

Marcos Antonio Vaz Salles

[marcos \[dot\] vazsalles \[at\] inf \[dot\] ethz \[dot\] ch](mailto:marcos.vazsalles@inf.ethz.ch)

+41 76 401 6526 / +41 44 632-8403

Universitätstrasse 6 – CAB F 52

CH-8092 – Zurich – Switzerland

Member of the Association for Computing Machinery (ACM), of the Institute of Electrical and Electronics Engineers (IEEE), and of the Brazilian Computing Society (SBC)

Research Interests

- Pay-as-you-go Information Integration
- Dataspaces
- Personal Information Management

Current Affiliation

Research Assistant (Ph.D. Student) in Computer Science

ETH Zurich

May, 2005 – November, 2008

Working on the iMeMEX Dataspace Management System (www.imemex.org).

I have defended my PhD thesis November 7, 2008 and **will start as a postdoc at Cornell University in December, 2008.**

Education

M.S. in Informatics

PUC-Rio – Catholic University of Rio de Janeiro

March, 2002 – July, 2004

GPA: 1.0 / 1.0

Master's Thesis: "Autonomic Index Creation in Databases", supervised by Prof. Dr. Sérgio Lifschitz. The thesis involved implementation work in C/C++ with direct modifications to the PostgreSQL open-source DBMS.

B.S. in Computer Science

UNICAMP – State University of Campinas

February, 1995 – December, 1998

GPA: 0.92 / 1.0 (best GPA of his class)

Publications

Conference and Journal Papers

1. Jens Dittrich, Lukas Blunschi, Marcos Antonio Vaz Salles. Dwarfs in the Rearview Mirror: How Big are they really? International Conference on Very Large Databases (VLDB), August 2008, Auckland, New Zealand.

2. Marcos Antonio Vaz Salles, Jens Dittrich, Lukas Blunschi. Adding Structure to Web Search with iTrails. Workshop on Information Integration Methods, Architectures, and Systems (IIMAS; in conjunction with ICDE), April 2008, Cancún, México.
3. Jens Dittrich, Marcos Antonio Vaz Salles, Lukas Blunschi. Managing Personal Information Using iTrails. Personal Information Management Workshop (PIM; in conjunction with CHI), April 2008, Florence, Italy.
4. Marcos Antonio Vaz Salles, Jens-Peter Dittrich, Shant Kirakos Karakashian, Olivier René Girard, Lukas Blunschi. iTrails: Pay-as-you-go Information Integration in Dataspaces. International Conference on Very Large Databases (VLDB), September 2007, Vienna, Austria.
5. Cristian Duda, Jens-Peter Dittrich, Björn Jarisch, Donald Kossmann, Marcos Antonio Vaz Salles. Bringing Precision to Desktop Search: A Predicate-based Desktop Search Architecture (Poster Paper). IEEE International Conference on Data Engineering (ICDE), April 2007, Istanbul, Turkey.
6. Jens-Peter Dittrich, Lukas Blunschi, Markus Färber, Olivier René Girard, Shant Kirakos Karakashian, Marcos Antonio Vaz Salles. From Personal Desktops to Personal Dataspaces: A Report on Building the iMeMex Personal Dataspace Management System. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW), March 2007, Aachen, Germany.
7. Marcos Antonio Vaz Salles, Sérgio Lifschitz, Eduardo Maria Terra Morelli. Towards Autonomic Index Maintenance. Brazilian Symposium on Databases (SBBDB), October 2006, Florianópolis, Brazil.
8. Marcos Antonio Vaz Salles, Jens-Peter Dittrich. iMeMex: A Platform for Personal Dataspace Management. VLDB Ph.D. Workshop, September 2006, Seoul, South Korea.
9. Jens-Peter Dittrich, Marcos Antonio Vaz Salles. iDM: A Unified and Versatile Data Model for Personal Information Management. International Conference on Very Large Databases (VLDB), September 2006, Seoul, South Korea.
10. S. Lifschitz, M. Salles. Autonomic Index Management (Poster Paper). IEEE International Conference on Autonomic Computing (ICAC), June 2005, Seattle.
11. R. Costa, S. Lifschitz, M. Noronha, M. Salles. Implementation of an Agent Architecture for Automated Index Tuning. International Workshop on Self-Managing Database Systems (SMDB; in conjunction with ICDE), April 2005, Tokyo, Japan.
12. R. Costa, S. Lifschitz, M. Salles. Index Self-tuning with Agent-based Databases. CLEI Electronic Journal – Special Issue of Best Papers presented at CLEI'2002, December 2003, Volume 6, Number 1, 22 pages, <http://www.clei.cl/cleiej/paper.php?id=88>.
13. M. Salles, S. Lifschitz. An Agent-based Architecture for the Index Self-tuning Problem in Relational Databases (in Portuguese), Ph.D. and M.S. Workshop, Brazilian Symposium on Databases (SBBDB), October 2003, Manaus, Brazil
14. M. Salles, C. Medeiros, F. Pires, J. Oliveira. Development of a Computer Aided Geographic Database Design System, Brazilian Symposium on Databases (SBBDB), October 1998, Maringá, Brazil

15. M. Salles, C. Medeiros, F. Pires, J. Oliveira. Development of a Computer Aided Geographic Database Design System (in Portuguese), GIS Brazil 98, May 1998, Curitiba, Brazil

Demos

1. Lukas Blunski, Jens-Peter Dittrich, Olivier René Girard, Shant Kirakos Karakashian, Marcos Antonio Vaz Salles. A Dataspace Odyssey: The iMeMex Personal Dataspace Management System. Conference on Innovative Data Systems Research (CIDR), January 2007, Asilomar, USA.
2. Jens-Peter Dittrich, Marcos Antonio Vaz Salles, Donald Kossmann, Lukas Blunski. iMeMex: Escapes from the Personal Information Jungle. International Conference on Very Large Databases (VLDB), August 2005, Trodheim, Norway.
3. M. Salles, S. Lifschitz. A Software Agent for Index Creation on PostgreSQL (in Portuguese). I Demo Session – Brazilian Symposium on Databases (SBBD), October 2004, Brasília, Brazil. This demo was awarded the 2nd place prize among other 10 demos presented at the session.

Tech Reports

1. Jens-Peter Dittrich, Cristian Duda, Björn Jarisch, Donald Kossmann, Marcos Antonio Vaz Salles. Bringing Precision to Desktop Search: A Predicate-based Desktop Search Architecture. Tech Report, ETH Zurich, November 2006.
2. S. Lifschitz, A. Milanés, M. Salles. State of the Art in Relational DBMS Self-tuning (in Portuguese). Tech Report PUC-RioInf.MCC35/04, PUC-Rio, October 2004.

Selected Presentations

1. Pay-as-you-go Information Integration in Personal and Social Dataspace. PhD Defense, November 2008, Zurich, Switzerland
2. iTrails: Pay-as-you-go Information Integration in Dataspace. International Conference on Very Large Databases (VLDB), September 2007, Vienna, Austria.
3. From Personal Desktops to Personal Dataspace: A Report on Building the iMeMex Personal Dataspace Management System. GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web (BTW), March 2007, Aachen, Germany.
4. iMeMex: A Platform for Personal Dataspace Management. VLDB Ph.D. Workshop, September 2006, Seoul, South Korea.
5. Autonomic Index Creation in Databases. DBIS/IKS Groups' Research Seminar at ETH Zurich, May 2005, Zurich, Switzerland.

Previous Work Experience

- **Software Architect**
GAPSO Advanced Planning
A software house specialized in Advanced Planning Systems (APS)
August, 2004 – March, 2005
Started a software development process improvement program in the company.
Acted as technical lead on a helicopter logistics planning project for Petrobras, Brazil's largest oil company.
Technologies used: Oracle Database, ERWin, Rational Rose, Java, XML, C/C++.
- **Founding Partner**
INMETRICS
A start-up specialized in IT systems monitoring
January, 2002 – May, 2004
Acted on the design and implementation of monitoring tools for databases and operating systems. Worked on adapting the company's technology to the IT infrastructure used in Brazilian financial institutions, with an emphasis on settlement systems. Worked on the adoption of the Rational Unified Process (RUP) for the company's software development process. Participated on an Oracle Forms migration and development project for a large telecom firm. Worked on pre-sales of HP and Oracle solutions on Cobra Technologies, a systems integrator owned by one of Brazil's largest public banks.
Technologies used: Oracle Database, Oracle Forms, Oracle Reports, Rational Rose, Java, XML.
- **Systems Analyst**
Software Design Informática
A medium-sized systems integrator and Oracle Partner
March, 1999 – January, 2002
Lead the company's Financial Market Solutions area in the Rio de Janeiro's office. Planned, installed and configured systems for savings accounts, loans, open market, payments, transfers and settlements in several investment banks. Acted on process redefinition to improve maintenance and installation activities for the company's banking systems. Participated on the development of Internet and Email Banking systems. Lead a team of analysts on the installation of a settlement solution in response to changes on the sector's legal requirements.
Technologies used: Oracle Database, Oracle Designer, Oracle Forms, Oracle Reports, ERWin, Java, XML.
- **Programmer**
Software Design Informática
A medium-sized systems integrator and Oracle Partner
June, 1998 – March, 1999
Developed tailor-made systems for clients such as HBO Brazil and Eaton Transmissions.
Technologies used: Oracle Database, Oracle Designer, Oracle Forms, Oracle Reports.

Teaching and Mentoring Experience

- Co-supervision of semester and master theses
ETH Zurich, 2005 – present day
 - Stefan Hildenbrand, Indexing under High Update Rates, Master Thesis, in progress
 - Sandro Blum, A Generic Merge-Based Dynamic Indexing Framework for iMeMex, Master Thesis
 - André Schmidt, Semi-Automatic Trail Creation in iMeMex, Master Thesis
 - Steven Fluck, Scripting Language Plug-ins for iMeMex, Semester Thesis
 - Tobias Abt, iMeMex Indexing Framework, Semester Thesis
 - Lukas Blunsi, Cost-based Optimization in iMeMex, Master Thesis
 - Olivier Girard, Query Processing and Pay-as-you-go Information Integration in iMeMex, Master Thesis
 - Markus Färber, Verteilte Dataspaces mit iMeMex, Master Thesis
 - Sibylle Dürr, iMeMex Web Client, Semester Thesis
 - Sandro Blum, Native OS Integration of iMeMex, Semester Thesis
 - Aarno Aukia, Backup and Recovery in iMeMex, Semester Thesis
 - Philip Stutz, Testing and Deployment of iMeMex, Semester Thesis
 - Olivier Girard, RSS/ATOM Integration in iMeMex, Semester Thesis
 - Georg Troxler, A Resource-aware Thread Scheduler for iMeMex, Semester Thesis
 - Marco Steybe, XML on the Desktop using iMeMex, Semester Thesis
 - Markus Färber, Distributed/P2P iMeMex, Semester Thesis
- Teaching Assistant of the Information Systems Laboratory lecture
ETH Zurich, FS 2007
- Teaching Assistant of the Data Warehouses lecture
ETH Zurich, SS 2007 and SS 2006
- Teaching Assistant of the Architecture and Implementation of Database Systems lecture
ETH Zurich, WS 2006/2007 and WS 2005/2006
- Lecturer and Coordinator of the Database Tuning and Administration I and II courses
Extension Program (CCE) – PUC-Rio, 2002 – 2005
- Lecturer of the “Oracle Administration” and “Oracle Tuning” courses
Petrobras Corporate University – in partnership with PUC-Rio
These courses were repeated four times during 2004 and once in 2005

Other Professional Activities

- External reviewer for the conferences: ICSSOC 2005, ICDE 2006, ICDCS 2006, SIGMOD 2006, CIKM 2007, SIGMOD 2007, VLDB 2007 – Demo Track, EDBT 2008, SIGMOD 2008 – Repeatability Committee, EDBT 2009.
- PC member for IIMAS 2008.

Selected Courses and Certifications

- Didactic Workshop
This three day course included video footage aiming to improve teaching skills.
ETH Zurich, October 2006
- Introduction to CMMI: Staged and Continuous
This SEI-certified course enables participation in CMMI appraisal teams.
Rio de Janeiro, January 2005

Languages

- Portuguese: Native
- English: Proficient
Obtained certificates from Michigan University, USA, and from Cambridge University, England (CAE).
- German (A2-B1), French, Spanish (A1): Beginner

References

Prof. Dr. Donald Kossmann
Professor, Department of Computer Science, ETH Zurich, Switzerland
Email: donald.kossmann@inf.ethz.ch

Dr. Jens-Peter Dittrich
Senior Scientist, Department of Computer Science, ETH Zurich, Switzerland
Email: jens.dittrich@inf.ethz.ch

Prof. Dr. Sérgio Lifschitz
Associate Professor, Department of Informatics, PUC-Rio, Brazil
Email: sergio@inf.puc-rio.br