

Gheorghii Guzun

Assistant Professor of Computer Engineering, San José State University

Research area

My research interests are at the intersection of data management and machine learning, and include topics such as, algorithm optimization for machine learning, energy efficiency in data intensive applications, data compression and quantization, large-scale indexing, data analysis and exploration. Many of the research projects are focused on scalability and performance of large data management systems, with a special interest in hardware-driven software design for data analytics, machine learning, and distributed systems. Our work has addressed critical applications such as improving compute infrastructure monitoring systems, energy efficient data centers, and fast modeling of possible flood events to improve intervention response times.

Education

2011–2016 **University of Iowa**, *Ph.D. Electrical and Computer Engineering - August 2016*, Iowa City, IA, USA.

Thesis: Distributed Indexing and Scalable Query Processing for Interactive Big Data Explorations

Advisor: Dr. Guadalupe M. Canahuate

2006–2010 **Technical University of Moldova**, *B.E. Electrical Engineering - June 2010*, Chisinau, Moldova.

Full Government Scholarship, Graduated with honors

Professional Experience

2017–present **Assistant Professor**, *Department of Computer Engineering*, San Jose State University, San Jose, CA.

2016–2017 **Adjunct Assistant Professor**, *Department of Electrical and Computer Engineering*, The University of Iowa, Iowa City, IA.

2016–2017 **Research Assistant**, *Department of Electrical and Computer Engineering, Iowa Institute of Hydraulic Research (IIHR)*, The University of Iowa, Iowa City, IA.

2012–2016 **Graduate Research Assistant**, *Department of Electrical and Computer Engineering*, The University of Iowa, Iowa City, IA.

2012–2016 **Graduate Teaching Assistant**, *Department of Electrical and Computer Engineering*, The University of Iowa, Iowa City, IA.

2010–2011 **Performance and Automation QA Engineer**, *TD Ameritrade*, Chisinau, Moldova,

Awards and Recognitions

2023-2028 CAREER: Scalable and Adaptable Sparsity-driven Methods for more Efficient AI Systems, \$550,306, *PI: Gheorghii Guzun*, National Science Foundation (NSF). Role: **Principal Investigator**

2023 Featured work in SJSU Research Foundation Annual Report (2024), San Jose State University Research Foundation. Role: **Principal Investigator**

2018-2019 Predictive Analytics for SJSU Admission Yield Estimation, \$16,000, *PIs: Gheorghii Guzun, David Anastasiu* Provost Office, San José State. University, San José, CA. Role: **Co-Principal Investigator**

2017-2023 College of Engineering and Department of Computer Engineering Professional Development Grants, \$12,500, San José State University, San José, CA.

2018-2023 Davidson Student Scholar Grants, \$6000 (supporting graduate students' research), San José State University, San José, CA. Role: **Advisor**

2017 University Grants Academy, 20% Release time, San José State University, San José, CA

2013 San Diego Supercomputer Center Summer Institute Scholarship, University of California, San Diego, CA

Teaching at San José State University

2017–2020 **CMPE-130**, Advanced Algorithm Design *F-2017, S-2018, S-2019, S-2020, S-2022*

2023 **CMPE-138**, Database Systems I *F-2023*

2018 **CMPE-139**, Fundamentals of Data Mining *F-2018*

2018-2020 **CMPE-255**, Data Mining *S-2018, F-2018, S-2020, F-2020, S-2021, F2021, F-2022, S-2023, F-2023*

2019 **CMPE-256**, Large Scale Analytics *S-2019, F-2019*

2022 **CMPE-256**, Advanced Data Mining *S-2022*

2023 **CMPE-266**, Big Data Engineering and Analytics *S-2023*

2020-2022 **CMPE-297**, Big Data Algorithms *F-2020, F-2022*

Teaching at University of Iowa

2017 **ENGR:1300**, Introduction to Engineering Computing *Spr 2017*

2016 **ENGR:2730**, Computers in Engineering *Fall 2016*

2011 **ECE:3540**, Communication Networks (Teaching Assistant) *Fall 2011*

Peer Reviewed Publications

Journal Publications

1. Yinghua Qin and **Guzun, Gheorghii**. "Faster Multidimensional Data Queries on Infrastructure Monitoring Systems" *Big Data Research Journal*: (2022) 1-23. **Impact factor: 3.74**

2. **Guzun, Gheorghi**, and Guadalupe Canahuate. "High-Dimensional Similarity Searches using Query Driven Dynamic Quantization and Distributed Indexing" *Distributed and Parallel Databases (2019)*: 1-32. **Impact factor: 1.147**
3. **Guzun, Gheorghi**, and Guadalupe Canahuate. "Hybrid query optimization for hard-to-compress bit-vectors." *The VLDB Journal* 25.3 (2016): 339-354. **Impact factor: 2.69**
4. **Guzun, Gheorghi**, and Guadalupe Canahuate. "Performance evaluation of word-aligned compression methods for bitmap indices." *Knowledge and Information Systems* 48.2 (2016): 277-304. **Impact factor: 2.25**
5. **Guzun, Gheorghi**, Joel Tosado, and Guadalupe Canahuate. "Slicing the dimensionality: Top-k query processing for high-dimensional spaces." *Transactions on Large-Scale Data-and Knowledge-Centered Systems* XIV. Springer Berlin Heidelberg, 2014. 26-50.

Conference Proceedings

1. Arpita Vats, **Gheorghi Guzun**, and David Anastasiu. "CLP: A Platform for Competitive Learning" *Proceedings of the European Conference on Technology Enhanced Learning*. Toulouse, France, 2022. EC-TEL 2022.
2. Yinghua Qin and **Guzun, Gheorghi** "Multidimensional Preference Query Optimization on Infrastructure Monitoring Systems" *In Big Data (Big Data), 2019 IEEE International Conference on*, Los Angeles, CA, 2019. IEEE, 2019.
3. **Guzun, Gheorghi**, and Guadalupe Canahuate. "Distributed query-aware quantization for high-dimensional similarity searches." *Proceedings of the 21st international conference on Extending Database Technology*. Vienna, Austria 2018. ACM EDBT 2018.
4. **Guzun, Gheorghi**, and Guadalupe Canahuate. "Supporting Dynamic Quantization for High-Dimensional Data Analytics." *Proceedings of the ExploreDB'17*. Chicago, IL, USA, 2017. ACM SIGMOD, 2017.
5. **Guzun, Gheorghi**, Josiah C. McClurg, Guadalupe Canahuate, and Raghuraman Mudumbai. "Power efficient big data analytics algorithms through low-level operations." *In Big Data (Big Data), 2016 IEEE International Conference on*, pp. 355-361. Washington, DC, USA, 2016. IEEE, 2016.
6. Tosado, Joel E., **Gheorghi Guzun**, Guadalupe Canahuate, and Ricardo Mantilla. "On-demand aggregation of gridded data over user-specified spatio-temporal domains." *In Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, p. 62. San Francisco, CA, USA, 2016. ACM, 2016.
7. **Guzun, Gheorghi**, Guadalupe Canahuate, and David Chiu. "A two-phase mapreduce algorithm for scalable preference queries over high-dimensional data." *In Proceedings of the 20th International Database Engineering & Applications Symposium*, pp. 43-52. Montreal, QC, Canada, 2016. ACM, 2016.
8. **Guzun, Gheorghi**, Joel E. Tosado, and Guadalupe Canahuate. "Scalable preference queries for high-dimensional data using map-reduce." *In Big Data (Big Data), 2015 IEEE International Conference on*, pp. 2243-2252. Santa Clara, CA, 2015. IEEE, 2015.

9. **Guzun, Gheorgi**, Guadalupe Canahuate, David Chiu, and Jason Sawin. "A tunable compression framework for bitmap indices." *In Data Engineering (ICDE), 2014 IEEE 30th International Conference on*, pp. 484-495. Chicago, IL, USA, 2014. IEEE ICDE, 2014.
10. Doan, Fredton, David Chiu, Brasil Perez Lukes, Jason Sawin, **Gheorgi Guzun**, and Guadalupe Canahuate. "Dynamic bitmap index recompression through workload-based optimizations." *In Proceedings of the 17th International Database Engineering & Applications Symposium*, pp. 96-105. Barcelona, Spain, 2013. ACM, 2013.

Posters

1. David Anastasiu and **Guzun, Gheorgi**. "Using Competitive Learning to Increase Student Engagement" *Academic Technology Expo: ATXPO 2021, Stanford, CA*

Under review and working papers

1. David C. Anastasiu and **Guzun, Gheorgi**. "Active Learning through Competition on Continuous Improvement Problems", *Submitted to International Journal of Educational Technology in Higher Education*, 2023
2. Anurag Patro and **Guzun, Gheorgi**. "Dynamic Quantization for Density Based Clustering in High-dimensional Spaces"
3. **Meera Tresa, Cindy Zhang, Guzun, Gheorgi**. "Leveraging sparsity in DNN data with compressed bit vectors"
4. Keegan P. Shay, Guadalupe Canahuate and **Gheorgi Guzun**. "An Ensemble for Survival Prognosis of Oropharyngeal Cancer Patients using Nearest Neighbors"

Talks and Presentations

- Oct. 2023 *RSCA in Five Speaker: Leveraging Sparsity in AI Systems*, San Jose State University, San Jose, CA, USA
- Apr. 2021 *Accelerating Data Analytics*, Technical University of Moldova, Chisinau, MD
- Feb. 2019 *Predictive Analytics for University Admission Yield Estimation*, Information Technology Department, San Jose State University, San Jose, CA
- May 2018 *Distributed Query-Aware Quantization for High-Dimensional Similarity Searches*, NorCal DB Day 2018, Oracle Conference Center, Redwood Shores, Redwood City, CA
- Mar. 2018 *Query aware quantization for high dimensional similarity searches*, EDBT 2018, Vienna, Austria
- Dec. 2016 *Power efficient big data analytics algorithms through low-level operations*, IEEE BigData 2016, Washington DC, USA
- April 2016 *Distributed nearest neighbor query processing for high dimensional data*, Research Open House, College of Engineering, University of Iowa, Iowa City, IA
- Oct. 2015 *Scalable preference queries for high-dimensional data using map-reduce*, IEEE Big-Data 2015, Santa Clara, CA

- April 2015 *Hybrid query optimization for advanced analytics queries using high density bit-vectors*, Research Open House, College of Engineering, University of Iowa, Iowa City, IA
- April 2014 *Data adaptive compression framework for bit-vector indices*, Research Open House, College of Engineering, University of Iowa, Iowa City, IA
- April 2014 *A tunable compression framework for bitmap indices*, IEEE ICDE 2014, Chicago, IL
- Aug. 2013 *Large-scale indexing using compressed bit-vectors* University of California - San Diego, La Jolla, CA
- April 2013 *V.A.L. - Variable Aligned Length compression for bitmap indices* Research Open House, College of Engineering, University of Iowa, Iowa City, IA

Student Supervision

Masters Thesis

- S2019-F2019 Yinghua Qin: *Multidimensional Query Optimizations for Infrastructure Performance Monitoring Systems*
- F2018-F2019 Utsav Jain: *Using Blockchain Technology for Creating The Organ Procurement and Transplant Network*

Doctoral thesis committee

- o Arina Alexei 2022.

Professional Service

Grant Proposal Reviews

- o *Proposal Review Panelist*: National Science Foundation (NSF). CISE, 2023
- o *Proposal Reviewer*: Chilean funding agency for scientific and technological research, 2019

Journal Reviewer

- o IEEE Access, 2022, 2023
- o IEEE Transactions on Knowledge and Data Engineering, 2019, 2020, 2021
- o Concurrency and Computation: Practice and Experience, 2021, 2022
- o Software Practice and Experience, 2016, 2017

Conference Organizing Committees

- o ECCO 2021: International Conference on Electronics, Communications and Computing, 2021-Advisory Committee Member
- o CBDCOM2020: IEEE International Conference on Cloud and Big Data Computing 2020 - Liaison & Publicity Chair

Conference Program Committee Member

- o ACM RecSys: The ACM Conference on Recommender Systems, 2023
- o TheWebConf: The Web Conference, 2022, 2023

- KDD: ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022
- WSDM: ACM International Conference on Web Search and Data Mining, 2022, 2023
- HiPC: IEEE International Conference on High Performance Computing, Data, & Analytics, 2022
- BDS: International Conference on Big Data Service and Applications, 2020, 2021, 2022
- GeoProcessing: International Conference on Advanced Geographic Information Systems, Applications, and Services, 2019
- EDBT: International Conference on Extending Database Technology, Best Demonstration Award, 2018
- AIFU: International Conference on Artificial Intelligence and Applications, 2018
- DMDB: International Conference on Data Mining and Database, 2018
- CCSEA: International Conference on Computer Science, Engineering and Applications, 2018

Service to the University, San Jose State University

2019-2020 Chair of the Faculty Diversity Committee

2018-2019 Member of the Faculty Diversity Committee

Service to the College of Engineering, San Jose State University

2018-2019 Contributed to setting-up a College of Engineering (CoE) High Performance Computer (HPC): Set-up the Lustre distributed file system for the CoE HPC

2019 Organizer: College of Engineering in partnership with USPS Hackathon "Delivering the future"

Service to the Department of Computer Engineering, San Jose State University

2023-2024 Member of the Graduate Curriculum Committee

2022-2023 Member of the Alumni and Industry Outreach Committee

2019-2022 Chair of the Alumni and Industry Outreach Committee

2020,2023 Member of the CMPE Faculty Recruitment Committee

2018-2019 Member of the Undergraduate Curriculum Committee

2017-2018 Member of the Graduate Curriculum Committee

Affiliations to Professional Organizations

- Association for Computing Machinery (ACM).
- Institute of Electrical and Electronics Engineers (IEEE).