Victor Giannakouris

• vg292@cornell.edu • (+1) 917 945 6484
• Google Scholar. Citations: 87, h-index: 4
• LinkedIn

INTERESTS	I am interested in database systems, and especially in query optimization for federated query execution engines. I have research and industry experience on implementing database system components, ranging from indexes/storage engines (TileDB, MonetDB) to federated query optimizers for Spark SQL. Lately I have been working on leveraging machine learning in federated query optimization.

EDUCATION

Cornell University, Ithaca, NY, USA

- Ph.D., Computer Science   2021 – 2026
  - Area: Database Systems & Machine Learning
  - Advisor: Immanuel Trummer

National and Kapodistrian University of Athens, Athens, Greece

- M.Sc., Computing Systems   Oct 2018 – Jul 2021
  - Thesis: An Online Sample Index for Approximate Query Processing
  - Advisor: Yannis Ioannidis

Ionian University, Department of Informatics, Corfu, Greece

- B.Sc., Informatics   Sep 2012- May 2017
  - Thesis: Optimization of Distributed SQL Query Execution Over Multiple-Engine Environments
  - Advisor: Dimitrios Tsoumakos

EMPLOYMENT

Cornell University, Ithaca, NY, USA

- Graduate Research Assistant   Aug 2021 – Present
  - TA: CS 4321 - Database Systems Implementation (Fall 2022)
  - TA: CS 4410 - Introduction to Operating Systems (Fall 2021, Spring 2022)

IBM Research, Almaden, CA, US

- Research Intern   May 2022 – Aug 2022
  - Database Internals

TileDB, Boston, MA, US

- Senior Software Engineer   Nov 2019 – Jul 2021
  - Database Internals, Storage Engines

Unravel Data Systems, Palo Alto, CA, US

- Software Engineer   Oct 2018 – Oct 2019
  - Big Data Application Performance Management (APM)
  - Developed Unravel Platform's Cloud Migration component (link)

AISERA, Palo Alto, CA, US

- Data Engineer   Apr 2018 – Sep 2018
  - Development of data collection microservices on the company's platform's data ingestion component.
  - Tools: Kubernetes, AWS, Google Cloud Platform, Java, Python

BEAT, Athens

- Big Data Software Engineer   Sep 2017 – Apr 2018
  - Development of streaming and ETL jobs over large-scale data.
  - Tuning and optimization of Spark SQL batch jobs on query and storage level.
  - Tools: Spark, YARN, Hadoop, Hive, Presto, Scala, Python

Computing Systems Laboratory (CSLab), NTUA, Greece

- Researcher   Jul 2015 – Aug 2017
  - R&D on Big Data Management, Databases and Multi-Engine Analytics.
  - Participated in four publications.
  - Development of a middleware optimization framework for Spark SQL, achieving execution time improvement for queries over heterogeneous environments up to one order of magnitude.
  - Participated in the EU Funded Project ASAP(A Scalable Analytics Platform).

isMOOD Data Technology Services

- Software Engineer   Apr 2014 – Sep 2015
  - Data collection and integration into a unified source
  - Development, scheduling & optimization of web crawlers/spiders

- Tools: Scrapy, Python, Elasticsearch, MongoDB

AWARDS	▪ ACM SIGMOD 2019 Student Travel Award	Jun 2019
	▪ IEEE BigData 2016 Student Travel Award	Dec 2016

SKILLS
- **Programming Languages:** Java, Scala, Python, C/C++, SQL
- **Databases:** Hive, Presto, MemSQL, PostgreSQL, MonetDB
- **Data Management - Analytics:** Spark, Hadoop, scikit-learn
- **Infrastructure**: Kubernetes, AWS, Openstack

CERTIFICATIONS	▪ Object Oriented Data Structures in C++	Coursera - UIUC
	▪ SQL (Mini-Course)	Stanford Online
	▪ CS190.1x: Scalable Machine Learning	edX - BerkeleyX
	▪ CS100.1x: Introduction to Big Data with Apache Spark	edX - BerkeleyX

PUBLICATIONS

CONFERENCES

[1] <u>V. Giannakouris</u>, I. Trummer, "Building learned federated query optimizers", *In proceedings of the VLDB 2022 Endowment (PhD Workshop)*, Sydney, Australia, Sep 2022.

[2] <u>V. Giannakouris</u>, A. Fernandez, A. Simitsis,and S. Babu, "Cost-Effective, Workload-Adaptive Migration of Big Data Applications to the Cloud," *In proceedings of the 2019 ACM SIGMOD/PODS, International Conference on Management of Data*, Amsterdam, The Netherlands, Jun 2019.

[3] <u>V. Giannakouris</u>, N. Papailiou, D. Tsoumakos,and N. Koziris, "MuSQLE: Distributed SQL Query Execution Over Multiple-Engine Environments," *In proceedings of the 2016 IEEE International Conference on Big Data (BigData 2016)*, Washington, DC, USA, Dec 2016.

[4] K. Doka, N. Papailiou, <u>V. Giannakouris</u>, D. Tsoumakos and N. Koziris, "Mix 'n Match Multi-Engine Analytics,," *In proceedings of the 2016 IEEE International Conference on Big Data (BigData 2016)*, Washington, DC, USA, Dec 2016.

[5] K. Doka, M. Filatov, <u>V. Giannakouris</u>, V. Kantere, N. Koziris, C. Mantas, N. Papailiou, V. Papaioannou and D. Tsoumakos, "Optimizing, Planning and Executing Analytics Workflows over Multiple Engines.," *In the 1st International Workshop on Multi-Engine Data AnaLytics (MEDAL 2016), in conjuction with the EDBT Conference*, Bordeaux, France, Dec 2016.

[6] N .Papailiou, K. Doka, <u>V. Giannakouris</u>, V. Papaioannou, D. Tsoumakos and N. Koziris, "Robust and Adaptive Multi-Engine Analytics using IReS (demo paper)," *In the 10th International Workshop on Enabling Real-Time Business Intelligence (BIRTE 2016), in conjunction with the VLDB Conference*, New Delhi, India, Mar 2016.

[6] <u>V. Giannakouris</u>, A. Plerou, and S. Sioutas, "CSMR: A Scalable Algorithm for Text Clustering with Cosine Similarity and MapReduce," *In the 10th International Conference on Artificial Intelligence Applications and Innovations*, Rhodes, Greece, Sep 2014.

JOURNAL PUBLICATIONS

[1] K. Doka, I. Mytilinis, N .Papailiou, <u>V. Giannakouris</u>, D. Tsoumakos and N. Koziris, "Multi-Engine Analytics with IReS," *IIn: Castellanos M., Chrysanthis P., Pelechrinis K. (eds) Real-Time Business Intelligence and Analytics. BIRTE 2015, BIRTE 2016, BIRTE 2017. Lecture Notes in Business Information Processing, vol 337. Springer, Cham*, New Delhi, India, Mar 2016.

[2] X. Evangelopoulos, <u>V. Giannakouris</u>, L. Iliadis, C. Makris, Y. Plegas, A. Plerou, S. Sioutas, "Evaluating Information Retrieval using Document Popularity: A case study using MapReduce," *Engineering Applications of Artificial Intelligence*, vol. 89, no. 2, pp. 4123–4133, Feb 2008.

LANGUAGES
- **English** (fluent), **Greek** (native)