

Curriculum Vitae: Dr. Andrei Lopatenko

+1 408 480-10-55

lopatenko@google.com

<http://andrei.lopatenko.com/>

- WORK ◇ **Sr. Software Engineer**, Google Inc., Mountain View, CA (Sep 2007 – present)
- EXPERIENCE Search Quality
- Working on the Quality of the Web Search
- ◇ **Software Engineer**, Google, Switzerland, Zurich (Sep 2006 – Sep 2007), transferred to USA
- Search Quality
- Working on Web Search
- ◇ **Assistant Professor (Researcher)** (Promoted), Free University of Bozen-Bolzano (Jan 2006 – Sep 2006)
- ◇ **Teaching and Research Assistant**, Free University of Bozen-Bolzano (October 2002 – Jan 2006)
- Developed an application "Data Cleaner" (C++, STL, Qt, Oracle) for Data Cleansing. The application cleanses database to make it satisfy a set of constrained imposed by a database administrator. In contrast to traditional database technologies which delete "dirty" data, Data Cleaner updates numerical attributes as minimum as possible. The development was motivated by census, bioinformatics and enterprise information integration applications. Designed efficient algorithms to process large inconsistent data sets. Algorithms includes efficient memory management, heuristics. Developed GUI application (Qt4) to administer data cleansing application manually and a set of file parsing and task management tools for command-line administration (flex, bison++). The application can be integrated with any ODBC complaint databases (unixODBC) and with Oracle by native Oracle libraries.
 - Developed an application ("coDBz") for P2P Data Integration (Java, JDBC, Java RMI, JXTA, Java DOM2, Swing) to perform efficient data exchange and query answering in dynamic large-scale networks of heterogenous relational databases (tuned for Oracle and PostgreSQL environment).
 - coDBz is the most efficient framework developed in Academia for query answering and data exchange over multiple database in dynamic networks
 - obtain a set of complexity results for query answering in P2P data integration system and developed efficient algorithms for different settings and network topologies: social and web networks, hierarchical enterprise networks, horizontal structured networks, random networks.
 - obtained a set of new results in computational complexity and algorithms for query answering and "database repair" problems: given an inconsistent in respect to a set of integrity constraints database, find the best repair of the database. The research is motivated by census, bioinformatics and enterprise data integration applications.
- Teaching duties: Introduction into Database and Data Modeling
- Developed lab. assignments for IBM DB2 and PostgreSQL exercises
 - Taught use of SQL database management systems and data modeling in EER.
- ◇ **Visiting Researcher**, The University of Queensland, Brisbane, Australia, Sep 2005, 1 month
- Visited Quantum Information Science Initiative.

- Adiabatic Quantum Computing
- ◇ **Visiting Researcher**, Carleton University, Ottawa, Canada, Feb. 2004, 1 month
 - Visited a research group of Prof. Bertossi
 - Have done a research in computational complexity and approximation of the consistent query answering
- ◇ **Programmer**, Vienna Technical University (TUWIEN), Vienna, Austria, (July 2001 - April 2002)
 - Developed University Information System (Oracle, JSP) for current current research information. Designed and tuned Oracle database for maximum performance under high workloads
 - Implemented ETL (Extract-Transform-Load) and Web Services(SOAP, XML) solutions for data exchange between universities and national-wide research institutions (AURIS-MM)
 - Performed data cleaning of 5 current research information databases (Oracle)
 - Performed ETL applications between a set of heterogenous applications for data exchange (Oracle)
 - Re-designed and unified 5 databases to implement efficient data exchange solutions (Oracle Designer)
 - As a chair of CERIF (international Special Interest Group) improved, cleaned and extended CERIF (Comon European Research Information Format) - relational database and XML format to store and exchange current research information.
- ◇ **Researcher**, Computing Centre of Russian Academy of Science, Moscow, Russia, (Sep 1997 – June 2001)
 - Developed information system for Russian Fund of Basic Research (Oracle Designer, MS SQL, JSP, Java). The system supports full-text and attribute-based search services integrated with documentflow and workflow.
 - Designed and Developed ETL (Extract-Transformation-Load) application for a semi-automatic data exchange between heterogeneous databases (Oracle, W3 XML, W3 RDF, OMG CWM formats, OMG MOF repository)
 - Developed data integration/exchange API (Java, Oracle, W3 RDF) compatible with Oracle ETL API (Java).
 - Developed software modules to describe data mapping and automatic data export from relational databases into RDF and XML data formats.
 - Developed Adaptive Document Flow Manager (Oracle Designer, Oracle DBMS, Oracle Reports, Java, JSP) and integrated it at both database (data transformation, replication and distributed transactions) and application (API) level with Documentum(Enterprise Management System based on Oracle RDBMS)
- ◇ **Software Engineer**, Quarta Ltd, Moscow, Russia, (1996-1997)
 - Developed Customer Relation Management System(MS SQL, Oracle, Visual Basic) for telecommunication company (cell phones and pagers)
 - Cleaned legacy data and imported them into newly designed CRM database (row files, MS Access). Performed ETL operations (Perl, T-SQL, PL/SQL).
 - Developed data analysis and decision support application for Oracle and MS SQL Server(SQL, MS Visual C++)
 - Maintained production Oracle database
 - Developed data mining application(Oracle, OLAP) for Internet Service Providers
- ◇ **DB Developer and Application Programmer**, Programbank Ltd, Moscow, Russia, (1994-1996)

- Designed a database (Oracle) and developed GUI Application (Gupta SQL Windows, C++) for currency exchange operations in financial sector (banks)
- Configured and Tuned Oracle Databases
- Performed a range of ETL operations, to import legacy data. Developed good ETL skills.

- EDUCATION
- ◇ **The University of Manchester**, Manchester, UK. 2002-2006.
(former: The Victoria University of Manchester.)
Ph.D. in Computer Science (awarded Jan 2007).
Dissertation: *Logic Based Data Integration*.
Supervisor: *Prof. Enrico Franconi*.
 - ◇ **Moscow State University of Physics and Technology (PhysTech)**, Moscow, Russia. 1994-1997.
Master of Science.
Major: Applied Mathematics and Information Technology. Speciality: Control Theory and Operational Research. Minor: Nonlinear Optics
Master's project: Maintenance of Data Integrity in Distributed Databases.
Average Gr. 3.7/4
 - ◇ **Moscow State University of Physics and Technology (PhysTech)**, Moscow, Russia. 1991-1994
Bachelor Of Science.
Major: Theor. Physics, Minor: Molecular and Chemical Physics.
Bachelor project ("The State Examination"): Squeezed Light. Theoretical Foundations and Modeling by Means of Non-quantum Physics *under supervision by A. V. Massalov, Lebedev Physical Institute. Average Gr. 3.95/4*
- AWARDS, INTERNSHIPS
- ◇ **2005.** "Quantum Information Visiting Fellowship" to visit *Quantum Information Science group, the University of Queensland, Australia*. September 2005, duration: 5 weeks.
 - ◇ **2005.** The WUN (Worldwide University Network) grant to visit *UCSD, San Diego, California, USA (Prof. Victor Vianu)*, 2075 GB pounds. Expected time is Jan 2006, duration: 4 weeks (did not visit due to clash with local duties).
 - ◇ **2003.** The WUN (Worldwide University Network) grant to visit *Univ. of Washington, Seattle, Washington, USA (Prof. Alon Halevy)*. 2000 GB pounds. July 2003, duration: 4 weeks.
 - ◇ **2002.** Scholarship of the University of Manchester, Computer Science Departments. Fees: 9500 GB pounds per years, scholarship: 7800 GB pounds per year, total 17300 GB pounds per years.
- RESEARCH INTERESTS
- Logic Based Data Integration, Inconsistent Databases, Approximation Algorithms, Data Integration and Query Processing for Large Scale networks (including Social Networks), Dynamic Algorithms, PCP theorem "holographics proofs", Quantum Computing, Adiabatic Quantum Computing
- RESEARCH PROJECTS
- ◇ KnowledgeWeb (FP6-507482)
Advisor: Prof. Enrico Franconi.
Query Answering for Description Logics.
 - ◇ coDBz (Data Coordination for Heterogeneous Databases)
Advisor: Prof. Enrico Franconi
An efficient data integration for large-scale networks of heterogeneous databases
- SELECTED PUBLICATIONS
- Leopoldo E. Bertossi, Loreto Bravo, Enrico Franconi, Andrei Lopatenko: The complexity and approximation of fixing numerical attributes in databases under integrity constraints. *Inf. Syst.* 33(4-5): 407-434 (2008)

^oCurrent as at: Jan 2008

Andrei Lopatenko, Loreto Bravo, "Efficient Approximation Algorithms for Repairing Inconsistent Databases", IEEE ICDE 2007 (23d International Conference on Data Engineering)

Andrei Lopatenko, Leopoldo Bertossi, Complexity of Consistent Query Answering in Databases under Cardinality-Based and Incremental Repair Semantics, ICDT 2007 (11th International Conference on Database Theory) [arXiv:cs.DB/0604002]

Leopoldo Bertossi, Loreto Bravo, Enrico Franconi and Andrei Lopatenko, "Complexity and Approximation of Fixing Numerical Attributes in Databases Under Integrity Constraints", Elsevier's "Information Systems" (Special Issue: Best Papers of DBPL'2005) [arXiv:cs.DB/0503032]

Andrei Lopatenko, Leopoldo Bertossi, Consistent Query Answering By Minimal-Size Repairs, The 2nd International Workshop on Logical Aspects and Applications of Integrity Constraints, 2006

Leopoldo Bertossi, Loreto Bravo, Enrico Franconi and Andrei Lopatenko, "Complexity and Approximation of Fixing Numerical Attributes in Databases Under Integrity Constraints", Tenth International Symposium on Database Programming Languages (DBPL) (co-located with VLDB), 2005

Leopoldo Bertossi, Loreto Bravo, Enrico Franconi and Andrei Lopatenko. "Fixing Inconsistent Databases by Updating Numerical Attributes" , Workshop on Logical Aspects and Application of Integrity Constraints, 2005

Leopoldo Bertossi, Loreto Bravo, Enrico Franconi and Andrei Lopatenko, "Data Cleansing for Numerical Data Sets", Thirteenth Italian Symposium on ADVANCED DATABASE SYSTEMS, (SEBD), 2005

Enrico Franconi, Andrei Lopatenko, "The coDBz Information Integration System for Autonomous Data Sources", Proceedings of the Open Interop Workshop on Enterprise Modelling and Ontologies for Interoperability, Co-located with CAiSE'05 Conference, Porto (Portugal), 2005

Andrei Lopatenko. "Query Answering for LAV and constraints under Exact View Assumption". "Workshop on Exchange and Integration of Data" collocated with CASCON 2004, the 14th Annual IBM Centers for Advanced Studies Conference, IBM Toronto Lab, October 2004

Enrico Franconi, Gabriel Kuper, Andrei Lopatenko, Ilya Zaihrayeu. Queries and Updates in the coDB Peer to Peer Database System, Proc. of Very Large Databases, 2004, (VLDB'04)

Andrei Lopatenko. Data Exchange and Query Answering Under Exact View Assumption in Local As View Data Integration System. AAAI Technical Report, 2004

Andrei Lopatenko. Query Answering Under Exact View Assumption in Local As View Data Integration System. Doctorial Consortium of Ninth International Conference on the Principles of Knowledge Representation and Reasoning (KR-2004), British Columbia, Canada, June 2004, [acceptance rate = 13/34]

Enrico Franconi, Gabriel Kuper, Andrei Lopatenko, Ilya Zaihrayeu. The coDB Robust Peer to Peer Database System, The Second Workshop on Semantics in Peer-to-Peer and Grid Computing at the Thirteenth International World Wide Web Conference, May 17, 2004, New York, USA

Enrico Franconi, Gabriel Kuper, Andrei Lopatenko, Ilya Zaihrayeu. A distributed algorithm for robust data sharing and updates in P2P database networks. "International Workshop on Peer-to-Peer Computing and Databases", collocated with EDBT 2004 [acceptance rate = 12/28]

Enrico Franconi, Gabriel Kuper, Andrei Lopatenko, Ilya Zaihrayeu. The coDB Robust Peer to Peer Database System, SEBD-2004, 12th Italian Symposium on Advanced Database Systems

Leopoldo Bertossi, Enrico Franconi and Andrei Lopatenko. Reconstructing Repairs of Census Data Forms as Database Repairs. Technical Report, Free University of Bozen-Bolzano, 2003.

Enrico Franconi, Gabriel Kuper, Andrei Lopatenko, Luciano Serafini, A Robust Logical and Computational Characterisation of Peer-to-Peer Database Systems pdf, "International Workshop On Databases, Information Systems and Peer-to-Peer Computing", collocated with VLDB'03, LNCS 2944, 2003.

Anne Asserson, Keith G. Jeffery, Andrei Lopatenko, CERIF: Past, Present and Future: An Overview, Gaining Insight from Research Information, 6th International Conference on Current Research Information Systems, CRIS-2002

Andrei Lopatenko, Anne Asserson, Keith G. Jeffery, CERIF - Information Retrieval of Research Information in a Distributed Heterogeneous Environment, 6th International Conference on Current Research Information Systems, CRIS-2002

Keith G. Jeffery, Anne Asserson, Andrei Lopatenko, Comparative Study of Metadata for Scientific Information: The place of CERIF in CRISs and Scientific Repositories, 6th International Conference on Current Research Information Systems, CRIS-2002

Andrei Lopatenko, Information retrieval in Current Research Information Systems, The Workshop on Knowledge Markup and Semantic Annotation at K-CAP'2001

Andrei Lopatenko, Semantic Web Content Accessibility Guidelines for Current Research Information Systems (CRIS), 2001, [arXiv:cs.NI/0107035]

INVITED
TALKS,
SEMINARS

A. Lopatenko, L. Bertossi, "Static, Incremental and Parameterized Complexity of Consistent Query Answering in Databases under Cardinality-Based Semantics", Invited talk to "Second International Workshop on Exchange and Integration of Data", Brixen-Bressanone, Italy, June 2006

L. Bertossi, A. Lopatenko " Cardinality-based semantics for consistent query answering: incremental and parameterized complexity", Invited talk at the "Logic and Databases" Workshop at the Isaac Newton Institute for Mathematical Sciences, Cambridge Univ., Feb-March 2006.

A. Lopatenko, "A robust logical and computational characterization of Peer-to-peer database systems", Feb 5, 2004, Carleton Institute for Computer Science, Carleton University, Ottawa, Canada

A. Lopatenko, "A query answering under Exact Views Assumption in data integration systems", Washington University, Seattle, July 23, 2004

PROF.
ACTIVITIES,
MEMBERSHIP

External Refereeing:

- JoDS, Springer Journal on Data Semantics
- ODBASE 2006, The 5th International Conference on Ontologies, DataBases, and Applications of Semantics
- ACM PODS 2006, , 25th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems
- AiML (Advanced in Modal Logic), 2006
- LAAIC, Logical Aspects and Applications of Integrity Constraints,2006
- ESWC, The 3d Annual European Semantic Web Conference
- IIDB, International Workshop on Inconsistency and Incompleteness in Databases, collocated with EDBT 2006
- IEEE ICDE-06, The 22nd International Conference on Data Engineering
- ACM PODS-05, 24th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems
- DBPL-05, Tenth International Symposium on Database Programming Languages (DBPL) (collocated with VLDB)
- SSTD05, 9th International Symposium on Spatial and Temporal Databases
- DL'05, 2005 International Workshop on Description Logics, collocated with IJCAI-2005

- ESWC'05, 2nd European Semantic Web Conference, 2005
- SBIA-2004, XVII Brazilian Symposium on Artificial Intelligence
- ECDM-2004, Third International Workshop on Evolution and Change in Data Management, collocated with ER 2004
- P2PKM, The First International Workshop on Peer-to-Peer Knowledge Management, collocated with MobiQuitous 2004

In the organizing committee of the 2nd International Workshop on Exchange and Integration of Data (EID-2006); Brixen-Bressanone, Italy, 9-11 June 2006.

PC Chair of "P2P Data Integration For Current Research Information systems" workshop, May 14, 2004, collocated with euroCRIS-2004, Antwerpen, Belgium.

PC Chair of "CERIF-2000 (Common European Research Information Format)", a workshop of 6th International Conference on Current Research Information Systems, Kassel, Germany, August 30, 2002

Membership: ACM SIGACT (Algorithms and Computation Theory), ACM SIGMOD (Management of Data), EATCS (European Association for Theoretical Computer Science).

- SKILLS
- ◇ Excellent knowledge of C++ . Good knowledge of Java, awk, bash (shell scripting). Excellent knowledge of STL (STLPort) and generic programming paradigm.
 - ◇ Extensive knowledge of Oracle DBMS (development, tuning) 10 years, knowledge of MS SQL Server (5 years: 1996-2001), open source DBMS (7 years): PostgreSQL, MySQL. Extensive experience in PL/SQL
 - ◇ Advanced knowledge of relation database design and database tuning for maximum performance (Oracle DBMS, EER, Oracle Designer)
 - ◇ Sophisticated understanding of algorithm theory (especially graph and distributed algorithms) and practical experience in implementation of efficient algorithms in C++ and Java. Good understanding of integer and linear programming and theory of approximation algorithms. Basic knowledge of convex optimization theory.
 - ◇ Good knowledge of development tools: CVS for version control, GCC toolchain, KDevelop and Apple's XCode(C++ development), NetBeans (Java Development)
 - ◇ Excellent skills in organizing database courses (EER conceptual modeling and relational database design, query processing, physical design of database(indexes, tables, files), transaction processing, replication and distributed transactions) and database labs. 4 years
 - ◇ Good knowledge of XML, XML parsing, generating and querying(DOM2 parsers, XPath), XML-Database data exchange (Oracle XML tools, ETL operations)
 - ◇ Linux (9 years), Mac OS X, MS Windows
 - ◇ MCSE (Microsoft Certified System Engineer), 1997, specialization: Windows NT and Microsoft SQL Server
 - ◇ Passed two Oracle exams: Oracle Database, and PL/SQL programming, 1999, score: 95
 - ◇ Experienced in a multinational environment

- SUMMER SCHOOLS
- EDBT Summer School in XML and Databases, Sep 2004, Italy, 6 courses
 - Lipari Summer School in Mobile Networks(Algorithms and Systems), Aug 2004: 8 courses, average grade A
 - European Summer School in Logic and Computation, 2003
 - European Summer School in Logic and Computation, 2002

- PERSONAL INFORMATION
- Nationality: Russian Federation
 - Marital Status: Married
 - Date of Birth: July 8, 1974

REFERENCE Prof. Enrico Franconi, Free University of Bozen-Bolzano, Italy, email: franconi@inf.unibz.it,
tel: +39 0471 016129

Prof. Leopoldo Bertossi, Carleton University, Canada, email: bertossi@scs.carleton.ca, tel:
+1 613 520 2600 x 1627

Dr. Michelangelo Diligenti, Google, Switzerland, email: michi@google.com