

# Viktor Sanca

Updated: December 2023.

He/Him  
e-mail: [viktor.sanca@epfl.ch](mailto:viktor.sanca@epfl.ch), [viktor.sanca@gmail.com](mailto:viktor.sanca@gmail.com)  
phone: +41 78 649 91 02  
citizenship: Serbian, Hungarian (EU, Swiss work permit)  
website: [www.viktorsanca.com](http://www.viktorsanca.com)

## Education

2018 – *Expected 08.2024* EDIC Doctoral School, Data-Intensive Applications and Systems Laboratory  
PhD program, EPFL, Lausanne, Switzerland  
Advisor: Prof. Anastasia Ailamaki ([anastasia.ailamaki@epfl.ch](mailto:anastasia.ailamaki@epfl.ch))  
2017 - 2018 Master's Research Scholar Program in Computer Science  
DIAS Lab, EPFL, Lausanne, Switzerland  
2013 - 2017 Bachelor with Honors in Electrical and Computer Engineering  
Computing and Control Engineering, Applied Computer Science and Informatics  
Faculty of Technical Sciences, Novi Sad, Serbia. GPA: 10.00/10.00.

## Publications, Conferences, and Invited Talks

- 2023 Context-Enhanced Relational Operators with Vector Embeddings  
Viktor Sanca, Manos Chatzakis, and Anastasia Ailamaki. **Under submission.** Preprint: <https://arxiv.org/abs/2312.01476>  
LARS: Compute and Data Access Efficient Approximations with Layout-Aware Reuse for Sampling  
Viktor Sanca and Anastasia Ailamaki. **Under preparation.**  
Post-Moore's Law Fusion: High-Bandwidth Memory, Accelerators, and Native Half-Precision Processing for CPU-Local Analytics  
Viktor Sanca and Anastasia Ailamaki.  
ADMS – collocated with VLDB, 2023. <https://ceur-ws.org/Vol-3462/ADMS1.pdf> + conference talk.  
E-Scan: Consuming Contextual Data with Model Plugins  
Viktor Sanca and Anastasia Ailamaki.  
CDMS – collocated with VLDB, 2023. <https://ceur-ws.org/Vol-3462/CDMS11.pdf> + conference talk.  
Improving K-means Clustering using Speculation  
Stefan Igescu, Viktor Sanca, Eleni Zapridou, and Anastasia Ailamaki.  
AIDB – collocated with VLDB, 2023. <https://ceur-ws.org/Vol-3462/AIDB2.pdf> + conference talk.  
Chaosity: Understanding Contemporary NUMA-architectures  
Hamish Nicholson, Andreea Nica, Aunn Raza, Viktor Sanca, and Anastasia Ailamaki.  
TPC-TC – collocated with VLDB, 2023. Preprint: <https://infoscience.epfl.ch/record/305948?ln=en>  
LAQy: Efficient and Reusable Query Approximations via Lazy Sampling  
Viktor Sanca, Periklis Chrysogelos, and Anastasia Ailamaki.  
SIGMOD, 2023. <https://dl.acm.org/doi/10.1145/3589319> + conference talk.  
Analytical Engines with Context-Rich Processing: Towards Efficient Next-Generation Analytics  
Viktor Sanca and Anastasia Ailamaki.  
ICDE'23, Special vision paper track. <https://ieeexplore.ieee.org/abstract/document/10184882> + conference talk.  
System, Workload, and Context-Conscious Analytics  
Invited talk at Lyon 1 University, France (Angela Bonifati)  
2022 Sampling-Based AQP in Modern Analytical Engines  
Viktor Sanca and Anastasia Ailamaki.  
DaMoN – collocated with SIGMOD, 2022. <https://dl.acm.org/doi/10.1145/3533737.3535095> + conference talk.  
Accelerating Analytics Through Speculation and Approximation  
Invited talk at Huawei Research Center, Edinburgh (Stratis Viglas and Nikos Ntarmos)  
2021 Accelerating Complex Analytics Using Speculation  
Panagiotis Sioulas, Viktor Sanca, Ioannis Mytilinis, and Anastasia Ailamaki.  
CIDR, 2021. [https://www.cidrdb.org/cidr2021/papers/cidr2021\\_paper03.pdf](https://www.cidrdb.org/cidr2021/papers/cidr2021_paper03.pdf)

## Ongoing Projects Lead

**Efficient, Reusable, and Workload-Conscious Approximate Analytics for Modern Systems**  
Workload-adaptive, online data summaries designed for using high-bandwidth storage and many-core processors.

**Hybrid Model-Relational Analytics: Context-Enhanced Operators and Holistic Optimization**  
Enhancing relational operators with ML-based vector context and designing logical and physical optimizations.

**Resource-Efficient Speculative Algorithms: Novel Data and Task Parallelization Opportunities**  
Speculation combines approximation and repair that enable breaking the serial, data-parallel-only execution order.

**Adapting Relational and Vector Data Processing to Heterogeneous Compute and Memory**  
Navigating the changing heterogeneous hardware landscape with novel compute capabilities and memory hierarchies.

## Fellowships and Scholarships

---

- 2018 - 2019 EPFL EDIC Doctoral Fellowship  
EDIC Doctoral School Fellowship, EPFL
- 2017 - 2018 EPFL Computer Science Research Scholars Program  
Data-Intensive Applications and Systems Lab, EPFL
- 2015 - 2017 University of Novi Sad Scholarship  
Fund for the Facilitation of Progress of Young Researchers, Talented Students, and Artists
- 2014 - 2015 Serbian Ministry of Education, Science, and Technological Progress Fellowship  
State-awarded scholarship

## Awards

---

- 2020 - 2022 Distinguished service award  
Awarded by EDIC Doctoral School, EPFL
- 2022 Teaching assistant award  
Awarded by EDIC Doctoral School, EPFL
- 2017 The best student at the Faculty of Technical Sciences  
For the class that started in 2013/2014 – among all the sections of the Faculty  
Awarded by the Faculty of Technical Sciences, University of Novi Sad  
  
Exceptional award for overall undergraduate studies  
Awarded by the University of Novi Sad (obtained perfect GPA of 10.00/10.00)  
  
Momčilo Momo Novković Charter  
“For the enthusiasm and exceptional results in curricular and extracurricular activities during studies, for dedication to pedagogical work and contribution in promotion of Faculty of Technical Sciences on the national and international level” – bachelor category  
Awarded annually to a student per study level of the Faculty of Technical Sciences
- 2014 - 2016 Exceptional award for accomplishments in studies  
Awarded by the University of Novi Sad (maintaining a GPA of 10.00/10.00)
- 2015 University award for the student scientific and research paper  
*The semantics of Programming Languages*. Awarded by the University of Novi Sad

## Student Theses Supervised

---

- 2023 Sequential Pattern Mining in Very Large Data Streams  
Sebastien Ollquist, Master’s Thesis – Swisscom. Co-supervisor at DIAS lab.
- 2022 - 2023 Distance-Based Anomaly Detection  
Youssef Saied, Master’s Thesis – done at Oracle Zurich. Co-supervisor at DIAS lab.
- 2019 In-Memory Graph Query Runtime Inside Relational Databases  
Ciprian Baetu, Master’s Thesis – done at Oracle Labs Zurich. Co-supervisor at DIAS lab.

## Teaching

---

- 2023 Design of a new undergraduate course: *Data Intensive Systems* (CS-300)  
Professors: Anastasia Ailamaki and Sanidhya Kashyap. Spring semester.  
Assisting in designing an evolution of the *Introduction to Database Systems* (CS-322), emphasizing practical work to reinforce the theoretical concepts, and with a greater focus on synthesizing data management and operating systems.
- 2020 - 2023 Head teaching assistant: *Introduction to Database Systems* (CS-322)  
Professors: Anastasia Ailamaki and Christoph Koch. Spring semester.  
Creating and improving the teaching materials, exams, and infrastructure for 270 students. Mentoring, guiding, and managing 6 junior teaching assistants, and transferring the course materials to support fully online and hybrid teaching and examination for reduced on-site presence.
- 2019 - 2020 Initiated and co-designed *Machine Learning for Database Systems* (CS-726)  
Professors: Anastasia Ailamaki and Christoph Koch. Fall semester.
- 2019 - 2021 Teaching assistant: *Information, Computation, Communication* course (CS-119d)  
Professor: Jean-Cedric Chappelier. Fall semester.
- 2019 Teaching assistant: *Introduction to Database Systems* (CS-322)  
Professors: Anastasia Ailamaki and Christoph Koch. Spring semester.

## Professional Activities

---

2019 - EPFL Computer Science Ph.D. Student Association: Committee Member and Organizer

EPFL, Lausanne, Switzerland. [epic.epfl.ch](https://epic.epfl.ch)

Faculty and industry talk organizer to bring together the faculty, alums, industry, and student community to exchange ideas and experiences. This provides a doctoral-school level platform to get together and interact with distinguished faculty members more directly and informally and a forum for discussing challenges and opportunities in industry and organizations.

Supporting and participating in a program to help improve the graduate school applications of underprivileged students who want but currently lack mentorship opportunities of graduate peers (<https://epic-guide.github.io/ramp>).

2019 - 2021 EU H2020 Project: Sustainable Data Lakes for Extreme-Scale Analytics

EPFL, Lausanne, Switzerland. [www.smartdatalake.eu](http://www.smartdatalake.eu)

Research and development of the in-house high-performance heterogeneous analytical engine Proteus ([www.proteusdb.com](http://www.proteusdb.com)) to enable storage tiering, approximate query processing, and high-performance integration with project components of other participating research and industrial partners. Designing, presenting, and discussing the project steps with the participating partners and preparing, reviewing, and participating in project reporting and presentations, leading to successful project evaluation by a panel of expert reviewers.

2017 - 2018 EU ERC 2017 Proof of Concept: ViDaR: R-enabled large-scale data analytics in ViDa

EPFL, Lausanne, Switzerland. Supervisors: Odysseas Papapetrou, Tahir Azim

Developed an R-based interface for the high-performance in-house analytical engine Proteus to enable faster data analytics with low coding overhead for scientific users while keeping interoperability with the existing R ecosystem. Explored the full system stack primitives written in C++/LLVM for high-performance hardware-conscious analytics, query optimization in Apache Calcite, and the language primitives of the R programming language.

2016 Summer@EPFL: Usability+performance with data structures wrappers for Python

EPFL, Lausanne, Switzerland. Data-Intensive Applications and Systems Lab. Supervisor: Darius Šidlauskas

Enhancing the usability of spatial index research prototypes in C++ for Python users by creating efficient code wrappers tested with Jupyter Notebooks. This approach demonstrated how existing high-performance codebases can be made more user-friendly without needing to rewrite them in the users' preferred languages or platforms.

## Languages

---

Serbian:	Native speaker
English:	Fluent Professional Use ( <i>Certificate in Advanced English, University of Cambridge, grade: A, equivalent level C2</i> )
Hungarian:	Conversational
German:	Beginner
French:	Intermediate ( <i>EPFL Centre des Langues, level A2/B1</i> )

## Technical Skills and Memberships

---

System:	Unix, Windows
Programming:	C++, C, LLVM, VHDL, CUDA, Assembly, R, Java, Scala, Python, SQL, PL/SQL, JavaScript
Software:	Matlab, Oracle DBMS, MySQL DBMS, PostgreSQL, Intel VTune, familiar with Web and Cloud technologies
Memberships:	EPFL IC PhD Student Association Committee member (2020-2023), ACM, IEEE.

## Research Interests

---

Analytical Query Processing, Data Management Systems and Applications, ML for Systems, Systems for ML, Parallel and Distributed Systems, Adaptive Systems, Operating Systems, Cloud Computing, and Hardware-Software Co-Design.