

# Georgios Siachamis

[GitHub](#) | [Google Scholar](#) | [LinkedIn](#) | [Personal website](#) | [georgedrfisia@gmail.com](mailto:georgedrfisia@gmail.com)

## Education

---

### Ph.D. in Database Systems

Nov 2019 - Nov 2024

*Delft University of Technology*

*Delft, Netherlands*

**Title:** [Adaptivity for Streaming Dataflow Engines](#)

**Summary:** In my thesis, I investigated the adaptivity capabilities of modern streaming dataflow engines. I primarily focused on three main aspects of adaptivity: adaptivity to input distribution changes (adaptive streaming similarity joins), infrastructure failures (checkpointing for fault tolerance), and input rate changes (autoscaling). In this thesis, I performed two extensive experimental evaluations on checkpointing and autoscaling and proposed a novel solution for adaptive streaming similarity joins.

**Advisor:** Dr. Asterios Katsifodimos

### Diploma (M.Eng.) in Electrical and Computer Engineering

Apr 2014 - Oct 2019

*Technical University of Athens*

*Athens, Greece*

**Thesis:** ["Exploration of query similarity using machine learning"](#) (in Greek)

**Advisor:** Dr. Verena Kantere

GPA: 7.86/10

## Employment

---

### Research Engineer / Postdoctoral researcher

Sept 2024 - Present

*Inria, CEDAR*

*Saclay, France*

- Working on ontology-based data access and graph data lakes.

### Ph.D. Candidate

Nov 2019 - Apr 2024

*Delft University of Technology, Web Information Systems*

*Delft, Netherlands*

- Performing the research related to my PhD topic.
- Participating to the teaching and research activities of the university.

### Academic Consultant

Jan 2020 - Dec 2023

*ING Group, Global Analytics and Tech Infra*

*Amsterdam, Netherlands*

- Member of the [AI for FinTech Research](#) collaboration between ING and TU Delft.
- Bridging the gap between academic research and industry by participating in various team meetings and organizing internal seminars.

### Research Intern

Apr 2019 - Jun 2019

*Université de Cergy-Paris*

*Paris, France*

- Under the supervision of Dr. Aikaterini Tzompanaki
- Title: "Querying Multiversion Datasets"
- Summary: During this three-month internship, I extended the publicly available code of [OrpheusDB](#) to include all the versioning strategies discussed in the paper and designed and performed an extended experimental evaluation.

### Software Engineer

Jan 2015 - Feb 2016

*FOCUS ON DIGITAL SERVICES LTD.*

*Athens, Greece*

- Software development.
- Digital data processing.

# Teaching

---

Delft University of technology

April 2021 - Apr 2024

Faculty of Electrical Engineering, Mathematics and Computer Science

Delft, Netherlands

- April 2021 - April 2024: [IN4331: Web-scale Data Management](#), Teaching Assistant, Master level.

# Student Supervision

---

## Master Theses Supervision

---

- Gianni Wiemers ('23) "[Benchmarking checkpointing algorithms in Stream Processing Engines](#)"
- Theodoros Veneti ('23) "[Online state migration in modern stream processing engines](#)"
- Tomás Hernandez Quintanilla ('23) "[Experimental evaluation of distributed similarity joins in stream processing environments](#)"
- Job Kanis ('23) "[An experimental evaluation of auto-scaling techniques for distributed stream processing systems](#)"
- Kanya Paramita Koesoemo ('21) "[The State of Data Streaming Practices at ING](#)"
- Kyriakos Psarakis ('20) "[Holistic Schema Matching at Scale](#)"

# Languages

---

Greek (native speaker), English (professional working proficiency), French (elementary proficiency).

# Software

---

- [Valentine: \(Schema-\) Matching DataFrames Made Easy](#)
  - A Python package for capturing potential relationships among columns of different tabular datasets, which are given as pandas DataFrames.
  - Contributions: Development of the algorithms included in the package.

# Service

---

- **External Reviewer:** ECMLPKDD 2024, ICDE 2024, SIGMOD 2023, SIGMOD Demo 2021, VLDB 2021
- **Reviewer:** ICDE 2025, EDBT Demo 2025, DEBS 2025

# Publications

---

## International Conferences

---

- [Evaluating Stream Processing Autoscalers](#)  
**G. Siachamis**, G.Christodoulou, K. Psarakis, M. Fragkoulis, A. van Deursen, A. Katsifodimos.  
ACM Conference on Distributed and Event-Based Systems (DEBS) 2024
- [CheckMate: Evaluating Checkpointing Protocols for Streaming Dataflows](#)  
**G. Siachamis**, K. Psarakis, M. Fragkoulis, A. van Deursen, P. Carbone, A. Katsifodimos.  
IEEE International Conference on Data Engineering (ICDE) 2024
- [Styx: Deterministic Transactional Serverless Functions on Streaming Dataflows](#)  
K. Psarakis, **G. Siachamis**, G. Christodoulou, M. Fragkoulis, A. Katsifodimos.  
Under Review - available on arXiv
- [Adaptive Distributed Streaming Similarity Joins](#)  
**G. Siachamis**, K. Psarakis, M. Fragkoulis, O. Papapetrou, A. van Deursen, A. Katsifodimos.  
ACM Conference on Distributed and Event-Based Systems (DEBS) 2023

- **Valentine in Action: Matching Tabular Data at Scale**

C. Koutras, K. Psarakis, **G. Siachamis**, A. Ionescu, M. Fragkoulis, A. Bonifati, A. Katsifodimos.  
In the Proceedings of the VLDB Endowment (PVLDB) 2021 (Vol.14, No. 12.) (demo track)

- **Valentine: Evaluating matching techniques for dataset discovery**

C. Koutras, **G. Siachamis**, A. Ionescu, K. Psarakis, J. Brons, M. Fragkoulis, C. Lofi, A. Bonifati, A. Katsifodimos.  
IEEE International Conference on Data Engineering (ICDE) 2021

## International Workshops

---

- **Towards Evaluating Stream Processing Autoscalers**

**G. Siachamis**, J. Kanis, W. Koper, K. Psarakis, M. Fragkoulis, A. van Deursen, A. Katsifodimos.  
International Workshop on Self-managing Database Systems (SMDB) 2023

## Technical Skills

---

- Programming Languages: Python (advanced), Java (advanced), C (intermediate), JavaScript (basic), HTML (basic).
- Big Data Processing Frameworks: Apache Flink, Apache Hadoop.
- Tools: Docker, Kubernetes, GIT.
- Databases: Postgres, MySQL, Elasticsearch, MongoDB