

# Christos Koutras

## Research Interests

Data Integration, Dataset Discovery, Schema Matching

## Education

- 2018– **PhD Candidate, Computer Science**, *Delft University of Technology (TU Delft)*, Netherlands.  
Research Topic: Scalable Data Integration  
Supervisors: Asterios Katsifodimos, Christoph Lofi, Geert-Jan Houben
- 2016–2018 **Master of Philosophy (MPhil), Computer Science**, *The Hong Kong University of Science & Technology (HKUST)*, Hong Kong.  
Research Topic: Databases, Data Management  
Thesis: Query Estimation Structures in Geo-Social Networks  
Supervisor: Dimitris Papadias  
GPA: 3.94/4
- 2009–2015 **Diploma (M.Eng), Electrical and Computer Engineering**, *National Technical University of Athens (NTUA)*, Greece.  
GPA : 8.24/10  
Major: Computer Science  
Thesis: Monitoring Continuous  $k$ -Nearest Neighbor Queries Over Moving Uncertain Objects  
Supervisors : Kostas Patroumpas, Yannis Vasileiou
- 2006–2009 **High School Degree**, *1<sup>st</sup> High School of Xanthi*, Greece.  
Degree Grade : 19,9/20 (excellent)  
Panhellenic Qualifying Examinations GPA : 19664/20000

## Experience

- April–July 2022 **Applied Scientist Intern**, *Amazon Web Services*, SF Bay Area, USA.  
Internship with the AWS GraphML team, where I worked towards semi-supervised matching among tabular data.
- November–December 2018–present **Course Coordinator**, *Delft University of Technology*, Netherlands.  
CSE1500: Web and Database Technology  
Responsible for organizing the course's DB lab sessions and teaching tutorials, while also for designing the DB assignments.
- February–April 2018–present **Course Coordinator**, *Delft University of Technology*, Netherlands.  
CSE1505: Information and Data Management  
Responsible for organizing the course's lab sessions and teaching tutorials. Also, for designing and grading lab assignments and exam tasks.
- February–May 2018 **Teaching Assistant**, *The Hong Kong University of Science & Technology*, Hong Kong.  
COMP 3311: Database Management Systems  
Responsible for teaching the tutorials of the second half of the course on topics such as : Functional Dependencies and Normalization, Memory Hierarchy and File Structures, Indexing and Hashing, Join Algorithms, Query Processing and Optimization
- September–December 2017 **Teaching Assistant**, *The Hong Kong University of Science & Technology*, Hong Kong.  
COMP 3711: Design and Analysis of Algorithms  
Responsible for designing and grading the course assignments on Greedy Algorithms, Dynamic Programming, Minimum Spanning Trees, Shortest Path Algorithms

July–August **Visiting Student**, *Oxford University*, United Kingdom.  
 2017 Worked on Query Optimization with Access Patterns under Integrity Constraints in collaboration with George Konstantinidis

February–**Teaching Assistant**, *The Hong Kong University of Science & Technology*, Hong Kong.  
 May COMP 3311: Database Management Systems  
 2017 Responsible for teaching the tutorials of the first half of the course on topics such as: ER diagrams, Relational Data Model, Relational Algebra, SQL, Functional Dependencies and Normalization

**External Reviewer.**  
 VLDB 2023, IEEE ICDE 2019, ACM SIGMOD 2018/2021, ACM SIGSPATIAL 2017, SSTD 2017

**PC Member.**  
 DBML 2022 (Workshop in conjunction with ICDE 2022)

## Publications

Under submission/TRL@NeurIPS 2022 **SiMa: Effective and Efficient Matching Across Data Silos Using Graph Neural Networks.**  
 Christos Koutras, Rihan Hai, Kyriakos Psarakis, Marios Fragkoulis, Asterios Katsifodimos

Under submission  
 arXiv **Data Lakes: A Survey of Functions and Systems.**  
 Rihan Hai, **Christos Koutras**, Christoph Quix, Matthias jarke

arXiv **Bridging the Gap between Data Integration and ML Systems.**  
 Rihan Hai, Yan Kang, **Christos Koutras**, Andra Ionescu, Asterios Katsifodimos

CIDR 2022 **Amalur: Next-generation Data Integration in Data Lakes**, *Conference on Innovative Data Systems Research (CIDR) 2022*.  
 Rihan Hai, **Christos Koutras**, Andra Ionescu, Marios Fragkoulis, Asterios Katsifodimos

VLDB 2021 **Valentine in Action: Matching Tabular Data at Scale**, *In Proceedings of VLDB 2021*, Copenhagen, Denmark.  
**Christos Koutras**, Kyriakos Psarakis, George Siachamis, Andra Ionescu, Marios Fragkoulis, Angela Bonifati, Asterios Katsifodimos

IEEE ICDE 2021 **Valentine: Evaluating Matching Techniques for Dataset Discovery**, *In Proceedings of IEEE ICDE 2021*, Chania, Greece.  
**Christos Koutras**, George Siachamis, Andra Ionescu, Kyriakos Psarakis, Jerry Brons, Marios Fragkoulis, Christoph Lofi, Angela Bonifati, Asterios Katsifodimos

SeaData Workshop 2020 **REMA: Graph Embeddings-based Relational Schema Matching**, *In Workshop Proceedings of the EDBT/ICDT 2020 Joint Conference*, Copenhagen, Denmark.  
**Christos Koutras**, Marios Fragkoulis, Asterios Katsifodimos, Christoph Lofi

VLDB Workshop 2019 **Data as a Language: A Novel Approach to Data Integration**, *Proceedings of the VLDB 2019 PhD Workshop*, Los Angeles, California, USA.  
**Christos Koutras**

SSDBM 2017 **Probabilistic k-Nearest Neighbor Monitoring of Moving Gaussians**, *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, Chicago, Illinois, USA.  
 Kostas Patroumpas, **Christos Koutras**

## Projects & Technical Skills

Valentine package  
 Core maintainer of `valentine` python package. With Valentine users can find matches among columns of pandas DataFrames.

Programming languages  
 Ability to develop and implement algorithms in JAVA, C/C++ and Python. Familiarity with MATLAB, PostGre SQL and Git/GitHub.

---

## Awards and Scholarships

- February **HKUST Postgraduate Scholarship.**  
2016 – June Postgraduate scholarship from The Hong Kong University of Science & Technology  
2018
- April 2008 **Hellenic Youth Parliament.**  
Selected as a substitute delegate of the town of Xanthi in Hellenic Youth Parliament through an essay competition

---

## Languages

- Greek **Native Speaker**
- English **Professional Working Proficiency** *CPE Certification, University of Michigan*
- German **Elementary Proficiency** *Zertifikat B1, Goethe Institut*