of data and processes across enterprises and societies". The focus was on cutting-edge research: web services, ontologies, semantic business process management, social aspects of IT, but also classical topics were present: information retrieval, mobile architectures, system design, and electronic government.

During the last 10 years the Conference steadily improved its rank. It started as a small mostly Polish-German conference and this year proceedings: of BIS 2007 were published as a volume in Lecture Notes in Computer Science series by Springer Verlag and of BIS 2006 were published as a post conference book also by Springer. It also deserves a name "international"; there were over 200 papers authored by 364 researchers from 40 countries. The Program Committee, consisting of 96 renowned scientists and experts from 22 countries, supported by 61 additional reviewers, accepted 49 papers (less than 25% acceptance rate). Almost 100 guests came to Poznań to attend the Conference, of which ca. 70 were from abroad, mostly from Germany and Italy, but there were also participants from such distant countries as Australia, Canada, China, Japan, Korea, and USA.

The Conference was inaugurated by two keynotes speeches. Prof. Alois Ferscha from Johannes Kepler University Linz (Austria) held an interesting presentation about "Digital artefacts" supported by demonstration of interaction with some of those presented digital objects. Next, Prof. A Min Tjoa from Vienna University of Technology (Austria) presented "Security Issues for the Use of Semantic Web in e-Commerce". Other days of the Conference also started with keynote addresses: Dr. John Domingue from The Open University (UK) introduced „Semantic Web Services and Business Processes: Approaches and Applications" and Dr. Wasim Sadiq from SAP Research, Australia tried to answer the question "Have BPMS delivered for

Processes like DBMS have for Data?” The Conference program was also enriched with several workshops: Ontology Evolution – OnE 2007, Social Aspects of the Web – SAW 2007, Time in Business Scenarios – TeBIS 2007, Web services interactions, quality and SLAs – WS-IQS 2007. On the last day there was also accompanying subconference devoted to electronic government: TCGOV 2007 - Towards Electronic Democracy Programme Conference on e-Government, organized by University of Applied Sciences Northwestern Switzerland.

BIS 2007 was sponsored by European Project SUPER (FP6-026850), EMISA, and several local IT companies from Poznań, members of Partner Club of the Department of Information Systems. BIS2007 was also supported by ACM Special Interest Group on Management Information Systems, Gesellschaft für Informatik, and Polish Society for Information Systems. Media patronage was held by Gazeta IT.

The next eleventh BIS will take place on May 5-7, 2008 in Innsbruck (Austria) - [http://bis.kie.ae.poznan.pl/11th\\_bis/](http://bis.kie.ae.poznan.pl/11th_bis/). We invite all researchers and practitioners who are interested in the topic!



From the left: Prof. Abramowicz (conf. chair), Prof. Ferscha (keynote speaker), Prof. Gorynia (vice rector of the Poznan University of Economics), Prof. A Min Tjoa (keynote speaker)



Participants of the conference in front of Wilson Park



Prof. Ferscha and one of the digital artefacts at the bottom of the screen

**Call For Participation  
5th International Conference on  
Business Process Management (BPM'07)**

**Brisbane, 24.–27. September 2007  
<http://bpm07.fit.qut.edu.au/>**



***Wissenschaftliches Programm***

***Montag, 24. September 2007***

**Workshops**

- 3rd Int'l Workshop on Business Process Intelligence
- 3rd Int'l Workshop on Business Process Design
- 1st Int'l Workshop on Collaborative Business Processes
- 1st Int'l Workshop on Process-oriented Systems in Healthcare
- 10th Int'l Workshop on Reference Modeling
- 2nd Int'l Workshop on Semantics in Web Services and Business Processes

***Dienstag, 25. September 2007***

**09:00 - 9:30**

**Welcome**

M. Rosemann, *General Conference Chair*

A. Sharma, *QUT Deputy Vice-Chancellor, Research & Commercialisation*

M. Choi, *Parliamentary Secretary, Queensland Government*

W. van der Aalst, *BPM Conference Steering Committee Chair*

P. Dadam, *BPM 2007 Program Co-Chair*

**09:30 - 10:30**

**Keynote Presentation**

S. Dale, Senior Vice President and Chief Technology Officer, SAP Asia-Pacific

*Holistic BPM: From Theory to Reality*

**11:00 - 12:30**

**Conference Session 1 - Business Process Maturity & Performance**

P. Willaert, J. van der Bergh, J. Willems, D. Deschoolmeester:

*The Process-Oriented Organisation: A Holistic View, Developing a Framework for Business Process Orientation Maturity*

S. Corea, A. Watters

*Challenges in Business Performance Measurement: The Case of a Corporate IT Function*

H. Reijers, M. Song, B. Jeong  
*On the Performance of Workflow Processes with Distributed Actors: Does Place Matter?*

**13:30 - 15:00**

**Conference Session 2 - Business Process Modeling**

H. Reijers, J. Mendling, J. Cardoso  
*What Makes Process Models Understandable?*

Chr. Wolter, A. Schaad  
*Modeling of Task-Based Authorization in BPMN*

M. Magnani, D. Montesi  
*BPMN: How Much Does It Cost? An Incremental Approach*

R. Bobrik, M. Reichert, T. Bauer  
*View-Based Process Visualization*

**15:30 – 17:30**

**Conference Session 3 - Case Studies**

S. Brahe  
*BPM on top of SOA: Experiences from the Financial Industry*

D. Heckl, J. Moormann  
*Matching Customer Processes with Business Processes of Banks: The Example of Small and Medium-Sized Enterprises as Bank Customers*

H. Reijers, M. Jansen-Vullers, M. zur Muehlen, W. Appl  
*Workflow Management Systems + Swarm Intelligence = Emergency Management Systems*

F. Stäber, J.P. Müller  
*Evaluating Peer-to-Peer for Loosely Coupled Business Collaboration: A Case Study*

***Mittwoch, 26. September 2007***

**08:45 - 09:00**

**Welcome**

M. Rosemann, *QUT*  
D. Carrington, *IDS Australia*

**09:00 - 10:00**

**Keynote Presentation**

S. Tieman, Vice President & Director, Estee Lauder, USA  
*Business Process Management: How to Close the Gap Between Rigor and Relevance*

**10:30 - 12:30**  
**Conference Session 4 - Compliance and Change**

S. Sadiq, G. Governatori, K. Namiri  
*Modeling Control Objectives for Business Process Compliance*

J. Kuester, K. Ryndina, H. Gall  
*Generation of Business Process Models for Object Life Cycle Compliance*

M. de Leoni, M. Mecella, G. de Giacomo  
*Highly Dynamic Adaptation in Process Management Systems through Execution Monitoring*

X. Zhao, C. Liu  
*Version Management in the Business Process Change Context*

**13:30 - 15:00**  
**Conference Session 5 - Process Configuration and Execution**

J. Nitzsche, T. van Lessen ,  
D. Karastoyanova, F. Leymann  
*BPEL light*

A. Ottensooser, A. Fekete  
*An Enactment-Engine Based on Use-Cases*

A. Lapouchnian, Y. Yu, J. Mylopoulos  
*Requirements-Driven Design and Configuration Management of Business Processes*

F. Gottschalk, W. van der Aalst, M.H. Jansen-Vullers  
*SAP WebFlow Made Configurable: Unifying Workflow Templates into a Configurable Model*

**15:30 - 16:30**  
**Oracle Keynote**

**16:30 - 17:30**  
**Panel Discussion - BPM 2010: Between Rigor and Relevance**

**10:30 - 12:30**  
**BPM Practice Day Welcome**  
Welcome to BPM Practice Day,  
R. Tregear

**BPM Practice Session 1:**  
**A Day in The Life of a Process Owner**  
B. Frew, Commonwealth Bank, Sydney, Australia

**13:30 – 15:00**  
**BPM Practice Session 2**

**15:30 – 16:30**  
**BPM Practice Session 3**

**Donnerstag, 27. September 2007**

**08:45 - 09:00**

**Welcome**

P. Dadam

**09:00 - 10:00**

**Keynote Presentation**

**10:30 - 12:30**

**Conference Session 6  
Formal Foundations of BPM**

N. Lohmann, P. Massuthe, K. Wolf  
*Behavioural Constraints for Services*

K. Bhattacharya, C. Gerede, R. Hull, R. Liu,  
J. Su  
*Towards Formal Analysis of Artifact-Centric  
Business Process Models*

G. Decker, M. Weske  
*Local Enforceability in Interaction Petri Nets*

K. van Hee, A. Serebrenik, N. Sidorova, M.  
Voorhoeve, J.M. van der Werf  
*Modeling with History-Dependent Petri Nets*

**10:30 - 12:30**

**BPM Tutorial - Business Process  
Management Standards**

M. zur Muehlen

**13:30 - 15:30**

**Conference Session 7  
Business Process Mining**

Chr. W. Günther, W. van der Aalst  
*Fuzzy Mining - Adaptive Process  
Simplification Based on Multi-Perspective  
Metrics*

E. Lamma, P. Mello, M. Montali, F. Riguzzi,  
S. Storari  
*Inducing Declarative Logic-Based Models  
from Labeled Traces*

D. Ferreira, M. Zacarias, M. Malheiros, P.  
Ferreira  
*Approaching Process Mining with  
Sequence Clustering: Experiments and  
Findings*

S. Mauser, R. Bergenthum, J. Desel, R.  
Lorenz  
*Process Mining Based on Regions of  
Languages*

**13:30 - 15:30**

**BPM Tutorial - Lifecycle Management  
for Dynamic Processes**

B. Weber  
S. Sadiq  
M. Reichert

**16:00 – 17:30**  
**Conference Session 8 - Semantic Issues in BPM**

SAP/HP Invited Talk  
*Deriving IT Configurations from Business Processes*

J. Recker, M. Indulska, P. Green  
*Extending Representation Analysis: BPMN User and Developer Perspectives*

P. Soffer, Y. Wand, M. Kaner  
*Semantic Analysis of Flow Patterns in Business Process Modeling*

A. Rodriguez, E. Fernandez-Medina, P. Mario  
*Towards CIM to PIM Transformation: From Secure Business Processes Defined in BPMN to Use Cases*

**16:00 - 17:30**  
**Demo Session**

K. Göser, M. Jurisch, H. Acker, U. Kreher, M. Lauer, S. Rinderle-Ma, M. Reichert, P. Dadam  
*Next-generation Process Management with ADEPT2*

A. Bog, F. Puhlmann, M. Weske  
*The PiVizTool: Simulating Choreographies with Dynamic Binding*

K. Ryndina, J. M. Kuster, H. Gall  
*A Tool for Integrating Object Life Cycle and Business Process Modelling*

J. Mendling  
*Combining a Reachability Graph and a Reduction Rule Approach for Verification of EPCs*

A. Erradi, P. Maheshwari, V. Tosic  
*Policy-Driven Middleware for Manageable and Adaptive Web Services Compositions*

**17:30**  
**Close of the BPM 2007 Conference**



## Call for Paper Modellierung 2008

<http://www.modellierung2008.org/>

12. bis 14. März - Berlin

Die "Modellierung" ist eine Fachtagung des Querschnittfachausschusses Modellierung der GI, die seit 1998 abgehalten wird. Gesucht werden Originalbeiträge hoher Qualität, die sich mit den Grundlagen, den Anwendungen oder den Werkzeugen der Software-, der Daten- oder der Unternehmensmodellierung befassen. Gebietsübergreifende Beiträge, insbesondere zwischen Theorie und Praxis, und die Präsentation neuer Ideen werden dem Charakter der Tagung entsprechend gern gesehen. Die Tagungssprache ist deutsch; Beiträge englischsprachender Konferenzteilnehmer sind jedoch gleichermaßen willkommen.

Die Tagung zeichnet sich durch lebendige Diskussionen und engagierte Rückmeldungen aus Wissenschaft und Praxis aus. Besondere Themen und Interessen werden durch spezielle Workshops abgedeckt. Zum traditionell rege besuchten Industrieforum kommt 2008 ein Doktorandensymposium hinzu, bei dem Nachwuchswissenschaftlern die Gelegenheit gegeben wird, vor einem engagierten Kollegium ihre Ideen und Forschungsansätze vorzutragen und so richtungsweisende Ratschläge zu erhalten. Um darüber hinaus die Beteiligung junger Wissenschaftler an der Tagung zu erhöhen, wird für den besten Beitrag, der aus einer Abschlussarbeit hervorgegangen ist, ein Preis vergeben werden.

Unter anderem die folgenden Themen werden durch die Tagung abgedeckt:

### Grundlagen

- Formale Grundlagen der Modellbildung
- Logik-basierte Modelle
- Ontologien in der Modellierung
- Modelltransformation
- Modellintegration
- Muster in der Modellbildung
- Generative Modelle
- Modellverifikation und -validierung
- Sicherung der Modellkonsistenz

### Modellierungsdomänen und Anwendungen

- Modelle für spezielle Anwendungsbereiche
- Unternehmensanwendungsintegration

### Methoden und Techniken

- Standardisierung
- Model Engineering
- Model-Driven Architecture
- Model-Driven Development
- Ganzheitliche Modellentwickl.
- Modellierung zur Verhaltensverifikation
- Aspektorientierte Modellierung
- Komponentenbasierte Modellierung
- Modellierung von Vorgehensweisen und Prozessen
- Anforderung an Modellierungsansätze und -sprachen
- Modellierung im Kontext von Zeit, Kosten, Qualität und Sicherheit
- Modellierung von IT-Landschaften
- Modellbasiertes Reengineering

- Geschäftsprozessmodellierung
- Benutzermodellierung
- Modellierung sicherheitskritischer Systeme
- Management und Qualitätssicherung von und mit Modellen
- Modellierung verteilter Handels- oder Planungssysteme
- Modellierung und IT-Controlling
- Refactoring von Modellen
- Integration der Evolution von Modellen in Vorgehensmodelle

#### Modellierungswerkzeuge

- Modellierungswerkzeuge
- Werkzeugentwicklung
- Modellanimation und -simulation
- Ausführbare Modelle

#### Sonstige

- Modellbildung in Ausbildung und Wissenstransfer
- Fallstricke der Modellierung

#### Publikation

Der Tagungsband wird in der Reihe GI Lecture Notes in Informatics (LNI) publiziert. Beiträge werden in Indizes wie dem DBLP gelistet. Autoren entsprechend geeigneter Arbeiten werden eingeladen, diese zur Einreichung bei Zeitschriften (Informatik – Forschung und Entwicklung, Software & System Modeling etc.) zu überarbeiten.

#### Ort der Tagung

Die Modellierung 2008 ist an der Humboldt-Universität zu Berlin zu Gast. Deren Campus Adlershof bietet ideale räumliche Voraussetzungen in Europas modernstem Technologiepark. Die Abende werden in den historischen Gebäuden der Humboldt-Universität gestaltet mit einem Empfang im Senatssaal der Universität und dem Konferenzdinner im Thaersaal.

#### Kategorien von Einreichungen

**Hauptkonferenz:** Als Einreichungen werden Langbeiträge (bis zu 16 Seiten) und Kurzbeiträge (bis zu 8 Seiten) in den Kategorien „Forschungsbeitrag“ und „Erfahrungsbericht“ akzeptiert. Alle Einreichungen werden von mindestens drei Gutachtern begutachtet; Arbeiten, die nicht gewissen Mindestqualitätsstandards genügen — dazu zählen neben dem Einhalten der Formatierungsregeln auch sprachliche Standards wie Rechtschreibung, Grammatik etc.; diese gelten auch für englischsprachige Einreichungen — können jedoch ohne weitere Begutachtung abgelehnt werden. Autoren wird die Gelegenheit gegeben, sich zu den ihre Einreichungen betreffenden Gutachten innerhalb von drei Tagen nach Erhalt derselben kurz (maximal 250 Wörter) zu äußern. Stellungnahmen werden von den Vorsitzenden des Programmkomitees entgegengenommen und, wenn sie sachlich begründet sind, an die betreffenden Gutachter weitergeleitet.

**Workshops:** In Workshops werden spezielle Themen im Kreis Interessierter mit starken formlosen Anteilen erörtert und diskutiert.

**Doktorandensymposium:** Mit dem Doktorandensymposium soll jungen Wissenschaftlern die Möglichkeit gegeben werden, ihr Vorhaben außerhalb des eigenen Umfelds zu präsentieren und so wertvolle Anregungen zu erhalten.

**Industrieforum:** Das Industrieforum bietet die Gelegenheit, unbeachtet wissenschaftlicher Kriterien über die industrielle Anwendung und Umsetzung von

Modellierungskonzepten, -techniken und -werkzeugen, den daraus gewonnenen Erfahrungen sowie insbesondere über aktuelle Probleme und Lösungsansätze zu berichten und sich auszutauschen. Einreichungen können in Textform oder als Foliensammlungen vorgenommen werden und werden durch die Industrieforum-Vorsitzenden ausgewählt.

### Wichtige Termine

- Einreichung von Beiträgen/Vorschlägen für Workshops: 7.10.2007
- Mitteilung der Gutachten: 27.11.2007
- Stellungnahme der Autoren: 30.11.2007
- Mitteilung der Entscheidung über die Annahme: 21.12.2007
- Einreichung der Endfassung: 11.1.2008
- Tagung: 12. – 14.3.2008

### Tagungsleitung

- Organisatorische Leitung: W. Reisig
- Öffentlichkeitsarbeit: Y. Sure
- Leitung des Programmkomitees: T. Kühne, F. Steimann
- Doktorandensymposium: R. Breu, A. Pretschner
- Workshops: M. Nüttgens
- Industrieforum: H. Biskup, B. Schätz

### Programmkomitee

- U. Assmann, TU Dresden
- C. Atkinson, Uni Mannheim
- Th. Baar, Swiss Federal Institute of Technology, Lausanne
- B. Bartsch-Spörl, BSR GmbH, München
- R. Breu, Uni Innsbruck
- J. Desel, Katholische Uni Eichstätt
- J. Ebert, Uni Koblenz-Landau
- G. Engels, Uni Paderborn
- U. Frank, Uni Duisburg-Essen
- M. Glinz, Uni Zürich, CH
- M. Gogolla, Uni Bremen
- N. Gronau, Uni Potsdam
- W. Hesse, Uni Marburg
- R. Hirschfeld, HPI Potsdam
- H. Hussmann, Uni München
- M. Jarke, RWTH Aachen
- J. Jürjens, TU München
- R. Kaschek, Massey University, NZ
- R. Kneuper, Darmstadt
- C. Kop, Uni Klagenfurt
- H. Lichter, RWTH Aachen
- P. Liggesmeyer, Uni Kaiserslautern
- H.C. Mayr, Uni Klagenfurt
- M. Minas, Uni der Bundeswehr München
- G. Müller-Luschnat, FAST GmbH, München
- M. Nüttgens, Uni Hamburg
- A. Oberweis, Uni Karlsruhe
- E. Ortner, TU Darmstadt
- B. Paech, Uni Heidelberg
- J. Philipps, validas AG, München
- U. Pletat, IBM Deutschland
- K. Pohl, Uni Duisburg-Essen
- A. Pretschner, ETH Zürich
- U. Reimer, University of Applied Sciences St. Gallen
- W. Reisig, Humboldt-Uni zu Berlin
- M. Riebisch, TU Ilmenau
- B. Rumpe, TU Braunschweig
- B. Schätz, TU München
- P. Schmitt, Uni Karlsruhe
- A. Schürr, TU Darmstadt
- S. Strahringer, European Business School
- M. von der Beeck, BMW AG
- G. Vossen, Uni Münster
- G. Wagner, TU Cottbus
- M. Weske, HPI Potsdam
- A. Winter, Uni Koblenz-Landau
- H. Züllighoven, Uni Hamburg
- A. Zündorf, Uni Kassel

+-----+  
**Fourth GI-Workshop XML4BPM**  
 XML Integration and Transformation for Business Process Management  
 held as a track of Multikonferenz Wirtschaftsinformatik 2008 in Munich (Germany),  
 28 February - 02 March 2008.  
<http://wi.wu-wien.ac.at/home/mendling/XML4BPM2008/>  
 +-----+

**Call for Papers**

XML technologies and applications have attracted increasing attention in the business process management (BPM) community in recent years. Integration aspects are becoming more and more important in this context. On the one hand, there are several interchange formats available to facilitate the exchange of business process models between tools and applications of different focus. These formats are well suited to serve as input to model-driven approaches for engineering of process-aware applications. XML transformations of process models are required to support these approaches in practice. On the other hand, web service technology and semantic web applications have a huge potential for run-time integration of process-aware applications and enterprise applications. These two trends require a good understanding of BPM- and related XML-standards as well as their interrelations. The workshop is dedicated to these two trends of business process management and business process modelling. Papers on ongoing and completed research, state of the art surveys, and reports on practical application of XML technologies in business process management and modelling are welcome. The workshop aims to identify current research directions as well as industry trends.

**Topics** include, but are not limited to:

- + Transformation of BPM-models and -schemas,
- + Model-driven development of BPM applications,
- + Integration of BPM applications,
- + Application of Web Services and Semantic Web technologies for BPM,
- + Metamodels, XML schemas, and ontologies for BPM,
- + Definition and application of XML-based reference models for BPM,
- + Evaluation and comparison of BPM standards,
- + BPEL, WS-CDL, BPSS, PNML, EPML, XPDL, XMI, etc. and their application,
- + Inter-organizational document exchange (e.g. XML-EDI, xCBL, etc.).

**Important Dates:**

Submission: 30 September 2007  
 Notification: 16 November 2007  
 Final Version: 21 December 2007  
 Workshop: between 26 and 28 February 2008

**Submission:**

Papers must be written in English. Each paper will be reviewed by at least two program committee members. Please submit your paper via the submission system on the easychair website. Moreover, please use the guidelines of Lecture Notes in Informatics (LNI) to format your paper. We recommend the use of the Latex package available from the Lecture Notes in Informatics (LNI) website. There are also authors' guidelines and a word template available. Submissions should not exceed 12 pages, including an abstract of up to 150 words and a list of references. Accepted papers will be published as part of the Multikonferenz Wirtschaftsinformatik 2008 proceedings in the LNI series. Similar to the years before, we aim to collect best papers for publication in a special issue of an international journal.

**Program Committee:**

Schahram Dustdar, TU Vienna, A  
 Rony Flatscher, WU Vienna, A  
 Christian Günther, TU Eindhoven, NL  
 Martin Hepp, DERI Innsbruck, A  
 Christian Huemer, TU Vienna, A  
 Ekkart Kindler, DTU, DK  
 Harald Kühn, BOC GmbH, A  
 Frank Leymann, Uni Stuttgart, D  
 Jan Mendling (Co-Chair), QUT Brisbane, AUS

Markus Nüttgens (Co-Chair), Uni Hamburg, D  
 Andreas Oberweis, Uni Karlsruhe (TH), D  
 Manfred Reichert, Uni Twente, NL  
 Oliver Thomas, Uni Saarland, D  
 Ingo Weber, SAP Research, D  
 Mathias Weske, HPI Potsdam, D  
 Andreas Winter, Uni Mainz, D  
 Michael zur Muehlen, Stevens Institute of Technology, USA

**Organisation:**

Volodymyr Klochkov, Uni Hamburg, Germany

# ISTA 2008 Call for Papers

## 7th International Conference on Information System Technology and its Applications

University of Klagenfurt, Austria, April 22 - 25, 2008

The ISTA conference series was founded in 2001 as a joint effort of universities in Klagenfurt and Kharkiv. It soon became a well-respected event joining international researchers in the development, implementation, application and improvement of information systems. After a trip around the globe (i.e. being held in Austria, New Zealand, Ukraine, and USA) ISTA 2008 is back again to Klagenfurt.

### SCOPE

ISTA 2008 aims at bringing together practitioners, researchers, users, and information providers involved in commercial or industrial information systems to exchange ideas, present their recent research and discuss future collaboration. We invite submission of papers focusing on technology for subjects like the following ones:

- the information systems life cycle,
- mobile, pervasive, and ubiquitous information systems,
- components, COTS, and services,
- semantic Web and knowledge intensive systems,
- ontologies,
- e-learning,
- homeland security, and
- natural language processing.

All submitted papers will be peer-reviewed according to their relevance, significance, correctness, originality, and clarity. Submissions must represent novel contributions that neither have been published nor are submitted or under review for publication elsewhere. It is intended to publish the conference proceedings with Springer.

### ISTA 2008 KEY DATA

Abstract submission	12 Oct 2007	Camera ready paper	24 Jan 2008
Paper submission	1 Nov 2007	Early bird Registration	24 Jan 2008
Author notification	17 Dec 2007	Conference	22 – 25 Apr 2008

**Full information regarding ISTA 2008 is available via the UNISCON 2008 Web site:**

[http://www.uniscon2008.org/index.php?option=com\\_frontpage&Itemid=1](http://www.uniscon2008.org/index.php?option=com_frontpage&Itemid=1) .

## **PROGRAM COMMITTEE CO-CHAIRS:**

Roland Kaschek, Massey University, New Zealand  
Günther Fliedl, Alpen-Adria Universität Klagenfurt, Austria  
Mykola Tkachuk, National Technical University "KhPI", Kharkiv, Ukraine

## **PROGRAM COMMITTEE**

Witold Abramowicz, The Poznan Univ. of Economics, Poland  
Marko Anzelak, Alpen-Adria-Universität Klagenfurt, Austria  
Peter Bellström, Karlstad University, Sweden (pending)  
Stephane Bressan, National University of Singapore  
Ruth Breu, University of Innsbruck, Austria  
Andrew Burton-Jones, University of British Columbia, Canada (pending)  
Lois Delcambre, Portland State University, USA  
Anatoly Doroshenko, Institute of Software Systems, Ukraine  
Vadim Ermolayev, Zaporozhye State University, Ukraine  
Alexander Gelbukh, National Polytechnic Institute, Mexico  
Robert Goodwin, Flinders University, Adelaide, Australia  
Remigijus Gustas, Karlstad University, Sweden  
Jean Luc Hainaut, University of Namur, Belgium (pending)  
Wolfgang Hesse, Philipps University Marburg, Germany  
Wolfgang H. Janko, Univ. of Ec. And Bus. Admin., Austria  
Roland Kaschek, Massey University, New Zealand  
Jörg Kerschbaumer, Alpen-Adria-Univ. Klagenfurt, Austria  
Myoung Ho Kim, KAIST, Korea  
Kinshuk, Athabasca University, Canada  
Mikhail R. Kogalovsky, Market Economy Institute. RAS, Russia  
Christian Kop, Alpen-Adria-Universität Klagenfurt, Austria  
Stephen W. Liddle, Brigham Young University, Provo, USA  
Elisabeth Metais, Laboratory CEDRIC, Paris, France  
Igor Nekrestyanov, St. Petersburg State University, Russia  
Boris A. Novikov, St. Petersburg State University, Russia  
Andreas Oberweis, University of Frankfurt, Germany  
Maria E. Orłowska, Polish-Japanese Institute of IT, Warsaw, Poland  
Wolfgang Reisig, Humboldt-Universität zu Berlin, Germany  
Gunter Saake, Otto-von-Guericke-Univ, Magdeburg, Germany  
Demetrios G. Sampson, Univ. of Piraeus, Greece  
Joachim Schmidt, TU Hamburg Harburg, Germany(pending)  
Johannes Siedersleben, sdm München, Germany(pending)  
Vladimir Shekhovtsov, Nat. Technical Univ., Kharkiv, Ukraine  
Keng Siau, University of Nebraska-Lincoln, USA  
Markus Stumptner, University of South Australia, Australia  
Yuzura Tanaka, Hokkaido University, Japan  
Helmut Thoma, IBM Global Services, Switzerland  
Jürgen Vöhringer, Alpen-Adria-Universität Klagenfurt, Austria  
Georg Weber, Infineon Technologies, Austria  
Vyacheslav Wolfengagen, JurInfoR-MSU, Russia  
Joze Zupancic, University of Maribor, Slovenia

## Sprecher der mit der EMISA assoziierten Arbeitskreise und Fachgruppen

<p><b>Prof. Dr. Jörg Desel</b> <i>FG Petrinetze und verwandte Systemmodelle</i></p> <p>Katholische Universität Eichstätt Lehrstuhl für Angewandte Informatik Ostenstr. 14 D-85072 Eichstätt Tel. 08421-93-1712 Fax 08421-93-1789 E-Mail: joerg.desel@ku-eichstaett.de</p>	<p><b>Prof. Dr. Ulrich Frank</b> <i>FG Modellierung betriebl. Informationssysteme</i></p> <p>Universität Duisburg-Essen Lehrstuhl für Wirtschaftsinformatik und Unternehmensmodellierung Universitätsstr. 9 D-45141 Essen Tel.: +49(201) 183 4042 Fax: +49(201) 183 4011 E-Mail: ulrich.frank@uni-duisburg-essen.de</p>
<p><b>Horst Kremers</b> <i>Comittee on Data for Science and Technology (CODATA)</i></p> <p>Postfach 200548 D-13515 Berlin Tel. 030-3728587 Fax 030-3728587 E-Mail: office@horst-kremers.de</p>	<p><b>Prof. Dr. Markus Nüttgens</b> <i>AK Geschäftsprozessmanagement mit Ereignisgesteuerten Prozessketten</i></p> <p>Universität Hamburg WISO Fakultät, Wirtschaftsinformatik Von-Melle-Park 9 D-20146 Hamburg Tel. 040 42838 2792 Fax 040 42838 2780 E-Mail: markus.nuettgens@wiso.uni-hamburg.de</p>
<p><b>Prof. Dr. Bernd Müller</b> <i>AK Informationssysteme mit Open Source</i></p> <p>Fachhochschule Braunschweig/Wolfenbüttel Fachbereich Informatik Am Exer 2 D-38302 Wolfenbüttel Tel. 05331 939 6313 Fax. 05331 939 6002 E-Mail: bernd.mueller@fh-wolfenbuettel.de</p>	<p><b>Dr.-Ing. Eike Schallehn</b> <i>AK Grundlagen von Informationssystemen</i></p> <p>Otto-von-Guericke-Universität Magdeburg Fakultät für Informatik, AG Datenbanken Postfach 4120 D-39016 Magdeburg Tel.: ++49/391/67-12845 Fax: ++49/391/67-12020 E-Mail: eike@iti.cs.uni-magdeburg.de</p>

# EMISA Leitungsgremium

Das EMISA Leitungsgremium besteht aus den Sprechern der assoziierten Arbeitskreise und Fachgruppen (siehe vorige Seite) sowie den folgenden gewählten Mitgliedern:

<p><b>Prof. Dr. Mathias Weske</b> (<i>Sprecher des Leitungsgremiums</i>)</p> <p>Universität Potsdam Hasso-Plattner-Institut für Softwaresystemtechnik Prof. Dr. Helmert-Str. 2-3 D-14482 Potsdam Tel. 0331-5509-191 Fax 0331-5509-189 E-Mail: Mathias.Weske@hpi.uni-potsdam.de</p>	<p><b>Prof. Dr. Manfred Reichert</b> (Stellv. Sprecher, Redaktion EMISA FORUM)</p> <p>University of Twente Faculty of Electrical Engineering, Mathematics &amp; Computer Science Department Databases / Information Systems P.O. Box 217 7500 AE Enschede, Niederlande Tel: ++31 53 489 3705 E-Mail: m.u.reichert@cs.utwente.nl</p>
<p><b>Fernand Feltz</b></p> <p>Centre de Recherche Public - Gabriel Lippmann Département Informatique, Systèmes et Collaboration (ISC) 41, rue du Brill, L-4422 Belvaux, Luxembourg Tel. +352-470261-600, Fax +352-470264 E-Mail: feltz@lippmann.lu</p>	<p><b>Prof. Dr. Andreas Gadatsch</b></p> <p>Fachhochschule Bonn-Rhein-Sieg Fachbereich Wirtschaft Grantham-Allee 20 D-53757 Sankt Augustin Tel. +49-2241-865-129; Fax: +49-2241-865-8129 E-Mail: Andreas.Gadatsch@fh-bonn-rhein-sieg.de</p>
<p><b>Dr. Dominik Kuropka</b></p> <p>Universität Potsdam Hasso-Plattner-Institut für Softwaresystemtechnik Prof. Dr. Helmert-Str. 2-3 D-14482 Potsdam Tel. 0331-5509-193; Fax 0331-5509-189 E-mail: Dominik.Kuropka@hpi.uni-potsdam.de</p>	<p><b>Prof. Dr. Andreas Oberweis</b></p> <p>Universität Karlsruhe Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB) D-76128 Karlsruhe Tel. 49 (721) 608 4516; Fax. +49 (721) 608 4548 E-mail: oberweis@aifb.uni-karlsruhe.de</p>
<p><b>Dr. Hansjürgen Paul</b></p> <p>Institut Arbeit und Technik Abteilung Produktionssysteme Munscheidstraße 14 D-45886 Gelsenkirchen Tel. 0209-1707229; Fax 0209-1707245 Email: paul@iatge.de</p>	<p><b>Dr. Stefanie Rinderle</b></p> <p>Universität Ulm Institut für Datenbanken und Informationssysteme D-89069 Ulm Tel. +49 731 50 24229; Fax +49 731 50 24134 Email: stefanie.rinderle@uni-ulm.de</p>
<p><b>Professor Dr. Gottfried Vossen</b></p> <p>Westfälische Wilhelm-Universität Münster Database Group Leonardo-Campus 3 D-48149 Muenster Tel.: +49 251 83 38150; Fax: +49 251 83 38159 E-Mail: vossen@uni-muenster.de</p>	

# Impressum

Das EMISA FORUM ist das Mitteilungsblatt der Fachgruppe Entwicklungsmethoden für Informationssysteme und deren Anwendung (EMISA) im Fachausschuss Datenbanken und Informationssysteme der Gesellschaft für Informatik. Das EMISA FORUM erscheint zweimal jährlich, und zwar im Januar und im August. Es wird den Mitgliedern der Fachgruppe zugesandt. Mitglied kann jeder werden, der sich für die Zielsetzungen der Fachgruppe interessiert und sich mit einer stichwortartigen Angabe seiner Arbeitsgebiete anmeldet. Es wird ein jährlicher Mitgliedsbeitrag von 10 € für GI-Mitglieder und 16.50 € für Nicht-GI-Mitglieder erhoben. Durch das Mitteilungsblatt sollen aktuelle Informationen unter den Mitgliedern der Fachgruppe ausgetauscht werden. Die Herausgeber bitten daher alle Leser, sich möglichst rege an der Gestaltung des Blattes zu beteiligen. Es können für die Fachgruppe relevante Beiträge unter anderem zu folgenden Rubriken eingereicht werden:

- Fachbeiträge
- Berichte aus der Praxis
- Projektberichte (auch über laufende Projekte)
- Aktuelle Schlagwörter
- Vorstellung von Arbeitsgruppen
- Buchbesprechungen
- Aktuelle Publikationen, Dissertationen, Habilitationen
- Tagungsberichte
- Call for Papers, Einladungen, Programme
- Leserbriefe, E-Mail-Zuschriften

Mit der Zusendung eines Beitrags ist das Einverständnis zur Veröffentlichung im EMISA Forum verbunden. Jeder Beitrag wird ohne Begutachtung veröffentlicht (wenn er dem Themenspektrum der EMISA entspricht). Er gibt nur die individuelle Meinung des Autors wieder. Die Beiträge können in deutscher oder englischer Sprache abgefasst sein; sie sollten keine Seitennummerierung enthalten. Redaktionsschluss für die Januar-Ausgabe ist der 3. Dezember, für die August-Ausgabe der 6. Juli (falls nicht anders angekündigt).

**Herausgeber:** Leitungsgremium der Fachgruppe EMISA

**Auflage:** 1200

**Redaktion:** Prof. Dr. Manfred Reichert

University of Twente

Faculty of Electrical Engineering, Mathematics & Computer Science

Department Databases / Information Systems (Zilverling 3096)

P.O. Box 217

7500 AE Enschede, Niederlande

Tel: ++31 53 489 3705

E-Mail: m.u.reichert@cs.utwente.nl

**Redaktionsschluss für das nächste Heft: 3. Dezember 2007**