DR RENATA BOROVICA-GAJIC CV

Lecturer in Data Analytics
The University of Melbourne
School of Computing and Information Systems

Email: renata.borovica@unimelb.edu.au
Website: www.renata.borovica-gajic.com
Linkedin: http://ch.linkedin.com/in/renataborovica

EDUCATION

2010 – 2016	École Polytechnique Fédérale de Lausanne, Switzerland **PhD in Computer Science** Thesis Title: Toward timely, predictable and cost-effective data analytics **8 publications, 4 are CORE A* (including one with best poster runner-up award); 1 US patent; more than 150 career citations**
2003 – 2008	Faculty of Technical Sciences, University of Novi Sad, Serbia **Master in Electrical and Computer Engineering** (integrated MSc and BSc studies) Thesis Title: Historical Server Query Performance Improvement Based on Grid Processing of Dynamic Data **National award for the best MSc thesis in the field of mathematics and computer sciences**

PRESENT AND PAST APPOINTMENTS

2017-now	**The University of Melbourne, School of Computing and Information Systems (Australia)** Lecturer in Data Analytics (Level B)
2010-2016	**École Polytechnique Fédérale de Lausanne (Lausanne, Switzerland)** Research assistant in the Data-Intensive Applications and Systems Laboratory (DIAS) working under the supervision of Professor Anastasia Ailamaki
2013-2013	**MICROSOFT Research (Redmond, WA, USA), duration 3 months** Research Intern in the Data Management, Exploration and Mining Group supervised by Dr. Surajit Chaudhuri, Dr. Vivek Narasayya and Dr. Christian Konig
2005-2010	**DMS Group LTD (Novi Sad, Serbia)** Senior Software Engineer (Database Team)
2008-2009	**SIEMENS Energy Inc. (Minneapolis, Minnesota, USA)** Software Developer - Contractor (Database Team)

GRANTS AND AWARDS

- Early Career Researcher (ECR) Grant, University of Melbourne, 2018.
- EPFL I&C School achievement award for 2015.
- 2015 Travel award for attending the International Conference on Data Engineering (ICDE)
- Best poster runner-up award at ICDE 2015.
- Dositeja Fellowship from the Serbian Government for students studying abroad, 2012-2013.
- EPFL I&C School fellowship student, 2010-2011
- Serbian national award “Mileva Maric Einstein” for the best master thesis in the field of mathematics and computer sciences in 2008.
- Best student of the graduation promotion at the Faculty of Technical Sciences in 2008.
- Proclaimed one of the best 100 students in Serbia by the Government of Serbia in 2007.
- Awards for excellent results achieved at the Faculty of Technical Sciences for 2003-2008 academic years.
- DMS Group Ltd. fellowship student (mentoring program), 2005-2008.

CONFERENCE PRESENTATIONS AND INVITED TALKS

- 2017, Dagstuhl, Germany — Smooth Scan: One Access Path to Rule them All
- 2016, Univ. of Melbourne, Australia — Toward timely, predictable and cost-effective data analytics
- 2016, Oracle Labs, Switzerland — Toward timely, predictable and cost-effective data analytics
- 2016, EPFL, Lausanne, Switzerland — Toward timely, predictable and cost-effective data analytics
- 2015, ICDE, Seoul, Korea — Smooth Scan: Statistics-oblivious Access Paths
- 2012, VLDB, Istanbul, Turkey — NoDB in Action: Adaptive Query Processing on Raw Data
- 2012, DBTest, Scottsdale, Arizona — Automated Physical Designers: What You See is (Not) What You Get

RESEARCH EXPERIENCE AND PRODUCTIVITY

Cost-effective data analysis with Skipper (2016-now)

- To decrease the cost of data analytics services, this project uses cold storage devices as a primary storage for enterprise and cloud-hosted databases, and builds Skipper, an end-to-end query processing framework, on top of it. Skipper shows that cloud vendors can halve the storage cost by offloading data to cold storage, while at the same time it offers a roadmap of a new-generation of database systems that can efficiently exploit properties of cold storage. Overall, Skipper enables substantial cost and energy savings for both cloud providers and users by using this new hardware technology, while at the same time providing efficient data analytics capabilities.
- Collaboration with EPFL, Switzerland and EURECOM, France.
- Outcomes: 1 **CORE A*** publication, follow-up collaboration with Google.

Predictable query performance with Smooth Operators (2015-now)

- Motivated by industrial database vendors' hunt for stable and predictable query performance, this project uses a radical idea of building database systems that morph automatically during query processing from one form into another to adjust to observed properties with respect to data and resource characteristics. For instance, when accessing data during query processing the access strategy called Smooth Scan is progressively transformed into an optimal form based on observed result distribution. Learning from data and quickly adjusting enables the database system to use near-optimal strategies, regardless of the data and query characteristics, hence increasing the system's stability.
- Collaboration with Harvard University, Snowflake Computing, and Google, USA.
- Outcomes: 2 **CORE A*** publications (**best poster runner-up**), an **organizer** of **Dagstuhl Seminar** on Robust Query Processing (2017).

Fast data exploration with NoDB (2012-2016)

- NoDB is a database system that enables fast data exploration by removing data loading as a prerequisite for data analysis. Instead, NoDB allows asking queries directly over files, while at the same time monitoring users' actions and tuning the system to provide good performance. As a result, NoDB significantly shortens the time to first insight, and enables wider database adoption by removing laborious decisions from the users and doing them as a side-effect of users' analysis.
- Collaboration with Raw Labs, Switzerland.
- Outcomes: **2 CORE A*** publications (**> 130** citations), **CACM Research Highlights** invitation, an US **patent**, and a **successful startup** inspired by this work.

TEACHING

- Database Systems (220 students), The University of Melbourne, 2018 s1.
- Database Systems (430 students) with Dr Ida Asadi Someh, The University of Melbourne, 2017 s2.
- Introduction to programming (Java course, 320 students) with Prof Jamila Sam. EPFL, 2014 s1.
- Programming I (C course, 50 students) with Prof Jean-Luc Desbiolles. EPFL, 2014 s2.
- Project in informatics (C++ course, 90 students) with Prof Jamila Sam. EPFL, 2013 s1.
- Computer-aided engineering II (C++ course, 200 students) with Prof Jamila Sam. EPFL, 2013 s2.
- Introduction to OO programming (Java course, 280 students) with Prof Jamila Sam. EPFL, 2012 s1.
- Computer-aided engineering (C++ course, 230 students) with Prof Jamila Sam. EPFL, 2012 s2.
- Introduction to programming (Java course, 230 students) with Prof Jamila Sam. EPFL, 2011 s1.
- Programming (Java course, 200 students) taught by Prof Thomas Lochmatter. EPFL, 2011 s2.

SUPERVISION

PhD Students

- Yixin Xu: "All Nearest Neighbor Applications", PhD underway, co-supervision with Prof Lars Kulik and Dr Jianzhong Qi. The University of Melbourne, 2017.
- Rajesh Chittor Sundaram: "Automated Cleaning of Spatial Data", PhD underway, co-supervision with Dr Martin Tomko and Dr Elham Naghi Zadeh. The University of Melbourne, 2017.

Master Students

- Kai Zhang: "Prefetching in Skipper", one semester project. The University of Melbourne, 2018.
- Rida Fatima: "Explaining Unexpected Query Performance using Machine Learning", three semesters master thesis. The University of Melbourne, 2018.
- Shashank Satish Parab: "Explaining Unexpected Query Performance using Machine Learning", three semesters master thesis. The University of Melbourne, 2018.
- Yuqian Shi: "Explaining Unexpected Query Performance using Machine Learning", three semesters master thesis. The University of Melbourne, 2018.
- Alejandro Rivas: "Multi-way joins in SparkSQL", master semester project. EPFL, Spring 2016.
- Dimitrios Sarigiannis: "Data Analytics on Cold Storage", master semester project. EPFL, Spring 2016.
- Lisa Zhou: "Implementing prefetching in PostgreSQL", master semester project. EPFL, Spring 2016.
- Arnold Lacko: "View matching for SQL workloads", summer intern project. EPFL, Summer 2015.
- Alejandro Naser: "View selection for SQL workloads", summer intern project. EPFL, Summer 2015.
- Alejandro Rivas: "Cardinality injection in PostgreSQL", bachelor semester project. EPFL, Spring 2014.

PUBLICATIONS

- R. Borovica-Gajic, S. Idreos, A. Ailamaki, M. Zukowski and C. Fraser: Smooth Scan: Robust Access Path Selection without Cardinality Estimation, The **VLDB Journal**, full length article, 25 pages, 2018 (to appear).
- Y. Xu, J. Qi, R. Borovica-Gajic, L. Kulik: Finding All Nearest Neighbors with a Single Graph Traversal. **DASFAA**, full length published article, p. 221-238, 2018.
- R. Borovica-Gajic, G. Graefe, A. Lee: Robust Performance in Database Query Processing (Dagstuhl Seminar 17222). **Dagstuhl Reports** 7(5), p. 169-180, 2017.
- R. Appuswamy, R. Borovica-Gajic, G. Graefe, A. Ailamaki: The Five-minute Rule Thirty Years Later and its Impact on the Storage Hierarchy. The 8th International Workshop on Accelerating Analytics and Data Management Systems Using Modern Processor and Storage Architectures (**ADMS**), full length published paper, 8 pages, 2017.
- R. Borovica-Gajic: Toward Timely, Predictable and Cost-effective Data Analytics. PhD Thesis, 227 pages, EPFL, 2016.
- A. Ailamaki, S. Idreos, I. Alagiannis, R. Borovica, and M. Branco: Query Management System and Engine Allowing for Efficient Query Execution on Raw Data Files. (US 9298754)
- R. Borovica-Gajic, R. Appuswamy, and A. Ailamaki: Cheap Data Analytics Using Cold Storage Devices. The 42nd International Conference on Very Large Data Bases (**VLDB**), full length published paper, vol. 9(12), p. 1029-1040, 2016.
- R. Borovica-Gajic, S. Idreos, A. Ailamaki, M. Zukowski and C. Fraser: Smooth Scan: Statistics-oblivious Access Paths. The 31st International Conference on Data Engineering (**ICDE**), full length published paper, p.315-326, 2015.
- I. Alagiannis, R. Borovica-Gajic, M. Branco, S. Idreos, and A. Ailamaki: NoDB: Efficient Query Execution on Raw Data Files. **Communications of the ACM, Research Highlights**, full length published article, vol. 58(12), p. 112-121, 2015.
- I. Alagiannis, R. Borovica, M. Branco, S. Idreos and A. Ailamaki: NoDB in Action: Adaptive Query Processing on Raw Data (demo). The 38th International Conference on Very Large Data Bases (**VLDB**), full length published demonstration, vol. 5(12), p. 1942-1945, 2012.
- R. Borovica, I. Alagiannis and A. Ailamaki: Automated Physical Designers: What You See is (Not) What You Get. The 5th International Workshop on Testing Database Systems (**DBTest**), full length published paper, 9 pages, 2012.
- I. Alagiannis, R. Borovica, M. Branco, S. Idreos and A. Ailamaki: NoDB: Efficient Query Execution on Raw Data Files. ACM International Conference on Management of Data (**SIGMOD**), full length published paper, p. 241-252, 2012.
- T. Heinis, M. Branco, I. Alagiannis, R. Borovica, F. Tauheed and A. Ailamaki: Challenges and Opportunities in Self-Managing Scientific Databases. IEEE Data Engineering Bulletin, full length published article, vol. 34(4), p. 44-52, (**DEB**), 2011.

LEADERSHIP & SERVICE

Membership

- IEEE member, 2015-
- ACM member, 2012-

Service

- Diversity and Inclusion Committee, The University of Melbourne, School of Engineering, 2017-
- Graduate Research Team, The University of Melbourne, School of Engineering, 2018-
- Future Leaders, The University of Melbourne, School of Engineering, 2018-
- Education Committee, The University of Melbourne, School of Engineering, 2018-

Conference Program committee

- 2018: SYSTOR, DBTest, WomENcourage
- 2019: ICDE, SIGMOD

Workshop organizer

- 2017, Dagstuhl Seminar 17222: “Robust Performance in Database Query Processing”

REFERENCES

Professor Anastasia Ailamaki
Full Professor, École Polytechnique Fédérale de Lausanne

Dr Surajit Chaudhuri
Deputy Managing Director, Microsoft Research

Professor Stratos Idreos
Assistant Professor, Harvard University

Dr Goetz Graefe
Principal Scientist, Google