ICIQ 2009

The 14th International Conference on Information Quality

Co-located with the German Information Quality Management Conference (GIQMC)

November 7-8, 2009.

Hasso Plattner Institute for IT Systems Engineering

University of Potsdam, Germany

The 14th International Conference on Information Quality, 2009

Co-located with the German Information Quality Management Conference (GIQMC)

http://www.iciq2009.org

November 7-8

Hasso Plattner Institute for IT Systems Engineering

Potsdam, Germany

Contents

Welcome Message

Felix Naumann

Dear participants of the 2009 edition of ICIQ,

welcome to the 14th International Conference on Information Quality 2009 (ICIQ'09). Since 1996 the ICIQ conferences were held annually at MIT in Cambridge. Now ICIQ has gone abroad: The Hasso Plattner Institute (HPI) in Potsdam, on the border of Berlin, is proud to host the first edition away from MIT. The HPI is privately funded and dedicated to teaching and research in the area of IT systems engineering. It is affiliated with the University of Potsdam. This year HPI was ranked among the top 4 university computer science institutes in Germany.

ICIQ provides a forum to exchange IQ knowledge and ideas and learn from each other. Practitioners and researchers will present findings and experience on topics such as IQ concepts, IQ management, case studies, best practices, cost/benefit analysis, IQ and data warehousing, IQ and e-business, policies and standards. In addition to the official program, there is plenty of opportunity for informal discussions. Highlights of this year's program include two keynote talks and a dinner talk at the conference banquet. The program features 18 talks grouped into six sessions and a poster session featuring six posters.

On behalf of all the organizers, I wish you a pleasant, insightful, and interesting stay during ICIQ at HPI. We have put together this booklet to provide helpful and interesting information about the conference, the venue, and the location at a glance. Please do not hesitate to contact me or any conference organizer for assistance and questions.

Enjoy your visit,

Felix Naumann

Hasso Plattner Institute at a Glance

A Profile of a Pioneering Institute

Hasso Plattner Institute for IT Systems Engineering (HPI) at the University of Potsdam is unique in Germany because it has been the first institute in Germany financed entirely by private funds, offering a unique degree in "IT Systems Engineering" as an alternative to conventional computer science programs.

Since 1999, HPI has awarded over than 200 Bachelor and more than 50 Master's degrees. The institute teaches more than 330 students in the design, development and control of complex IT systems. There are currently more than 50 professors, lecturers, and staff working at HPI.

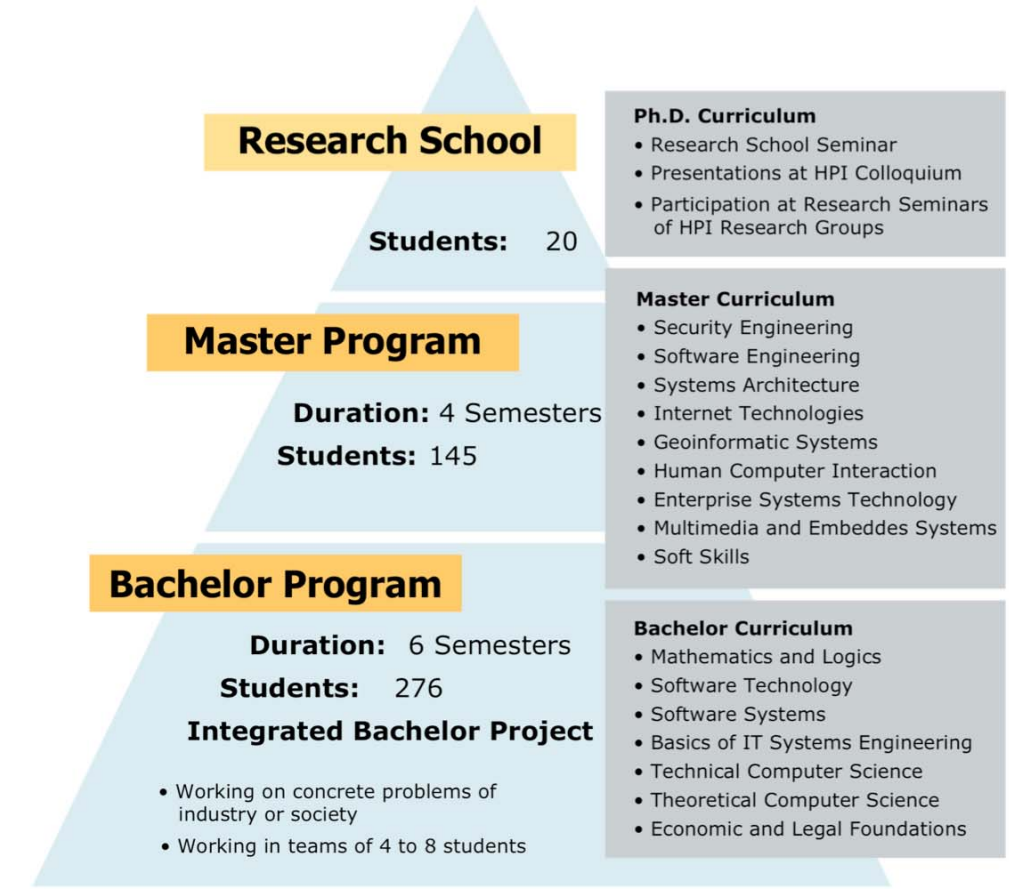
The HPI achieved excellent results in two major evaluations of IT schools in Germany, undertaken by the well-recognized Karriere magazine and the acknowledged Center for Higher Education Development (CHE). Both in the CHE-ranking of May 2006, and in the June 2006 edition of Karriere, HPI was ranked on the fourth place out of more than 100 institutions. Karriere refers to HPI as the "Shooting-Star in computer sciences".

HPI provides an outstanding student-teacher ratio, supported by latest technical and structural equipment. As a result, this pioneering Potsdam institute provides an entire new generation of scientific elites with the best study and research conditions available today.

One of the HPI's greatest priorities is the personal support of each individual student. Each year up to 80 of the best qualified applicants are accepted to the HPI where, in their first semester, they attend a seminar designed to guide them throughout their studies. A maximum of 40 students are accepted into the Master's Program each year. Every student is assigned a professor as a personal mentor; a sufficient number of computer workstations are available in the labs and seminar rooms. Tuition fees are not required.

In October 2005, the HPI started its Research School on "Service-Oriented Systems Engineering", a graduate school based on the model of the DFG (German Research Foundation) "Graduiertenkolleg" (Graduate School), which has more than 25 members.

The School of Design Thinking at the Hasso-Plattner-Institute began its program in the winter term of 2007/2008. Under the directorship of Prof. Ulrich Weinberg, this unique complementary program in Design Thinking (inventive development) is the first of its kind at a German university. Modelled on the famous d.school at Stanford University in California (USA), the one-year program in Design Thinking will enable students to develop particularly user-friendly IT-based products and services in multidisciplinary teams.



Research Groups

- Enterprise Platform and Integration Concepts
- Internet Technologies and Systems
- Operating Systems and Middleware
- Business Process Technology
- Human Computer Interaction
- Computer Graphics Systems
- System Analysis and Modeling
- Software Architecture
- Information Systems
- School of Design Thinking

A total of 71 Ph.D. Students in HPI Research Groups
20 graduates completed a Ph.D.

Status: 11.06.2009

Saturday, 7. Nov. 2009

Time \ Location	Lecture Hall (HS 2)	Lecture Hall (HS 3)
9:00-13:00	GIQMC	
13:00-14:00	Lunch	
14:00-15:00	**Keynote talk Jeff Jonas "Macro Trends in Data and Sensemaking" (HS 1)**	
15:00-15:30	Coffee break	
15:30-17:00	Parallel sessions	
	Session 1 "Master Data Management" Chair: Andy Koronios, University of Southern Australia	**Session 2 "Modeling & Metadata"** Chair: Ahmed Elmagarmid, Purdue University
	Master Data Management: Products and Research	Flexible and Generic Data Quality Metadata Exchange
	Information Management along the life cycle of data and application systems - challenges and solution approaches	Extending BPMN to Support the Modeling of Data Quality Issues
	Master Data Management Processes - A Petri-Net based solution that supports organizations with global and local MDM quality requirements	Data Quality through Conceptual Model Quality - Reconciling Researchers and Practitioners through a Customizable Quality Model
17:00-17:15	Short break	
17:15-18:45	Parallel sessions	
	Session 3 "Techniques for Improving IQ" Chair: Marcus Gebauer, DGIQ	**Session 4 "IQ Metrics & Assessment"** Chair: Elizabeth Pierce, University of Arkansas at Little Rock
	How to Screen a Data Stream - Quality-Driven Load Shedding in Sensor Data Streams	Identification of Business Oriented Data Quality Metrics
	SOG: A Synthetic Occupancy Generator to Support Entity Resolution Instruction and Research	A Framework for Economic-driven Assessment of Data Quality Decisions
	Computing Uncertain Key Indicators from Uncertain Data	A SOA-based Data Quality Assessment Framework in a Medical Science Center
18:45-19:15	Transfer to Exploratorium	
19:15-22:00	Conference banquet (GIQMC & ICIQ)	
	Dinner Speech Richard Wang "Challenges in Advancing Information Quality"	

Sunday, 8. Nov. 2009

Time \ Location	Lecture Hall (HS 2)	Lecture Hall (HS 3)
8:00-9:00	Breakfast	
9:00-9:30	Award Ceremony (HS1)	
9:30-10:30	**Keynote talk Niels Weigel "Lean and Agile Enterprise Information Management" (HS 1)**	
10:30-11:00	Coffee break	
11:00-13:00	Parallel sessions	
	Session 5 "Poster Flash" Chair: John Talburt, University of Arkansas at Little Rock	**Session 6 "Case Studies"** Chair: Boris Otto, University of St. Gallen
	Is DQ/IQ the Quality of Information? Two Views	Information Quality Issues in the Mortgage Banking Industry
	A Meta-model for Data Quality Management Simulation	The Quality of Monitoring Data in Civil Engineering Works
	Schema Based Deduplication	
	A Multi-Dimensional Model for Assessing the Quality of Answers in Social Q&A Sites	Data Quality Evaluation in an E-Business Environment: A Survey
	Multi-Source Object Identification With Constraints	
	ExplainIE - Explaining Information Extraction Systems	Multidimensional Management and Analysis of Quality Measures for CRM Applications at EDF
13:00-14:00	Lunch	
14:00-15:30	Parallel sessions	
	Session 7 "IQ in Web 2.0" Chair: Rolf Wigand, University of Arkansas at Little Rock	**Panel**
	An Empirical Study on Criteria for Assessing Information Quality in Corporate Wikis	
	Towards Assessing Information Quality in Knowledge Management in the Enterprise 2.0	
	Web-based Affiliation Matching	
15:30-15:45	Closing Remarks (HS 1)	

Organization

General Chair

Felix Naumann: Hasso Plattner Institute for IT Systems Engineering, Germany.

Program Chairs

- **Paul Bowen**: Florida State University, USA.
- **Ahmed K. Elmagarmid**: Purdue University, USA.
- **Hubert Oesterle**: University of St Gallen, Switzerland.
- **Kai-Uwe Sattler**: Ilmenau University of Technology, Germany.

Program Committee

- **Carlo Batini:** University of Milan, Italy.
- **Laure Berti-Equille:** University of Rennes, France.
- **Mokrane Bouzeghoub:** Universite de Versailles, France.
- **Ismael Caballero:** University of Castilla-La Mancha, Spain.
- **Tiziana Cartaci:** University of Roma La Sapienza, Italy.
- **Tamraparni Dasu:** AT&T Labs, USA.
- **Helena Galhardas:** Technical University of Lisbon, Portugal.
- **Floris Geerts:** University of Edinburgh, UK.
- **Michael Gertz:** University of Heidelberg, Germany.
- **Eberhard Hechler:** IBM, Germany.
- **Markus Helfert:** Dublin City University, Ireland.
- **Theodore Johnson:** AT&T Labs, USA.
- **Barbara Klein:** University of Michigan-Dearborn, USA.
- **Andrea Maurino:** Universita di Milano Bicocca, Italy.
- **Helinä Melkas:** Lappeenranta University of Technology, Finland.
- **Paolo Missier:** University of Manchester, UK.
- **Boris Otto:** University of St. Gallen, Switzerland.
- **Mourad Ouzzani:** Purdue University, USA.
- **Elizabeth Pierce:** University of Arkansas, Little Rock, USA.
- **Leo Pipino:** University of Massachusetts Lowell, USA.
- **Erhard Rahm:** University of Leipzig, Germany.
- **Friedrich Roithmayr:** University of Linz, Austria.
- **Monica Scannapieco:** University of Roma La Sapienza, Italy.
- **Yasuki Sekiguchi:** Hokkaido University, Japan.
- **Valerie Sessions:** Charleston Southern University, USA.
- **Kai Simon:** Gartner, Germany.
- **Divesh Srivastava:** AT&T Labs, USA.
- **Dan Suciu:** University of Washington, USA.
- **Panos Vassiliadis:** University of Ioannina, Greece.
- **Vassilios Verykios:** University of Thessaly, Greece.
- **Anette Weisbecker:** Fraunhofer Institute, Germany.

ICIQ Steering Committee

- **Felix Naumann:** Hasso Plattner Institute, Germany, (Chair).
- **Bruce Davidson**: Cedars-Sinai Health System, USA.
- **Jing Gao**: University of South Australia.
- **Marcus Gebauer**: DGIQ, Germany.
- **Andy Koronios**: University of Southern Australia.
- **John Talburt**: University of Arkansas at Little Rock, USA.
- **Richard Wang**: MIT, USA.

Proceedings Chair

- **Chistoph Böhm**: Hasso Plattner Institute for IT Systems Engineering, Germany.

Publicity Chair

- **Mohammed AbuJarour**: Hasso Plattner Institute for IT Systems Engineering, Germany.

Local Arrangements Chair

- **Tobias Vogel**: Hasso Plattner Institute for IT Systems Engineering, Germany.

Conference Venue

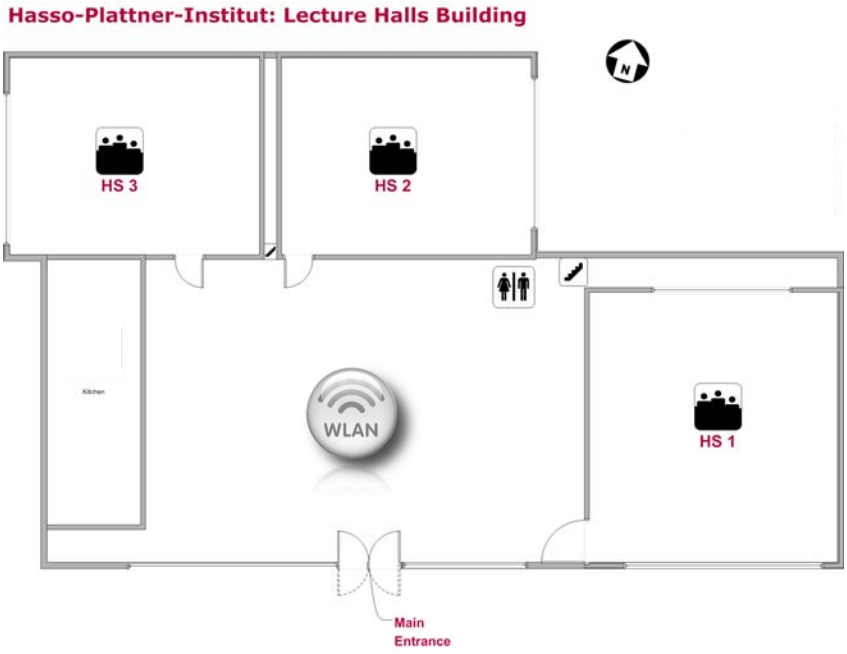
ICIQ 2009 is held at Hasso Plattner Institute, which is located in Potsdam-Babelsberg, on the edge of Berlin not far from the Film and Media Park and in the direct vicinity of a lake, the Griebnitzsee. The 30,000 square meter campus, set in green surroundings, is easily accessible by metropolitan train (S-Bahn) and bus. There are sufficient parking spaces for private cars. The park-like campus with its tall trees, small lake and water fountains offers an inspiring atmosphere for work and ideal surroundings for relaxing. Two buildings - the main HPI complex incorporating the cafeteria and the University of Potsdam's Institute of Informatics - and another building - the lecture-halls building with a total capacity of 490 audience - catch the eye with their simple yet elegant design.

All sessions of the ICIQ 2009 will be held in the lecture halls building of the HPI. This building has three lecture halls (HS1, HS2, and HS3). The plan of the lecture halls building is shown in the figure below. **WLAN** will be available in the conference venue during the conference.

Address:

Hasso Plattner Institute for IT-Systems Engineering
Prof.-Dr.-Helmert-Str. 2-3
D-14482 Potsdam
Germany

Phone: +49 331 5509 280



Getting There

Situated directly next to the Griebnitzsee metropolitan train (S-Bahn) station, the Hasso Plattner Institute in Potsdam-Babelsberg is superbly located in terms of accessibility. On S-Bahn line S7, it is just a short ride to both the capital Berlin and Potsdam's city center. The Griebnitzsee station is also served by local transport bus lines 694 and 696.

The lake-side campus is an impressive oasis of greenery set in beautifully tended grounds. Both the main building and the auditorium (lecture halls) building are located at Prof.-Dr.-Helmert-Straße 2-3.

① Hauptgebäude
- Institutsleitung und -verwaltung
- Fachgebiet Betriebssysteme und Middleware
- Fachgebiet Business Process Technologies
- Fachgebiet Internettechnologien und -systeme
- Fachgebiet Informationssysteme
- Fachgebiet Software-Architekturen
- Fachgebiet Systemanalyse und Modellierung

② HPI-Hörsaalgebäude

① Main Building
- Institute Management
- Operating Systems and Middleware Group
- Business Process Technology Group
- Internet Technologies and Systems Group
- Information Systems Group
- Software Architecture Group
- System Analysis and Modeling Group

② HPI- (Auditorium) Lecture Halls Buidling

Map: "Campus Griebnitzsee"

Directions from Griebnitzsee station

- To the **main building**: Leave the underpass in the direction of "Universität" (university), turn left at the bus loop, pass the level crossing and keep half right.
- To the **auditorium (Lecture Halls) building**: Leave the underpass in the direction of "Universität" (university), turn left at the bus loop, pass the level crossing and keep walking straight.

To find your way around the campus, please take a look at the "Campus Griebnitzsee" map.

Arrival from Berlin airports

- From **Tegel Airport** take bus line 109, destination Bahnhof Zoo, to S Charlottenburg (travel time approx. 15 min). Switch to S-Bahn line S7, destination Potsdam-Hauptbahnhof, and travel to Potsdam-Griebnitzsee (travel time approx. 20 min).
  Total travel time approx. 50 min; ticket for zones Berlin ABC required.
- **From Schönefeld Airport**, there is a footpath leading to the S-Bahn station S Flughafen Berlin-Schönefeld DB (approx. 400 m).

  - From that station, take Regional-Bahn (RB22), destination Potsdam Hauptbahnhof (Hbf), to Potsdam Hbf (travel time approx. 52 min). Switch to line S7 or S1 and travel to Potsdam-Griebnitzsee (travel time approx. 6 min).
    Total travel time approx. 1 hour; ticket for zones Berlin ABC required.
  - From that station, take Regional Express (RE 7), destination Belzig or Dessau, to S Wannsee (travel time approx. 50 min). Switch to line S7 or S1, destination Potsdam-Hauptbahnhof, and travel to Potsdam-Griebnitzsee (one station) (travel time approx. 4 minutes).
    Total travel time approx. 1 hour. Ticket for zones Berlin ABC required.

Arrival by train at Griebnitzsee metropolitan train (S-Bahn) station

- From **Berlin central station (Berlin Hauptbahnhof)** take S-Bahn line S7, destination Potsdam-Hauptbahnhof, and travel to Potsdam-Griebnitzsee (travel time approx. 35 min).
- From **Berlin Zoologischer Garten** take S-Bahn line S7, destination Potsdam-Hauptbahnhof, and travel to Potsdam-Griebnitzsee (travel time approx. 25 min).
- From **Potsdam central station (Potsdam Hauptbahnhof)** take S-Bahn line S7 or S1, destination Ostbahnhof or Potsdamer Platz , to Potsdam-Griebnitzsee (travel time approx. 10 min).

A taxi ride from Potsdam central station takes around 15 minutes and costs approx. 10 Euros.

Arrival by Taxi

We have negotiated a pick-up service and flat rates for the two Berlin Airports and train stations. Call *Taxi Kaiser* of Potsdam toll-free from Germany at 0800 600 1711 (+49 331 600 1711 from other countries) and mention "HPI".

- **Tegel airport**: 30€ (40 minutes)
- **Schönefeld airport**: 45€ (60 minutes)
- **Berlin Hauptbahnhof** (Central Station): 45€ (35 minutes)

Arrival by car

- From the A10 (Berlin beltway), change onto the A115 at the intersection "Dreieck Nuthetal".
- Leave the A115 at the Potsdam-Babelsberg exit (dual exit!) and continue in the direction of Potsdam-Babelsberg.
- Follow Nuthestraße towards the city center ("Zentrum") and exit at Wetzlarer Straße (Medienstadt Babelsberg).
- Turn right, pass the BMW car dealer and turn left in the direction of Babelsberg.
- At the next intersection (Großbeerenstraße/Wetzlarer Straße) keep going straight onto August-Bebel-Straße.
- After approx. 1.5 km, just before the rail underpass, turn right onto Prof.-Dr.-Helmert-Straße. At the end of the street (next to the bus loop) you will see the HPI on your right side.

In case your car navigation system cannot identify Prof.-Dr.-Helmert-Straße, please enter the following address instead: Stahnsdorfer Straße 156A, 14482 Potsdam.

Keynote Talks

We are very proud to present three distinguished invited speakers at ICIQ 2009. Jeff Jonas will hold the opening keynote for ICIQ on Saturday; Niels Weigel will open Sunday with his keynote talk; and Richard Wang will give a dinner speech at the conference banquet.

Jeff Jonas

Chief Scientist
IBM Entity Analytics Group
IBM Distinguished Engineer

"Macro Trends in Data and Sensemaking"
Saturday 2:00-3:00 p.m.

Richard Wang

Founder of ICIQ
Director of TDQM, MIT
Chief Data Quality Officer, US Army

"Challenges in Advancing Information Quality"
Saturday (dinner speech)

Niels Weigel

Senior Product Manager
Technology Solution Management - Information Management
SAP Business Objects

"Lean and Agile Enterprise Information Management"
Sunday 9:30-10:30 a.m.

Jeff Jonas

IBM

"Macro Trends in Data and Sensemaking"

Saturday 2:00-3:00 p.m.

Bio

Jeff Jonas is Chief Scientist, IBM Entity Analytics Group and an IBM Distinguished Engineer. The IBM Entity Analytics Group was formed based on technologies developed by Systems Research & Development (SRD), founded by Jonas in 1984, and acquired by IBM in January, 2005.

Prior to the acquisition Jonas lead SRD through the design and development of a number of unique systems including technology used by the Las Vegas gaming industry. One such innovation played a pivotal role in protecting the gaming industry from aggressive card count teams. The most notable known as the MIT team featured in the book "Bringing Down the House", and recent movie "21." Today, possibly half of the casinos in the world use technology created by Jonas and his SRD team. This work is frequently featured in documentaries appearing on, the Discovery Channel, Learning Channel and the Travel Channel.

Following an investment in 2001 by In-Q-Tel, the venture capital arm of the CIA; SRD began playing a role in America's national security and counterterrorism mission. One such contribution includes an analysis of the connections between the individual 9/11 terrorists. This link analysis is now taught in universities and has been widely cited by think tanks and the media, for example, an extensive one-on-one interview with Peter Jennings that aired on ABC PrimeTime.

Jonas designs next generation technology that helps organizations better leverage their enterprise-wide information assets. With particular interest in real-time "sensemaking" these innovative systems fundamentally improve enterprise intelligence which makes organizations smarter, more efficient and highly competitive.

Jonas is also somewhat unique as a technologist in that he frequently engages with the privacy and civil liberties community. With responsible innovation in mind, Jonas invented technology which enables organizations to discover records of common interest (e.g., identities) without the transfer of any privacy-invading content. This cryptographic-based technique known as "Anonymous Resolution" delivers new levels of privacy protection in areas of critical interest like; clinical health care research, bio-surveillance, aviation safety, homeland security, fraud detection and identity theft.

Jonas' work has received wide media attention from the Wall Street Journal, The Washington Post, to Fortune Magazine, MSNBC and National Public Radio. A highly sought after speaker, Jonas travels the globe discussing innovation, national security, and privacy with government leaders, industry executives, leading global think tanks, privacy advocacy groups, and policy research organizations, including the Center for Democracy and Technology, Heritage Foundation and the Markle Foundation. He is a Senior Associate at the Center for Strategic and International Studies.

Jonas periodically testifies on privacy and counterterrorism in such venues as the White House before the President's Privacy and Civil Liberties Oversight Board, the Department of Homeland Security's Data Privacy and Integrity Advisory Committee, and other federally convened commissions.

Jonas was briefly a quadriplegic in 1988 following a car accident. Today, he competes in Ironman triathlons around the world. He raised three wonderful children as a single father.

Richard Wang
MIT & US Army

"Challenges in Advancing Information Quality"

Saturday (dinner speech)

Bio

Richard Y. Wang is the Chief Data Quality Officer of the U.S. Army, on leave from MIT Information Quality (MITIQ) Program. He also holds an appointment as University Professor of Information Quality, University of Arkansas at Little Rock. Before heading the MITIQ program, Dr. Wang served as a professor at MIT for a decade. He received a Ph.D. in Information Technology from MIT.

Dr. Wang is the recipient of the 2005 DAMA International Achievement Award. Previous recipients of this award include Codd for inventing the Relational Data model, Chen for the Entity Relationship model, and Inman for data warehousing.

Dr. Wang has extensive interactions with industry and government, serving as an expert consultant, principal investigator, and advisor to execute enterprise data quality, data warehousing, and data governance projects.

Wang has put the term Information Quality on the intellectual map with myriad publications. In 1996, Prof. Wang organized the premier International Conference on Information Quality, which he has served as the general conference chair and currently serves as Chairman of the Board. Wang's books on information quality include Journey to Data Quality (MIT Press, 2006), Information Quality: Advances in Management Information Systems (M.E. Sharpe, 2005), Introduction to Information Quality (MITIQ Publications, 2005), Data Quality (Kluwer Academic, 2001), and Quality Information and Knowledge (Prentice Hall, 1999).

Prof. Wang has been instrumental in the establishment of the Ph.D. and Master of Science in Information Quality degree program at the University of Arkansas at Little Rock, the Stuart Madnick IQ Best Paper Award for the International Conference on Information Quality, the comprehensive IQ Ph.D. dissertations website, and the Donald Ballou & Harry Pazer IQ Ph.D. Dissertation Award.

Niels Weigel
SAP

"Lean and Agile Enterprise Information Management"

Sunday 9:30-10:30 a.m.

Abstract

Over the last decades information management has grown up from a small departmental initiative of individual tasks, such as documentation of database systems, re-active data cleansing, etc., to a broader rocket science of data governance aspects, such as data definitions and shared business vocabulary; metadata management; data modeling, data integration; data quality; data security; master data management; content management; and taxonomy design and maintenance. The growing amount of data and applications in an enterprise; multiple ways of information representation (structured and unstructured); more complicated business processes, where most of them need to be compliant with federal laws; global extension of business; and new local requirements are just some of the drivers to be mentioned to develop an information management strategy that covers all data, all applications, and all process of an enterprise.

Suddenly, during the last year with the financial markets crisis many things changed. It is still obvious to everybody that especially now it is the time to invest into one of the most valuable assets inside an organization, the data and the management of the data. But the pictures of enterprise-wide deployment of information management strategies are really bright in a tunnel that is limited by pressures of cost efficiency and reduced expenses. It is now essential to walk through the tunnel in smaller steps with defined milestones that produce additional value to the organization already during implementation steps. That said, it is time for a Lean and Agile Enterprise Information Management deployment!

Bio

Niels Weigel is Senior Solution Manager at SAP for Data Quality in the Enterprise Information Management Division. Working for about 15 years in several Software Vendor organizations, he has a broad experience on the market demands and requirements for solutions as well as project frameworks to solve the Enterprise Information Management challenges.

After studying Aerospace Engineering at the University of Stuttgart, finishing with his diploma thesis on "Fuzzy logic for adaptive position control and guidance of a solar airship", he identified new challenges in the area of International Data Quality Management at FUZZY! Informatik AG. After initial work in design and development of software solutions, he was leading the Consulting Services organization. As Head of Business Development he was responsible for the Product Management and was also setting up the FUZZY! DataCare Process, a methodology for successfully implementing Data Quality projects within an organization. After the acquisition of FUZZY! Informatik AG by Business Objects in 2007 he joined the Solution Management Team for Enterprise Information Management at the SAP BusinessObjects division and took over responsibility for SAP's International Address Cleansing solutions.

He is member of the board of the German Society for Information and Data Quality (DGIQ) and has successfully completed the IQ-2 Certified Information Quality Professional course at the MIT, Boston. He appeared on several international Information Quality Conferences (ICIQ, AusIQ), the German SAP User Group Conference DSAG and spoke at the University of Arkansas at Little Rock in the series "Distinguished Lecturers In Information Quality"

Research and Poster Papers

This year, 19 full research papers and 6 poster papers will be presented at the International Conference on Information Quality (ICIQ 2009).

Research Papers

Master Data Management

- **Master Data Management: Products and Research**

  Jochen Kokemüller, Fraunhofer IAO, Germany
  Anette Weisbecker, Fraunhofer IAO, Germany

- **Information Management Along the Lifecycle of Data and Application Systems: Challenges and Solution Approaches**

  Georg Fischer, SAP AG
  Axel Herbst, SAP AG

- **Master Data Management Processes - A Petri-Net based solution that supports organizations with global and local MDM quality requirements**

  Kai-Uwe Baryga, SYDECON GmbH, Germany

Modeling & Metadata

- **Flexible and Generic Data Quality Metadata Exchange**

  David Becker, The MITRE Corporation, USA
  John Jaster, Digital Prospectors Corporation, USA
  Jereme Kuperman, Illumination Works, USA

- **Extending BPMN to Support the Modeling of Data Quality Issues**

  Noelia Sánchez-Serrano, Alarcos Research Group, University of Castilla La Mancha, Spain
  Ismael Caballero, Alarcos Research Group, University of Castilla La Mancha, Spain
  Félix Garcían, Alarcos Research Group, University of Castilla La Mancha, Spain

- **Data Quality through Conceptual Model Quality - Reconciling Researchers and Practitioners through a Customizable Quality Model**

  Kashif Mehmood, ESSEC Business School, CEDRIC-CNAM, Paris, France
  Samira Si-Said Cherfi, CEDRIC-CNAM, Paris, France
  Isabelle Comyn-Wattiau, ESSEC Business School, CEDRIC-CNAM, Paris, France

Techniques for Improving IQ

- **How to Screen a Data Stream - Quality-Driven Load Shedding in Sensor Data Streams**

  Anja Klein, SAP Research Center, Dresden, Germany
  Gregor Hackenbroich, SAP Research Center, Dresden, Germany
  Wolfgang Lehner, University of Technology, Dresden, Germany

- **SOG: A Synthetic Occupancy Generator to Support Entity Resolution Instruction and Research**

  John R. Talburt, University of Arkansas at Little Rock, USA
  Yinle Zhou, University of Arkansas at Little Rock, USA
  Savitha Yalanadu Shivaiah, University of Arkansas at Little Rock, USA

- **Computing Uncertain Key Indicators from Uncertain Data**

  Carlos Rodríguez, University of Trento, Italy
  Florian Daniel, University of Trento, Italy
  Fabio Casati, University of Trento, Italy
  Cinzia Cappiello, Politecnico di Milano, Italy

IQ Metrics & Assessment

- **Identification of Business Oriented Data Quality Metrics**

  Boris Otto, University of St. Gallen, Switzerland
  Kai M. Hüner, University of St. Gallen, Switzerland
  Hubert Österle, University of St. Gallen, Switzerland

- **A Framework for Economic-driven Assessment of Data Quality Decisions**

  Adir Even, Ben-Gurion University of the Negev, Israel
  Marcus Kaiser, University of Augsburg, Germany

- **A SOA-based Data Quality Assessment Framework in a Medical Science Center**

  Yao Zhou, Charité Berlin, Germany
  Sabine Hanß, Charité Berlin, Germany
  Malte Cornils, Charité Berlin, Germany
  Claudia Hahn, Charité Berlin, Germany
  Sonja Niepage, Charité Berlin, Germany
  Thomas Schrader, Fachhochschule Brandenburg, Germany

Case Studies

- **Information Quality Issues in the Mortgage Banking Industry**

  Rolf T. Wigand, University of Arkansas at Little Rock, USA
  Jerry Wood, University of Arkansas at Little Rock, USA
  Yusuf Yiliyasi, University of Arkansas at Little Rock, USA

- **The Quality of Monitoring Data in Civil Engineering Works**

  Ana Lucas, Laboratório Nacional de Engenharia Civil, Portugal
  António Palma-dos-Reis, Instituto Superior de Economia e Gestão, Portugal
  Mário Caldeira, Instituto Superior de Economia e Gestão, Portugal

- **Data Quality Evaluation in an E-Business Environment: A Survey**

  Soumaya Ben Hassine-Guetari, Lyon University, France

- **Multidimensional Management and Analysis of Quality Measures for CRM Applications at EDF**

  Verónika Peralta, Université François Rabelais Tours, Université de Versailles, France
  Virginie Thion-Goasdoué, Université de Paris Dauphine, France
  Zoubida Kedad, Université de Versailles, France
  Laure Berti-Équille, Université de Rennes 1, France
  Isabelle Comyn-Wattiau, CNAM-CEDRIC, Paris, France
  Sylvaine Nugier, EDF R&D, France
  Samira Si-said-Cherfi, CNAM-CEDRIC, Paris, France

IQ in Web 2.0

- **An Empirical Study on Criteria for Assessing Information Quality in Corporate Wikis**

  Therese Friberg, University of Paderborn, Germany
  Wolfgang Reinhardt, University of Paderborn, Germany

- **Towards Assessing Information Quality in Knowledge Management in the Enterprise 2.0**

  Sven Ahlheid, Siemens AG, Germany
  Therese Friberg, University of Paderborn, Germany
  Gernot Graefe, Siemens AG, Germany
  Alexander Krebs, Siemens AG, Germany
  Jan-Philipp Müller, Siemens AG, Germany
  Dirk Schuster, Siemens AG, Germany

- **Web-based Affiliation Matching**

  David Aumueller, University of Leipzig, Germany
  Erhard Rahm, University of Leipzig, Germany

Poster Papers

Poster Papers

- **Is DQ/IQ the Quality of Information? Two Views**

  Zbigniew J Gackowski, California State University Stanislaus, USA

- **A Meta-model for Data Quality Management Simulation**

  Boris Otto, University of St. Gallen, Switzerland
  Kai M. Hüner, University of St. Gallen, Switzerland

- **Schema Based Deduplication**

  Pei Li, University of Milan, Bicocca, Italy
  Andrea Maurino, University of Milan, Bicocca, Italy

- **A Multi-Dimensional Model for Assessing the Quality of Answers in Social Q&A Sites**

  Zhemin Zhu, TU Darmstadt, Germany
  Delphine Bernhard, TU Darmstadt, Germany
  Iryna Gurevych, TU Darmstadt, Germany

- **Multi-Source Object Identification With Constraints**

  Matteo Di Gioia, DIS-Università di Roma "Sapienza", Italy
  Domenico Beneventano, DII-Università di Modena e Reggio Emilia, Italy
  Monica Scannapieco, Istituto Nazionale di Statistica, Italy

- **EXPLAINIE - Explaining Information Extraction Systems**

  Wojciech Barczynski, SAP Research CEC Dresden, Germany
  Falk Brauer, SAP Research CEC Dresden, Germany
  Adrian Mocan, SAP Research CEC Dresden, Germany

Full papers

Master Data Management: Products and Research

Jochen Kokemüller, Anette Weisbecker

Abstract

Master Data Management is the discipline of creating and maintaining high value, high quality master data. In this contribution we give a definition of this data category underlining its importance to an overall high level of cooperate data quality. The current situation of commercial master data management solutions is presented. It is based on the results on six systems of two surveys we conducted. Here we discuss the systems capabilities for information integration, data modeling and information quality. After the current situation we provide an outlook on future developments in the area of master data management and discuss the relevance of Peer-To-Peer technologies. On this behalf we go into some detail discussing the specialized architecture VIANA.

Information Management Along the Lifecycle of Data and Application Systems: Challenges and Solution Approaches

Georg Fischer, Axel Herbst

Abstract

A tried and tested method for improving the quality of information in thebusiness system and control data volume growth, is moving selected, eligible datafrom the application system to less expensive long-term storage. To ensure that theeligibility of the data is actively taken into account, methods such as archivabilitychecks must be part of this process. Descriptive policies play into the strategy, toallow for the compliant retention of the moved data along its entire life cycle. Afurther challenge is controlled destruction to complete the life cycle of dataaccording to the law. At the same time, the fact that often the life span of applicationsystems is shorter than that of the data itself, needs to be considered.In this presentation we show how long-standing archiving techniques for businessdata are challenged anew by the growing complexity of today's legal and businessrequirements, causing an evolution to information lifecycle management.

Flexible and Generic Data Quality Metadata Exchange

David Becker, John Jaster, Jereme Kuperman

Abstract

Data quality metadata frequently needs to be exchanged between various parties and tools engaged in themanagement of data quality (DQ). In this paper we describe a data quality metadata exchange (DQME) ExtensibleMarkup Language (XML) approach that addresses a number of data quality management fundamentals in a flexibleand generic manner. The approach provides for the exchange of XML messages covering DQ definitions, businessrule evaluations and DQ measurements. Each of these types of DQ metadata exchange operates within the contextof a corresponding set of processes and tools. The approach has been developed to support an Enterprise DataQuality Management Service (EDQMS) project being implemented for the United States Air Force (USAF)Operations Support community.

Extending BPMN to Support the Modeling of Data Quality Issues

Noelia Sánchez-Serrano, Ismael Caballero, Félix García

Abstract

Many companies have just realized about the increasing importance of dealing with data which have asuitable level of quality for their business operations, thus avoiding the errors and major pitfalls that could make anegative impact their performance. Various works agree that the means to achieve this goal is through thereengineering of organizational business processes, focusing attention on the most critical points, and bringing thedata and information quality requirements to fruition in the most explicit manner possible. BPMN provides a suitablenotation with which to represent business process issues, although it lacks the means to cope with specific dataquality requirements. IP-MAP, on the other hand, does have these means, but is not as widely accepted and used torepresent business process as is BPMN; in addition, BPMN is supported by different tools, which makes it moreusable in practice. The main aim of this paper is to analyze BPMN's capability for represent required Data and Information Quality issues for business processes by providing certain proposals of extensions based on IP-MAP. We discuss how to extend BPMN to the support the data quality provided by IP-MAP in order to obtain asufficiently rich notation. An example of the application of this extension is also presented.

Data Quality Through Conceptual Model Quality - Reconciling Researchers and Practitioners Through a Customizable Quality Model

Kashif Mehmood, Samira Si-Said Cherfi, Isabelle Comyn-Wattiau

Abstract

Data quality has emerged as an important and challenging topic in recent years. In this article we areaddressing the conceptual model quality since it has been widely accepted that better conceptual models producebetter information systems and thus implicitly improve the data quality. Unfortunately there is neither standard noragreed framework for managing conceptual models quality. This article presents an overview of existing approacheswith their advantages and limitations. It then proposes a comprehensive model for evaluating the quality ofconceptual models. A survey involving practitioners has been used as an initial validation. This validation exerciseaims to collect the responders' views on the holistic quality of the conceptual models in addition to their feedbackover the newly proposed model. The received feedback has been evaluated and incorporated to the quality model.Furthermore, we propose a general approach for quality evaluation and improvement.

How to Screen a Data Stream - Quality-Driven Load Shedding in Sensor Data Streams

Anja Klein, Gregor Hackenbroich, Wolfgang Lehner

Abstract

As most data stream sources exhibit bursty data rates, data stream management systems mustrecurrently cope with load spikes that exceed the average workload to a considerable degree. To guaranteelow-latency processing results, load has to be shed from the stream, when data rates overstress systemresources. There exist numerous load shedding strategies to delete excess data. However, the consequentdata loss leads to incomplete and/or inaccurate results during the ongoing stream processing.

In this paper, we present a novel quality-driven load shedding approach that screens the data stream to find and discard data items of minor quality. The data quality of stream processing results is maximized underthe adverse condition of data overload. After an introduction to data quality management in data streams,we define three data quality-driven load shedding algorithms, which minimize the approximation error ofaggregations and maximize the completeness of join processing results, respectively. Finally, we demonstratetheir superiority over existing load shedding techniques at real-life weather data.

SOG: A Synthetic Occupancy Generator to Support Entity Resolution Instruction and Research

John R. Talburt, Yinle Zhou, Savitha Yalanadu Shivaiah

Abstract

This paper reports on a project to develop SOG (Synthetic Occupancy Generator), a system to create realistic, but synthetic residential occupancy (name andaddress) histories as input for Entity Resolution (ER) processes. ER processes areintended to link records referencing the same, or related, real-world entities. Mostorganizations use some type of ER process to recognize their customers or clients acrossdifferent channels of contact such as name and address, telephone number, or emailaddress. However, growing concerns over customer privacy and identity theft have madeorganizations reluctant to publicly release personally-identifiable customer information.The result is that it can difficult to obtain actual occupancy information to use for studentexercises or to experiment with entity resolution methods and techniques. SOG wascreated to address this problem by providing a tool capable of automatically generating alarge number of realistic, but synthetic occupancy histories. SOG control parametersallow the user to customize certain features of the simulated occupancy histories. Theproject reported here is the first phase of a larger project. The second phase is to developtools that will systematically disrupt the SOG output to create ER test files that havevarying degrees of data quality and file formats.

Computing Uncertain Key Indicators from Uncertain Data

Carlos Rodriguez, Florian Daniel, Fabio Casati, Cinzia Cappiello

Abstract

Key indicators, such as key performance indicators or key compliance indicators are atthe heart of modern business intelligence applications. Key indicators are metrics, i.e., numbers, thathelp an organization to measure and assess how successful it is in reaching predefined goals (e.g.,lowering process execution times or increasing compliance with regulations), and typically the peoplelooking at them simply trust the values they see when taking decisions. However, it is importantto recognize that in real business environments we cannot always rely on fully trusted or certaindata, yet indicators are to be computed.

In this paper, we tackle the problem of computing uncertain indicators from uncertain data, we characterizethe problem in a modern business scenario (combining techniques from uncertain and probabilisticdata management), and we describe how we addressed and implemented the problem in aEuropean research project

Identification of Business Oriented Data Quality Metrics

Boris Otto, Kai M. Hüner, Hubert Oesterle

Abstract

Corporate data of poor quality can have a negative impact on the performance of businessprocesses and thereby the success of companies. Similar to machine tools corporate data show signsof wear (imaging a moving customer with a new address, for example) and have to be monitoredcontinuously for quality defects. Effective quality control of corporate data requires metrics thatmonitor potential data defects with the most significant impact on the performance of a company'sprocesses. However, due to company specific success factors and IT landscapes, it is hardly possibleto provide generic metrics that can be implemented without any adjustment. This paper presents amethod for the identification of business oriented data quality metrics. The presented approach takesinto account company specific requirements from both a business and an IT perspective. The method'sdesign and evaluation process is discussed in the context of three real-world cases.

A Framework for Economic-driven Assessment of Data Quality Decisions

Adir Even, Marcus Kaiser

Abstract

Economic perspectives have raised growing attention in recent data quality (DQ) literature, as studieshave associated DQ decisions with major cost-benefit tradeoffs. Despite the growing interest, DQ research has notyet developed a robust, agreed-upon view for assessing and studying the link between DQ and economic outcome.As a contribution, this study proposes a framework, which links costs to the decisions made in managing theinformation process and improving the DQ, and benefits to the use of information-product outcomes by dataconsumers. Considering past research contributions, we develop this framework further into a high-leveloptimization model that permits quantitative assessment of cost-benefit tradeoffs, towards economically-optimal DQdecisions. We demonstrate a possible use of the proposed framework and the derived model, and highlight theirpotential contribution to an economics-driven view of DQ issues in both research and practice.

A SOA-based Data Quality Assessment Framework in a Medical Science Center

Yao Zhou, Sabine Hanß, Malte Cornils, Claudia Hahn, Sonja Niepage,Thomas Schrader

Abstract

The Open European Nephrology Science Center is a metadata repository of clinical and research relateddata on nephrology diseases. Clinical data is source data for medical research. Medical research has specific andchanging quality requirements for source data, so the quality of clinical data has to be assessed before it can be usedfor medical research. Quality assessment requirements change from one medical research to another, hence, aflexible data quality assessment solution is needed. In this paper, we present a data quality assessment frameworkwhich supports flexible data quality assessment by allowing users to define their own data quality assessmentrequirements, and automatically searching for proper data quality assessment tools which fulfill the requirements.Development of the framework is based on an ontology and SOA. Components of this framework are the FormalQuality Requirement, the Service Repository and the Service Selection Process. Those components are discussed indetail and an example is given to show how the framework works.

Information Quality Issues in the Mortgage Banking Industry

Rolf T. Wigand, Jerry Wood, Yusuf Yiliyasi

Abstract

The current mortgage crisis and credit crunch presents the U.S. mortgage industry notonly with unprecedented challenges, but also with great opportunities for improvement.In this research, information quality (IQ) issues within the mortgage banking industry arereviewed and discussed. Through extensive literature reviews from a multitude ofmortgage industry resources, IQ issues within the industry are analyzed utilizing six ofthe most applicable Wang and Strong dimensions. These widely regarded IQ dimensionsprovide a framework by which IQ issues in the mortgage banking industry can beassessed. Specific IQ problems are indentified and possible means of improvement arediscussed. We focus especially on how the Mortgage Industry Standards MaintenanceOrganization (MISMO) standards and data specifications help improve IQ in thisindustry. This research concludes that industry wide standards such as MISMO'seMortgage standards can help provide extensive benefits to the mortgage bankingindustry providing firms are willing to accept and adopt these standards. And, moreimportantly, we hope that these conclusions and recommendations may make a small, butimportant contribution to the current malaise in the mortgage industry.

The Quality of Monitoring Data in Civil Engineering Works

Ana Lucas, António Palma-dos-Reis, Mário Caldeira

Abstract

This paper presents the research and discusses the findings concerning the identification and definition of quality dimensions related to the monitoring data obtained in the safety control of major civil engineering works, namely dams and bridges. The analysis of the behavior and the safety evaluation of these works essentially depend on the quality of the data collected by the sensors, named the *observation data*, besides the quality of the models and, last but not least, the knowledge and experience of the experts involved. In addition, LNEC 1 stores and preserves the observation data concerning major civil engineering works, which represents an important research source for LNEC and Universities and, thus, shapes itself as an e-Science scenario. Differently from other e-Science environment authors, we have not restricted our work to objective dimensions, whose measures can be calculated automatically. It appears, from the experts' opinions, that some subjective dimensions can largely enrich the quality information about archived data at the expense of some additional work in data curation.

In the development of this research we used an online version of the Delphi method with the Q-Sort technique, complemented with upstream interviews and a downstream meeting with the experts. We found out, and sorted by the importance level assigned by the experts, ten dimensions for the quality of observation data, which are *error-of-observation, coherence, relevancy, interpretability, timeliness, completeness, accessibility, appropriate amount of data, access security and preservation.*

Data Quality Evaluation in an E-Business Environment: A Survey

Soumaya Ben Hassine-Guetari

Abstract

Industrial databases and information systems are plagued by a plethora of quality issues especially in cooperative information systems under the use of external data files because of the lack of collaboration in internal integration processes of marketing and sales data which is still a biggest challenge facing businesses, and, above all, the lack of benchmarks and tools standardizing for the external files exchanges and handling. In fact, the quality of exchanged data is essential for developing service-based applications and correctly performing cooperative activities such in Business-to-Business (B-to-B) marketing operations where issues are no longer limited to individual erroneous records but also to the insufficiency of customer knowledge and files features when integrating external data files.

In this paper, we describe data assessment methods needed to assess and maintain the quality of an e-business activity. The aim is to define the basis of a data quality assessment framework that is suitable for evaluating and enhancing an e-business information system.

Multidimensional Management and Analysis of Quality Measures for CRM Applications in an Electricity Company

Verónika Peralta, Virginie Thion-Goasdoué, Zoubida Kedad, Laure Berti-Équille, Isabelle Comyn-Wattiau, Sylvaine Nugier, Samira Si-said-Cherfi

Abstract

This paper presents an approach integrating data quality into the business intelligence chain in the context of customer-relationship management (CRM) applications at EDF (Electricité de France), the major electricity company in France. The main contribution of this paper is the definition and instantiation of a generic multi-dimensional star-like model for storing, analyzing and capitalizing data quality indicators, measurements and metadata. This approach is illustrated through one of EDF's CRM applications, implementing domain-specific quality indicators and providing quality-driven information management as a business intelligence chain. The role of the data quality expert is highly emphasized.

An Empirical Study on Criteria for Assessing Information Quality in Corporate Wikis

Therese Friberg, Wolfgang Reinhardt

Abstract

Wikis gain more and more attention as tool for corporate knowledge management. The usage of corporate wikis differs from public wikis like the Wikipedia as there are hardly any wiki wars or copyright issues. Nevertheless the quality of the available articles is of high importance in corporate wikis as well as in public ones. This paper presents the results from an empirical study on criteria for assessing information quality of articles in corporate wikis. Therefore existing approaches for assessing information quality are evaluated and a specific wikiset of criteria is defined. This wiki-set was examined in a study with participants from 21 different German companies using wikis as essential part of their knowledge management toolbox. Furthermore this paper discusses various ways for the automatic and manual rating of information quality and the technical implementation of such an IQ-profile for wikis.

Towards Assessing Information Quality in Knowledge Management in The Enterprise 2.0

Sven Ahlheid, Therese Friberg, Gernot Gräfe, Alexander Krebs, Jan-Philipp Müller, Dirk Schuster

Abstract

With regard to the success stories of Web 2.0 based knowledge centers such as the online encyclopedia Wikipedia [54] companies have begun to enrich their corporate knowledge management with Web 2.0 technologies, hoping to benefit from increasing flows of information. Besides information quantity, the quality of information is a key factor determining the return on investment of such Enterprise 2.0 platforms. In this context we will discuss requirements for the concept of information quality, identify important differences to the Web 2.0 environment and also elaborate on the basic design of a system assessing information quality in an Enterprise 2.0 context. We will thereby integrate implicit user feedback and explain the key benefits of this novel approach.

Web-based Affiliation Matching

David Aumüller, Erhard Rahm

Abstract

Authors of scholarly publications state their affiliation in various forms. This kind of heterogeneity makes bibliographic analysis tasks on institutions impossible unless a comprehensive cleaning and consolidation of affiliation data is performed. We investigate automatic approaches to consolidate affiliation data to reduce manual work and support scalability of affiliation analysis. In particular, we propose to set up a reference database of affiliation strings found in publications. A key step in this task is the matching of different affiliation strings to determine whether or not they match. For affiliation matching we investigate web based similarity measures utilizing the cognitive power of current search engines. They determine the similarity of affiliations based on how the URLs in the result sets of affiliation web searches overlap. We evaluate the effectiveness of affiliation matching based on URL overlap as well as for the combined use with the Soft TF-IDF similarity measure.

Poster papers

Is DQ/IQ the Quality of Information? Two Views

Zbigniew J Gackowski

Abstract

This is an inquiry into quality dimensions commonly attributed to information based on threecases. Operations management and survey-based consumer views are contrasted. The majority of IQ dimensions address use-related aspects of operation factors that are determined by situation-specific users' needs and expectations. Thus, IQ is a short name for a broader meaning than the term implies. It misguides researchers in articulating research questions or hypotheses about IQ, tempting them into using aggregates of no-covariant dimensions as substitutes for elementary dimensions. The question is, which quality aspects are actual attributes of the value and format of information, and which pertain to the circumstances of its use? Then different questions need to be asked and different relationships examined.

A Meta-model for Data Quality Management Simulation

Boris Otto, Kai M. Hüner

Abstract

Data quality management initiatives could both help to prevent the occurrence of data defects and repair their effect. While such initiatives can reduce overall costs, they also cause costs for their development and implementation. Therefore, the overall aim is not to improve data quality by any means, but to ensure cost-efficiency. The paper proposes a meta-model for simulating data quality management, which can be used for planning of cost-efficient initiatives.

Multi-Source Object Identification With Constraints

Matteo Di Gioia, Domenico Beneventano, Monica Scannapieco

Abstract

The problem of identifying the manifold generated copies of an object is known as Object Identification (OI). Numerous solutions have been proposed to solve this task, based on the similarity between two objects. Most of these solutions are oriented to discover pairs of duplicates (pairs-oriented OI) rather than sets of similar objects (group-oriented OI), for which some clustering techniques are used. In this paper, we proposed a new technique, based on the concept of constraints, to resolve the group-oriented OI problem. It is composed of two phases: extraction phase and grouping. During the extraction phase constraints are extracted by analyzing data at hand. After that we have collected the constraints, we reason about those to find the groups of similar objects. The group-based OI technique we propose allows us to deal with multiple sources.

A Multi-Dimensional Model for Assessing the Quality of Answers in Social Q&A Sites

Zhemin Zhu, Delphine Bernhard, Iryna Gurevych

Abstract

The quality of user-generated content in Web 2.0 dramatically varies from professional to abusive. Quality assessment is therefore a critical problem in producing, managing and retrieving information in Web 2.0. In this paper, we develop a multi-dimensional model for assessing the quality of answers in social Q&A (Question & Answer) sites.

Schema Based Deduplication

Pei Li, Andrea Maurino

Abstract

In this paper we present a preliminary report on a domain independent strategy to reduce duplicated records by means of the knowledge stored in the schema. According to different kinds of relationships, we propose specific techniques to build and compare the knowledge networks by means of graph-based similarity techniques.

EXPLAINIE - Explaining Information Extraction Systems

Wojciech Barczynski, Falk Brauer, Adrian Mocan

Abstract

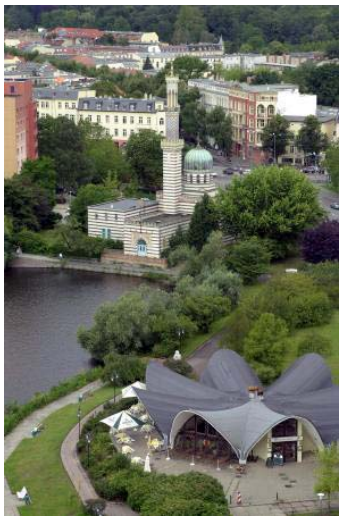
Business Intelligence (BI) over unstructured text is under intense scrutiny both in the industry and research. Recent work in this field includes automatic integrating of unstructured text into business analytics, model recognition, and probabilistic databases to handle uncertainty of Information Extraction (IE). However, still an open issue is how to handle IE quality, which is a part of ETL like process for the BI. Precision of IE is still too low for BI and, according to Sunita Sarawagi in recent survey on IE, we are still far from a comprehensive quality model for IE. Currently the BI user has neither methodology nor tools, which would help him to discover if the result is an unexpected fact or an error in IE. In this work we present preliminary results on developing methodology and tool (ExplainIE), which helps users to debug unexpected results. ExplainIE presents results within BI tool and auxiliary view on low level detail (e.g., entity graph). We consider two kinds of users: BI and IE expert.

Location

Since 1996, the ICIQ conferences were held annualy at MIT in Cambrige. This year, ICIQ goes abroad: The Hasso Plattner Institute (HPI) in Potsdam, on the border of Berlin hosts ICIQ 2009.

Potsdam

Without a doubt, Potsdam is one of the most beautiful cities in Germany. Adding to the cultural ambience, the capital city of the State of Brandenburg is a natural destination to satisfy varied interests and demands.



Potsdam - a city surrounded with water

Potsdam's most popular site is the *Schloß Sanssouci* palace, located in the park to which it gives its name. However, there is far more than the former summer residency of Frederick the Great and the Sanssouci park to make your visit worthwhile. The *Alexandrowka*, the *Holländisches Viertel* (the Dutch Quarter) and the Weavers' Quarter - the historic sections of the city - provide the flair of a city steeped in European tradition.

Sanssouci palace with the great fountain

The Babelsberg Filmpark, the Biosphäre nature experience exhibit, the National Horticulture Show park, known as the BUGA park and built for the National Horticulture Show in 2001, the tourist cruise ships, and the Krongut Bornstedt (the crown estate) are all special attractions which will make your visit to Potsdam a very pleasurable experience.

In recent years, a modern tourism framework has been developed to fulfill every wish and requirement. Contemporary hotels and guest houses, as well as private rooms, allow our guests to feel easily at home.

For further information, you may refer to these websites:

- Potsdam official website: http://www.potsdam.de
- http://www.potsdam-tourism.com

Potsdam: Attractions and Sightseeing

Sanssouci Park

Sanssouci Park is an ensemble of the garden and palaces that began to take shape under Frederick the Great in the 18th century and expanded in the 19th century under Frederick William IV. Sanssouci Palace, the summer residence of Frederick the Great, forms the central focus of the ensemble.

Dutch Quarter (Das Holländische Viertel)

These 134 Dutch style red brick houses were built by order of the "Soldier King", in an attempt to attract Dutch immigrants to Potsdam. Today, the quarter is home to many shops and restaurants with a unique flair.

Church of Peace (Friedenskirche)

Based on an Italian model, this church was built from 1845 to 1854. King Frederick William IV and Queen Elisabeth were laid to rest here. The adjacent mausoleum holds the tomb of Emperor Frederick III, Empress Victoria and King Frederick William I.

Cecilienhof Palace (Schloss Cecilienhof)

Cecilienhof Palace was built as residence of Crown Prince William and his wife Cecilie from 1913 to 1917 in the style of an English country manor. It became famous as the site of the Potsdam Conference in 1945.

Babelsberg Palace (Schloss Babelsberg)

Built from 1834 to 1835 in the neo-Gothic style according to plans by K. F. Schinkel. Due to on-going restoration work, the palace is partly open to visitors.

Steam Engine (Dampfmaschinenhaus)

From 1841 to 1843, L. Persius built a steam engine building designed to look like a mosque which served as the pumping station for the fountains in Sanssouci Park.

Babelsberg Filmpark

Visitors can immerse in the world of film and television. Experience breathtaking stunt shows and amazing film technology, stroll through film scenery, admire elaborate props and discover the ingenious tricks of Babelsberg's filmmakers.

Glienicke Palace (Schloss Glienicke)

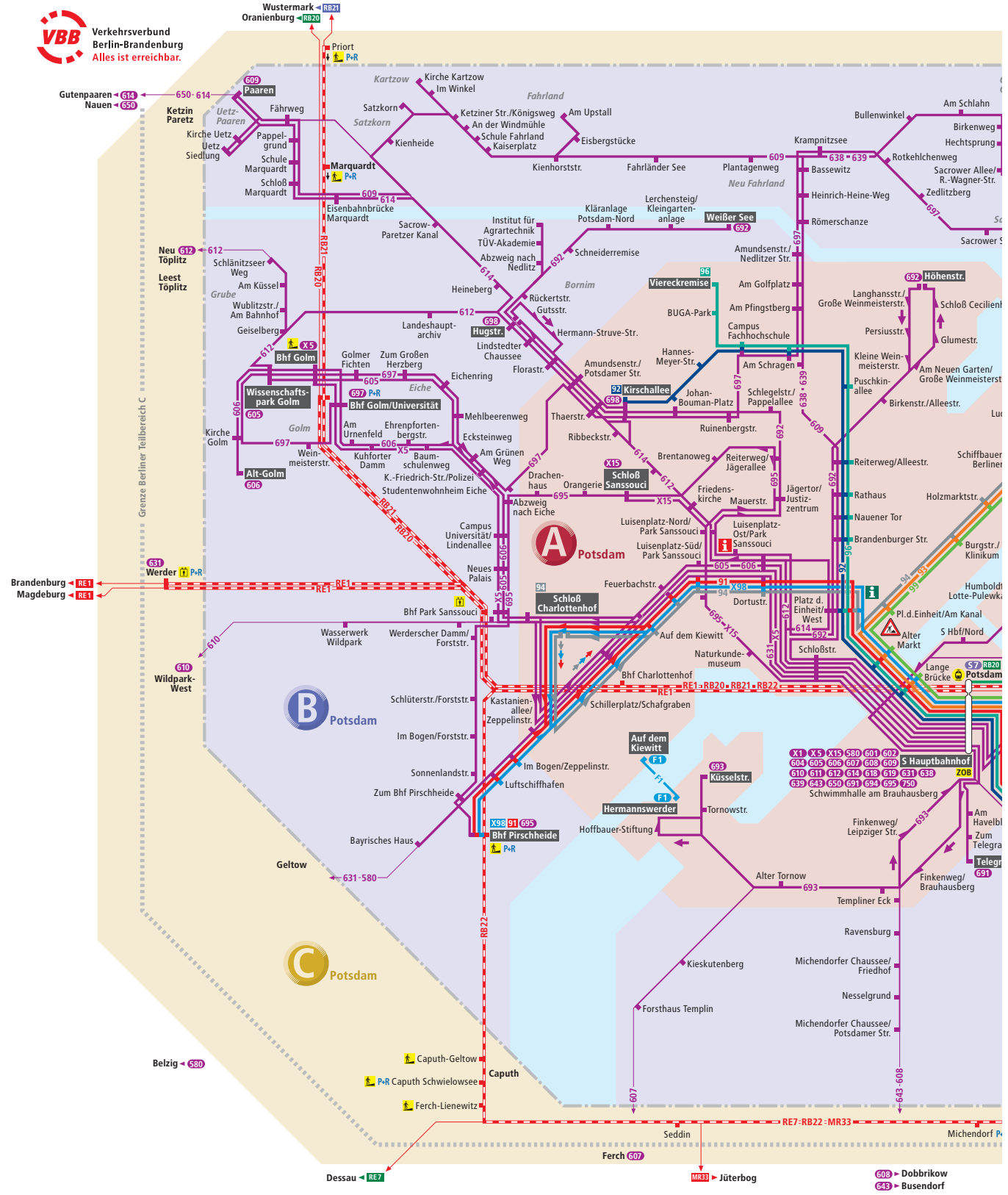
This neo-Classical palace was built from 1825 to 1828 by K. F. Schinkel. The Court Gardener's Museum is also located here.

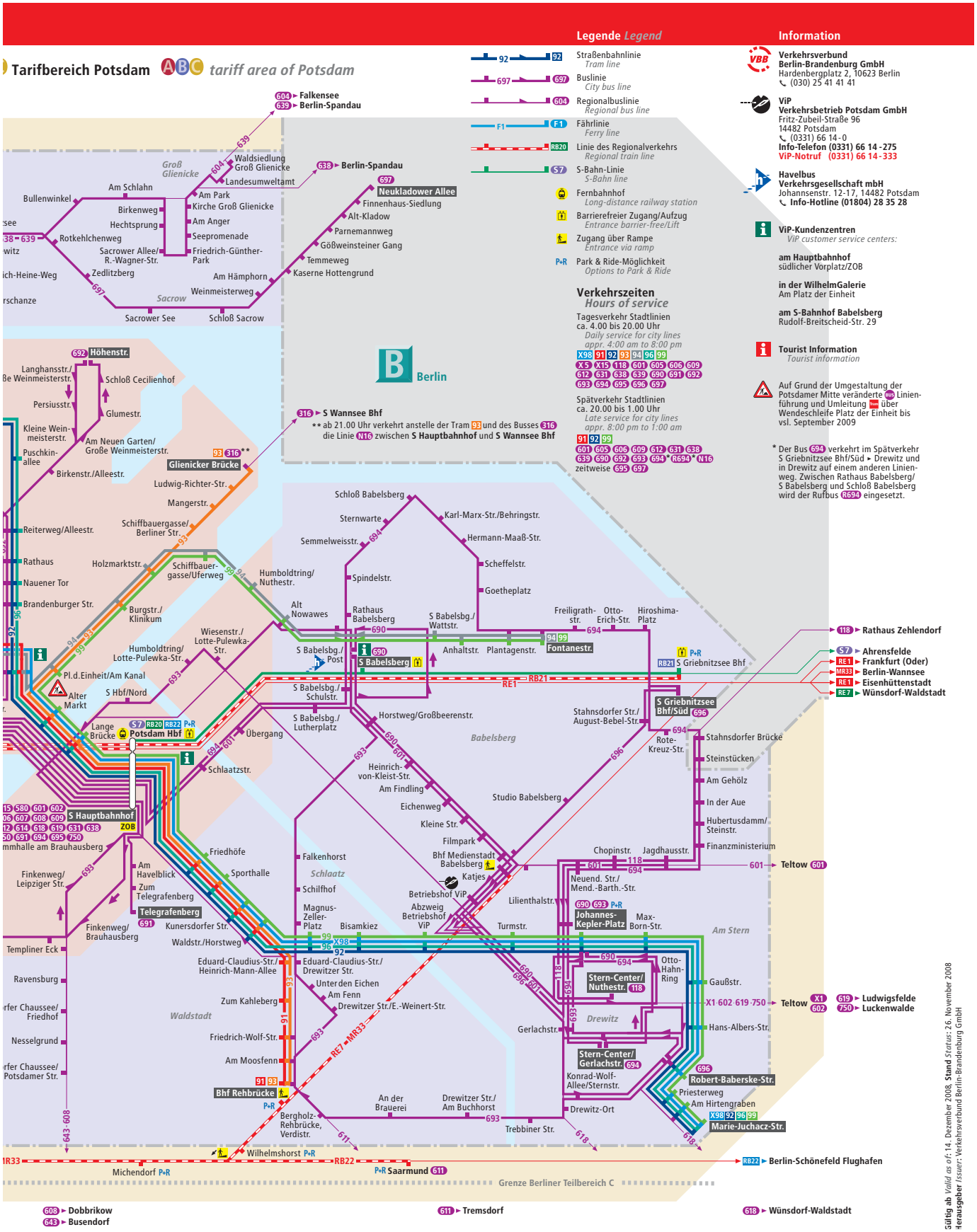
Potsdam

Tageliniennetz

Daytime Network

ABC Tarifbereich Potsdam





Gültig ab Valid as of: 14. Dezember 2008, Stand Status: 26. November 2008
Herausgeber Issuer: Verkehrsverbund Berlin-Brandenburg GmbH

Berlin

Berlin is the capital city and one of sixteen states of Germany. With a population of 3.4 million within its city borders, Berlin is Germany's largest city. It is the second most populous city and the eighth most populous urban area in the European Union. Located in northeastern Germany. It is the center of the Berlin-Brandenburg metropolitan area, comprising 5 million people from over 190 nations.

Berlin is a major center of culture, politics, media, and science in Europe. Its economy is primarily based on the service sector, encompassing a diverse range of creative industries, media corporations, environmental services, congress and convention venues. The city serves as a continental hub for air and rail transport, and is one of the most visited tourist destinations in the EU. Other industries include traffic engineering, optoelectronics, IT, pharmaceuticals, biomedical engineering, and biotechnology.

Located in eastern Germany, almost halfway between Paris and Moscow, Berlin is laid out on an epic scale. When the city-state of Berlin was incorporated in 1920, it swallowed towns and villages far beyond the downtown area around the two main rivers, the Spree and the Havel. Each of its boroughs has distinctive characteristics. Charlottenburg, Schöneberg, and Kreuzberg are popular areas in the west, and to the east, Prenlzauer Berg and Friedrichshain are favored residential and nightlife neighborhoods. Modern urban commercial centers such as Potsdamer Platz and Leipziger Platz still feel like odd insertions between the historically developed quarters surrounding them.

Berlin: Attractions and Sightseeing

Berlin has many sights worth seeing. Here is only a brief list. For further information in several languages, please refer to http://www.visitberlin.de

Experience Berlin by public transportation

Buses 100 and 200

Berlin's city centre is very extensive and cannot be fully explored on foot. It is worth experiencing certain stretches by public transport. The number 100 and 200 buses drive past many sights from Zoologischer Garten to Alexanderplatz, so that a journey on them is already like a sightseeing tour.

S-Bahn "Stadtbahn"

You can also catch a first glimpse of many sights from the S-Bahn on its raised tracks. On the stretch between Zoologischer Garten and Alexanderplatz you will pass by the Tiergarten with the Siegessäule: (Victory Column), Reichstag: (Parliament building) and government buildings and glide between the buildings on Museum Island.

Taxis in Berlin

Alongside the well developed network of buses and trains in Berlin, there are almost 7000 taxis ready to take residents and their visitors all over the city. Finding a taxi is easy: whether just hailing one down, at a taxi rank or by calling in advance.

Berliner Dom

The Berliner Dom (Berlin Cathedral), completed in 1905, is Berlin's largest and most important Protestant church as well as the sepulchre of the Prussian Hohenzollern dynasty.

Brandenburger Tor

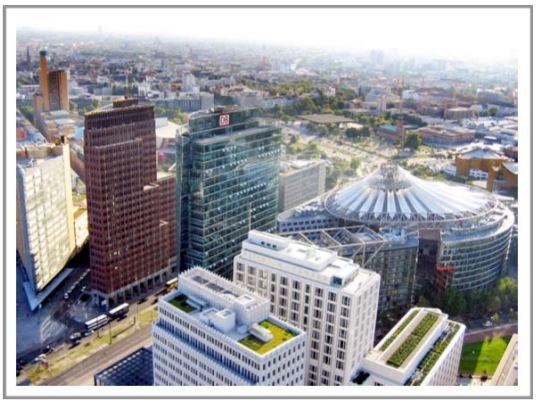
The Brandenburg Gate is one of Berlin's most important monuments – a landmark and symbol all in one with over two hundred years of history

Checkpoint Charlie

Checkpoint Charlie, along with Glienicker Brücke (Glienicker Bridge) was the best known border-crossing of Cold War days.

Jüdisches Museum (Jewish Museum)

A timeless monument to Jewish history and life in Germany, Daniel Libeskind's Berlin Juedisches Museum is one of the world's undisputed museums and architectural gems



Potsdamer Platz and Sony Center

Berlin's Potsdamer Platz is the most striking example of the urban renewal that turned Berlin into the 'New Berlin' in the 1990s. Sony Centeris a Sony-sponsored building complex opened in 2000.

Museumsinsel

Berlin's Museumsinsel (Museum Island) is a unique ensemble of five museums, including the Pergamon Museum - built a the small island in Berlin's Spree River between 1824 and 1930

Fernsehturm and Alexanderplatz

Everyone is supposed to remember that Berlin's Fernsehturm (TV Tower) is 365m high and is the tallest building in Berlin. Fernsehturm is located at Alexanderplatz.

Siegessäule

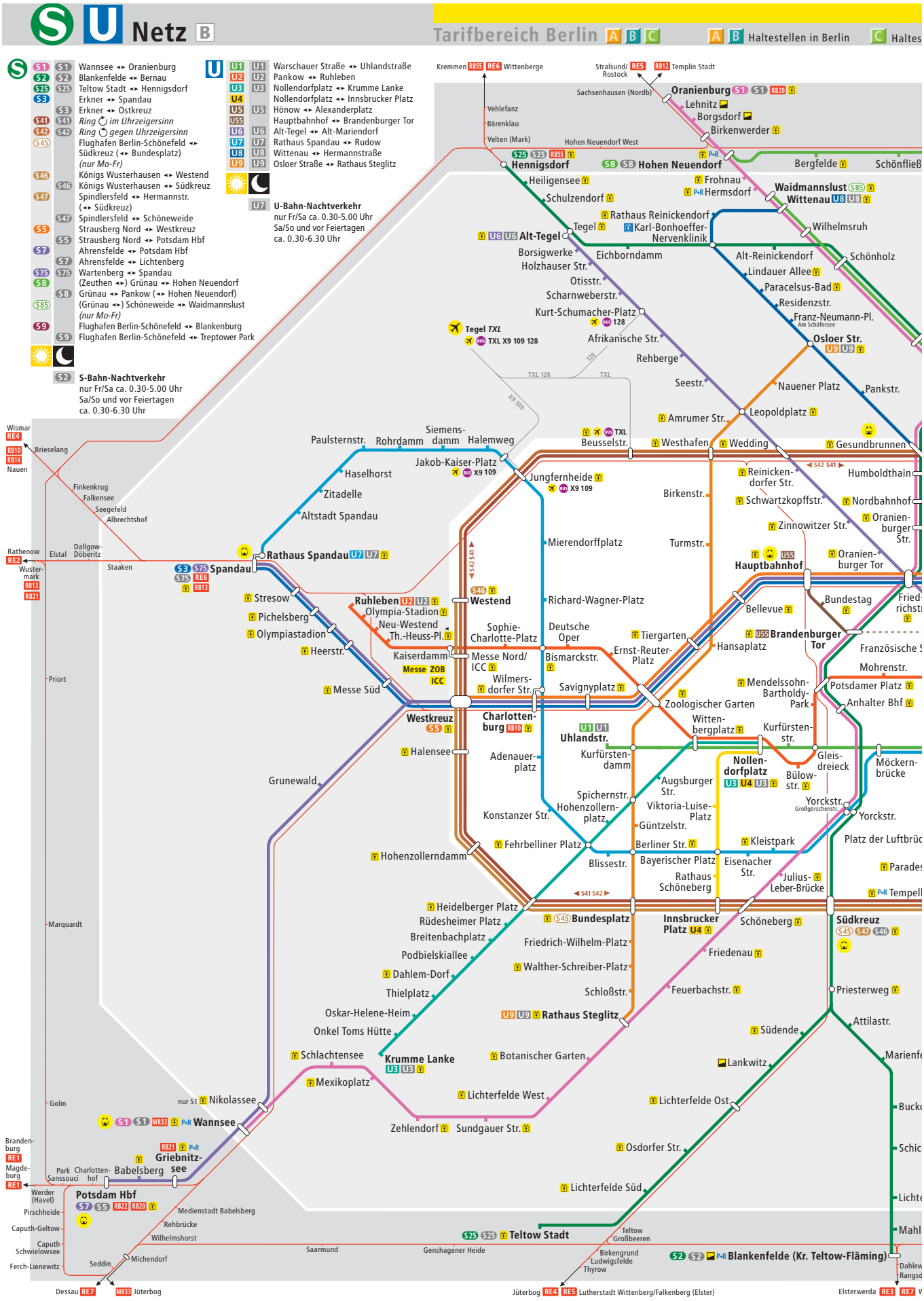
Berlin's Siegessäule - Victory Column - is another of Berlin's monuments that has reinvented itself through the ages - from symbol of Prussian military victory in the 19th century to a favourite tourist spot today.

Kaiser Wilhelm Gedächtniskirche

The Gedächtniskirche or Kaiser Wilhelm Memorial Church is the symbolic centre of West Berlin, an anti-war memorial to peace and reconciliation. Following Allied bombing during WWII, the original, west tower has remained standing as a ruin and is popularly named the "hollow tooth".

Reichstag

Following German reunification on October 3, 1990 the Bundestag (German Federal Parliament) decided, one year later, to reinstate the Reichstag as the seat of Parliament in Berlin, the restored capital of reunited Germany.


S Netz B
Tarifbereich Berlin A B C
A B Haltestellen in Berlin
C Haltes
S1 Wannsee ◂▸ Oranienburg
S2 Blankenfelde ◂▸ Bernau
S25 Teltow Stadt ◂▸ Hennigsdorf
S3 Erkner ◂▸ Spandau
S3 Erkner ◂▸ Ostkreuz
S41 Ring im Uhrzeigersinn
S42 Ring gegen Uhrzeigersinn
S45 Flughafen Berlin-Schönefeld ◂▸ Südkreuz (◂▸ Bundesplatz) (nur Mo-Fr)
S46 Königs Wusterhausen ◂▸ Westend
S46 Königs Wusterhausen ◂▸ Südkreuz
S47 Spindlersfeld ◂▸ Hermannstr. (◂▸ Südkreuz)
S47 Spindlersfeld ◂▸ Schöneweide
S5 Strausberg Nord ◂▸ Westkreuz
S5 Strausberg Nord ◂▸ Potsdam Hbf
S7 Ahrensfelde ◂▸ Potsdam Hbf
S7 Ahrensfelde ◂▸ Lichtenberg
S75 Wartenberg ◂▸ Spandau
S8 (Zeuthen ◂▸) Grünau ◂▸ Hohen Neuendorf
S8 Grünau ◂▸ Pankow (◂▸ Hohen Neuendorf)
S85 (Grünau ◂▸) Schöneweide ◂▸ Waidmannslust (nur Mo-Fr)
S9 Flughafen Berlin-Schönefeld ◂▸ Blankenburg
S9 Flughafen Berlin-Schönefeld ◂▸ Treptower Park
S2 S-Bahn-Nachtverkehr nur Fr/Sa ca. 0.30-5.00 Uhr Sa/So und vor Feiertagen ca. 0.30-6.30 Uhr
U1 U1 Warschauer Straße ◂▸ Uhlandstraße
U2 U2 Pankow ◂▸ Ruhleben
U3 U3 Nollendorfplatz ◂▸ Krumme Lanke
U4 Nollendorfplatz ◂▸ Innsbrucker Platz
U5 U5 Hönow ◂▸ Alexanderplatz
U55 Hauptbahnhof ◂▸ Brandenburger Tor
U6 U6 Alt-Tegel ◂▸ Alt-Mariendorf
U7 U7 Rathaus Spandau ◂▸ Rudow
U8 U8 Wittenau ◂▸ Hermannstraße
U9 U9 Osloer Straße ◂▸ Rathaus Steglitz
U7 U-Bahn-Nachtverkehr nur Fr/Sa ca. 0.30-5.00 Uhr Sa/So und vor Feiertagen ca. 0.30-6.30 Uhr
Kremmen RB55 RE6 Wittenberge
Stralsund/ Rostock RE5 RB12 Templin Stadt
Sachsenhausen (Nordb)
Oranienburg S1 S1 RB20
Lehnitz
Borgsdorf
Birkenwerder
Vehlefanz
Bärenklau
Velten (Mark)
Hohen Neuendorf West
S25 S25 RB55 Hennigsdorf
S8 S8 Hohen Neuendorf
Bergfelde
Schönfließ
Heiligensee
Frohnau
Hermsdorf
Waidmannslust S85
Wittenau U8 U8
Schulzendorf
Rathaus Reinickendorf
Tegel
Karl-Bonhoeffer-Nervenklinik
Wilhelmsruh
U6 U6 Alt-Tegel
Borsigwerke
Holzhauser Str.
Eichborndamm
Alt-Reinickendorf
Schönholz
Otisstr.
Lindauer Allee
Scharnweberstr.
Paracelsus-Bad
Residenzstr.
Kurt-Schumacher-Platz
Franz-Neumann-Pl.
Am Schäfersee
Tegel TXL
TXL X9 109 128
Afrikanische Str.
Osloer Str. U9 U9
Rehberge
Seestr.
Nauener Platz
Pankstr.
Leopoldplatz
Amrumer Str.
Wismar RE4
Brieselang
RB10 RB14 Nauen
Paulsternstr.
Rohrdamm
Siemens-damm
Halemweg
Beusselstr. TXL
Westhafen
Wedding
Gesundbrunnen
Jakob-Kaiser-Platz X9 109
Haselhorst
Zitadelle
Altstadt Spandau
Jungfernheide X9 109
Reinicken-dorfer Str.
Humboldthain
Schwartzkopffstr.
Nordbahnhof
Birkenstr.
Zinnowitzer Str.
Oranien-burger Str.
Finkenkrug
Falkensee
Seegefeld
Albrechtshof
Mierendorffplatz
Turmstr.
Hauptbahnhof U55
Oranien-burger Tor
Rathenow RE2
Elstal
Dallgow-Döberitz
Wuster-mark RB13 RB21
Staaken
S3 S75 Spandau
S75 RE6 RB13
Rathaus Spandau U7 U7
S42 S41
Stresow
Pichelsberg
Olympiastadion
Heerstr.
Ruhleben U2 U2
Olympia-Stadion
Neu-Westend
Th.-Heuss-Pl.
S46 Westend
Richard-Wagner-Platz
Bundestag
Bellevue
Fried-richstr.
Sophie-Charlotte-Platz
Deutsche Oper
Tiergarten
Ernst-Reuter-Platz
Hansaplatz
U55 Brandenburger Tor
Französische S
Kaiserdamm
Messe ZOB ICC
Messe Nord/ ICC
Bismarckstr.
Wilmers-dorfer Str.
Savignyplatz
Mohrenstr.
Mendelssohn-Bartholdy-Park
Potsdamer Platz
Anhalter Bhf
Priort
Messe Süd
Westkreuz S5
Charlotten-burg RB10
Zoologischer Garten
Witten-bergplatz
U1 U1 Uhlandstr.
Kurfürsten-str.
Halensee
Adenauer-platz
Kurfürsten-damm
Nollen-dorfplatz U3 U4 U2
Gleis-dreieck
Bülow-str.
Möckern-brücke
Grunewald
Augsburger Str.
Spichernstr.
Hohenzollern-platz
Viktoria-Luise-Platz
Yorckstr. Großgörschenstr.
Yorckstr.
Konstanzer Str.
Güntzelstr.
Fehrbelliner Platz
Berliner Str.
Kleistpark
Platz der Luftbrücke
Hohenzollerndamm
Blissestr.
Bayerischer Platz
Eisenacher Str.
Rathaus Schöneberg
Julius-Leber-Brücke
Parades
P+R Tempelh
S41 S42
Heidelberger Platz
Rüdesheimer Platz
S45 Bundesplatz
Innsbrucker Platz U4
Schöneberg
Südkreuz S45 S47 S46
Breitenbachplatz
Friedrich-Wilhelm-Platz
Friedenau
Podbielskiallee
Walther-Schreiber-Platz
Dahlem-Dorf
Thielplatz
Schloßstr.
Feuerbachstr.
Priesterweg
Oskar-Helene-Heim
U9 U9 Rathaus Steglitz
Attilastr.
Onkel Toms Hütte
Südende
Schlachtensee
Krumme Lanke U3 U3
Botanischer Garten
Lankwitz
Marienfe
Mexikoplatz
Lichterfelde West
Lichterfelde Ost
Golm
nur S1 Nikolassee
S1 S1 MR33 P+R Wannsee
Zehlendorf
Sundgauer Str.
Osdorfer Str.
Brandenburg RE1
Magdeburg RE1
RB21 P+R Griebnitzsee
Park Sanssouci
Charlotten-hof
Babelsberg
Lichterfelde Süd
Werder (Havel)
Potsdam Hbf
S7 S5 RB22 RB20
Pirschheide
Medienstadt Babelsberg
Caputh-Geltow
Rehbrücke
Wilhelmshorst
Caputh Schwielowsee
Saarmund
Genshagener Heide
S25 S25 Teltow Stadt
Teltow Großbeeren
Ferch-Lienewitz
Seddin
Michendorf
Birkengrund Ludwigsfelde Thyrow
S2 S2 P+R Blankenfelde (Kr. Teltow-Fläming)
Dessau RE7
MR33 Jüterbog
Jüterbog RE4 RE5 Lutherstadt Wittenberg/Falkenberg (Elster)
Elsterwerda RE3 RE7


estellen in Brandenburg
www.BVG.de www.s-bahn-berlin.de
Legende
S+U-Bahn-Linie
Umsteigemöglichkeit
Bus-Anbindung zum Flughafen
Fernbahnhof
ZOB Zentraler Omnibusbahnhof
P+R Parkplatz
Barrierefrei durch Berlin
Barrierefreier Zugang, S+U Aufzüge
Barrierefreier Zugang/ Umsteigebahnhof, nur S Aufzug
Barrierefreier Zugang/ Umsteigebahnhof, nur U Aufzug
S+U Zugang über Rampe
Bezeichnung der Bahnhöfe unter Fortlassung der Tarifbezeichnung Berlin bzw. Potsdam
Service
BVG
Berliner Verkehrsbetriebe (BVG)
© Bereich
Vertrieb und Marketing (FVM)
Kundenbindung/Fahrgastinformation
10096 Berlin
BVG Call Center: (030) 19 44 9
DB BAHN Bahn Berlin
S-Bahn Kundentelefon
(030) 29 74 33 33
Stand: 31. August 2009
Groß Schönebeck (Schorfheide)
Stralsund/Schwedt (Oder)
Frankfurt (Oder)
Schmachtenhagen
Wensickendorf
Zühlsdorf
Wandlitzsee
Wandlitz
Basdorf
Schönwalde (Barnim)
Schönerlinde
Rüdnitz
Bernau
Bernau-Friedenstal
Zepernick (bei Bernau)
Röntgental
Buch
Karow
Mühlenbeck-Mönchmühle
Blankenburg
Pankow-Heinersdorf
Pankow
Wollankstr.
Vinetastr.
Bornholmer Str.
Schönhauser Allee
Prenzlauer Allee
Greifswalder Str.
Voltastr.
Bernauer Str.
Eberswalder Str.
Senefelderplatz
Rosa-Luxemburg-Pl.
Rosenthaler Platz
Weinmeisterstr.
Alexanderplatz
Hackescher Markt
Schillingstr.
Strausberger Platz
Weberwiese
Jannowitzbrücke
Frankfurter Tor
Samariterstr.
Klosterstr.
Märkisches Museum
Stadtmitte
Ostbahnhof
Warschauer Str.
Ostkreuz
Hausvogteiplatz
Spittelmarkt
Heinrich-Heine-Str.
Kochstr.
Moritzplatz
Hallesches Tor
Kottbusser Tor
Prinzenstr.
Schlesisches Tor
Görlitzer Bhf
Schönleinstr.
Mehringdamm
Hermannplatz
Gneisenaustr.
Südstern
Rathaus Neukölln
Boddinstr.
Karl-Marx-Str.
Leinestr.
Hermannstr.
Neukölln
Alt-Tempelhof
Kaiserin-Augusta-Str.
Ullsteinstr.
Westphalweg
Alt-Mariendorf
Grenzallee
Blaschkoallee
Parchimer Allee
Britz-Süd
Johannisthaler Chaussee
Lipschitzallee
Wutzkyallee
Zwickauer Damm
Rudow
X7 171 N7
Landsberger Allee
Storkower Str.
Frankfurter Allee
Magdalenenstr.
Lichtenberg
Nöldnerplatz
Tierpark
Friedrichsfelde Ost
Friedrichsfelde
Rummelsburg
Betriebsbahnhof Rummelsburg
Karlshorst
Treptower Park
Plänterwald
Sonnenallee
Köllnische Heide
Baumschulenweg
Schöneweide
Betriebsbahnhof Schöneweide
Adlershof
Grünau
Altglienicke
Grünbergallee
Eichwalde
Zeuthen
Wildau
Niederlehme
Zernsdorf
Kablow
Königs Wusterhausen
Frankfurt (Oder)
Cottbus
Senftenberg
Bitte beachten Sie, dass der Ort Schönefeld einschließlich Flughafen und Bahnhof dem Teilbereich C des VBB Tarifbereiches Berlin zugeordnet ist.
Flughafen Berlin-Schönefeld
Schönefeld SXF
X7 171 N7
Hohenschönhausen
Wartenberg
Gehrenseestr.
Ahrensfelde
Werneuchen
Seefeld (Mark)
Blumberg
Ahrensfelde Nord
Ahrensfelde Friedhof
Springpfuhl
Poelchaustr.
Marzahn
Raoul-Wallenberg-Str.
Mehrower Allee
Strausberg Nord
Strausberg Stadt
Hegermühle
Hönow
Louis-Lewin-Str.
Hellersdorf
Cottbusser Platz
Petershagen Nord
Neue Grottkauer Str.
Fredersdorf
Kaulsdorf-Nord
Neuenhagen
Strausberg
Kostrzyn
Biesdorf
Wuhletal
Kaulsdorf
Mahlsdorf
Birkenstein
Hoppegarten (Mark)
Elsterwerdaer Platz
Biesdorf-Süd
Wuhlheide
Köpenick
Hirschgarten
Friedrichshagen
Rahnsdorf
Wilhelmshagen
Erkner
Eisenhüttenstadt
Fangschleuse
A
B
C
Wünsdorf-Waldstadt

Banquet on Saturday

Time: 19:15-22:00

Exploratorium Potsdam

While visiting the Exploratorium Potsdam, prepare yourself for some real miracles. Science is a real-life experience here, being both tangible and understandable. The roughly 100 exhibits to touch and play with are not only fun, but also educational at the same time. And the best part? Visitors are active particpants in this spectacle.

Ever wondered how a chocolate-covered cream cake explodes? Or how a water-jet transports light? Or even why an electric current makes your fingers prickle? The Exploratorium Potsdam answers these and many more questions around natural phenomena.

Located in Babelsberg, a growing center of the media industry, this interactive world is a true adventure land, where exciting experiments await curious explorers. The Exploratium Potsdam is open to children, students, and families.

Special seasonal topic areas are being developed for any time of the year in cooperation with scientists from Potsdam.

A **bus transfer** will be arranged from the HPI to the Exploratorium and back for all participants.

Address:

Wetzlarer Straße 46
14482 Potsdam
Tel.: +49(0)331 877 36 28
http://www.exploratorium-potsdam.de

Essential German Words and Expressions

Living in Germany without a sufficient knowledge of the German language is not easy. In this section, we provide some essential German words and expressions that could help you moving around and eliminate any complications and inconvenience.

Pronunciation

German is a very close relative of English, so pronunciation isn't too terribly difficult. Just keep in mind that vowels are crisp and clear, so the German o is really one simple sound, not like the oh-oo-wuh of English!

Some important vowel sounds to remember are: *ei* as in line, *ie* as in lean, *ö* as in worm, but without the *r*, *ü* as in tea, but with the lips rounded, *ä* as in get, and *eu* or *äu* as in boy. The consonant combination *sch* is pronounced as in shoe, and *sp* and *st* are pronounced as shp and sht. You'll sometimes see *ß*, which is pronounced like boss. Finally, *z* is pronounced like cats. A few letters have different pronunciation in German; German *j* is pronounced as an English *y*, *v* as *f*, *w* as *v*.

German	English	German	English
Ja/ Nein	Yes / No	Hallo	Hello
Danke	Thank you!	Tschüss!	Bye bye!
Wie geht es Ihnen?	How are you?	Frau/Herr	Ms./Mr.
Gut	Good	Freut mich	Nice to meet you
Entschuldigung	Sorry / Excuse me	Hilfe!	Help!
Wie heißen Sie?	What's your name?	Freitag/Samstag/ Sonntag	Friday/Saturday/Sunday
Bahnhof	(train) Station	U-Bahn / S-Bahn	Subway / Urban train
Fleisch	Meat	Hähnchen	Chicken
Wasser	Water	Ich bin Vegetarier	I am vegetarian
Sprechen Sie Englisch?	Do you speak English?	Wie spät ist es?	What time is it?
Wo ist die U-Bahn?	Where is the subway?	Kann ich ins Internet gehen?	Can I go online?
Wieviel kostet das?	How much does that cost?	Können Sie mir helfen?	Can you help me?
Gibt es hier eine öffentliche Telefonzelle?	Is there a public phone here?	Wo ist das WC?	Where is the bathroom?
Woher kommen Sie?	Where do you come from?	Ich komme aus Deutschland	I come from Germany

15th International Conference on Information Quality (ICIQ-2010)

UALR, Little Rock, Arkansas USA
November 12-14, 2010
http://ualr.edu/iciq2010

BACKGROUND

The *International Conference on Information Quality* (ICIQ) attracts researchers and practitioners from the academic, public and private sectors. Traditionally held on the MIT Campus in Cambridge, Massachusetts, the 2010 conference will be hosted by the Donaghey College of Engineering and Information Technology at the University of Arkansas at Little Rock (UALR), the first university to offer graduate degree programs in information quality.

The conference program will include tracks of research papers, practice-oriented papers, and panel sessions. Accepted papers will compete for the *Stuart E. Madnick IQ Best Paper Award* (US$1,000). Further, the best paper and other high-quality papers will be considered for fast-track to the *ACM Journal of Data and Information Quality*. The 15th *International Conference on Information Quality* will be held at UALR from November 12th (Friday, 5-6:30 p.m.) to November 14th (Sunday noon), 2010.

TOPICS OF INTEREST

Enterprise Architecture Deployment
Social Aspects of IQ
Knowledge Strategies in IQ
Entity Identity Resolution and Management
Metadata and IQ
Business Intelligence and IQ
Data Governance and IQ
IQ Concepts, Tools, Metrics, Measures, and Models
IQ Policies and Standards, IQ Assessment
IQ Practices: Case Studies and Experience Reports
IQ Product Experience Reports
Cost/Benefit Analysis of IQ and IQ Improvement
Information Product Theories and Practices
IQ in Databases, the Web, and e-Business
Data Warehouses and Data Mining
Corporate Household Data
IQ in Scientific Data Management
Data Cleansing and Reconciliation
IQ Education and Curriculum Development
Trust, Knowledge, and Society in the IQ Context
Customer Data Integration
IQ and Formal Semantic Analysis
IQ in Situational Awareness
IQ in Sensor Networks
Information Fusion

ICIQ-2010 IMPORTANT DATES

July 2, 2010:	Submission deadline (firm)
August 13, 2010:	Notification of acceptance
August 29, 2010:	Early registration deadline
September 5:	Camera-ready copy due
November 12-14:	Conference

SUBMISSION GUIDELINES

Each submission must be identified as a completed academic paper, research-in-progress, or practice-oriented paper. Practice-oriented papers may be submitted as PowerPoint presentations. Please download the paper or PowerPoint template from http://ualr.edu/iciq2010.

ICIQ-2010 CONFERENCE CHAIRS

John R. Talburt, University of Arkansas at Little Rock (jrtalburt@ualr.edu)

Andy Koronios, University of South Australia, Adelaide (Andy.Koronios@unisa.edu.au)

Ying Su, Institute of Scientific and Technical Information of China, Beijing (suy.istic@gmail.com)

ICIQ-2010 PROGRAM CHAIRS

Elizabeth Pierce, University of Arkansas at Little Rock (expierce@ualr.edu)

Ismael Caballero Muñoz-Reja, Universidad de Castilla-La Mancha (Ismael.Caballero@uclm.es)

Anne Marie Smith, EWSolutions/North Centeral University (asmith@manieri.com)

Diane Strong, Worcester Polytechnic Institute (dstrong@wpi.edu)

Poster Sessions

ICIQ-2010 will also include poster sessions, providing the opportunity to present early research results, work-in-progress and to engage in one-on-one, informal discussions with colleagues and peers.

Little Rock Attractions

The capital city of Arkansas located on the banks of the Arkansas River boasts a newly developed River Market district between the Capital Building and the Clinton Presidential Library. The River Market is home to several restaurants, bars, shops, parks, museums, historic buildings, and the new River Rail Trolley System. Only a short drive from Little Rock is Hot Springs National Park with historic bathhouses built over natural thermal hot springs.

As of June 19, 2009

Notes

Notes

Notes

International Conference on Information Quality 2009

November 7-8, 2009

Hasso-Plattner-Institut
IT Systems Engineering
University of Potsdam
Prof.-Dr.-Herlmert-Str. 2
14482 Potsdam
Germany

http://www.iciq2009.org

http://wwww.hpi.uni-potsdam.de