DIMITRIOS KARAPIPERIS CURRICULUM VITAE

Affiliation: School of Science and Technology, International Hellenic University.
Address: Kato Scholari, 57 500, Thessaloniki, Greece.
E-mail: dkarapiperis@ihu.edu.gr, karapiperis121@gmail.com
Tel: +30 23920 65096 (landline), +30 6974 601618 (cell).

EDUCATION

2013 – 2016. **PhD degree** in Computer Science, awarded by the School of Science and Technology of the Hellenic Open University, Greece. My research interests lie in the field of Entity Resolution (Record Linkage), where I focus on developing similarity algorithms, data structures, approximation schemes, and scalable (distributed) solutions that rely on randomization schemes.

2000 – 2001. **MSc degree** in Software Engineering, awarded by the Computer Science Department of the University of York, UK. Scholarship awarded from the State Scholarships Foundation of Greece (IKY) after participating in national exams.

1995 – 1999. **BSc degree** awarded by the Technological Educational Institute of Thessaloniki, Greece, Faculty of Applied Technology, Information Technology Department.

ACADEMIC EMPLOYMENT

10/2021 – present. Hellenic Open University, Lecturer (tutor) at the School of Science and Technology (course taught in English).

Courses:

- Foundation in Computer Science (Data Mining and Machine Learning techniques)

1/2020 –present. International Hellenic University, Greece. Adjunct lecturer at the School of Science and Technology in the post-graduate programs (all courses taught in English).

Courses:

- Knowledge Management in the Web
- Big Data and Cloud Computing
- Java/Python Programming Language
- Operating Systems (Linux Shell)
- ICT Essentials
- Web Programming (HTML, Javascript-Jquery, CSS, PHP)
- Exploratory Data Analysis and Visualization

10/2019 – 7/2021. University of Western Macedonia, Greece. Adjunct lecturer at the Department of Communication and Digital Media.

Courses:

- Data Technologies
- Database Management
- Educational Digital Tools

3/2017 – 3/2020. Hellenic Open University, Greece. Post-doctoral associate at the School of Science and Technology.

Research work:

- Design of similarity algorithms, data structures, and approximation schemes for privacy-preserving record linkage.
- Design and development of large-scale systems which rely on various randomization schemes (e.g., LSH, count-min sketches) for performing privacy-preserving record linkage using voluminous data.

3/2017 – present. I have supervised 20 post-graduate theses on various topics, including Big Data management and design of efficient algorithms.

RESEARCH PUBLICATIONS

1. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios," MultiBlock: A Scalable Iterative Approach for Progressive Entity Resolution", IEEE **Big Data**, 2021.
2. D. Vatsalan, R. Bhaskar, A. Gkoulalas-Divanis, D. Karapiperis, "Privacy Preserving Textual Data Encoding", **Big Data**, 2021.
3. A. Gkoulalas-Divanis, D. Vatsalan, D. Karapiperis, M. Kantarcioglu, "Modern Privacy-Preserving Record Linkage Techniques: An Overview", **TIFS**, 2021
4. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "Summarizing and linking electronic health records", **DPDB**, 2021.
5. D. Karapiperis, A. Gkoulalas-Divanis "An Unbiased Estimator for Hamming LSH Blocking", **ISC2**, 2021.
6. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios,"Efficient Record Linkage in Data Streams", IEEE **Big Data**, 2020.
7. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "Large-Scale Distributed Linkage of Records Containing Spatio-Temporal Information", **ISC2**, 2020.
8. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "Linkage of Spatio-Temporal Records and Trajectories", **ISC2**, 2019.
9. D. Vatsalan, D. Karapiperis, V.S. Verykios, "Privacy-Preserving Record Linkage", **Encyclopedia of Big Data Technologies**, 2019.
10. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "FEMRL: A Framework for Large-Scale Privacy-Preserving Linkage of Patients' Electronic Health Records", **ISC2**, 2018.
11. D Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "FEDERAL: a framework for distance-aware privacy-preserving record linkage", **TKDE**, 30 (2), p. 292-304, 2018.
12. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "Fast schemes for online record linkage", **DMKD**, 32(5), p. 1229 -1250, 2018.
13. D. Karapiperis, A. Gkoulalas-Divanis, V.S. Verykios, "Summarization Algorithms for Record Linkage", **EDBT**, p. 73-84, 2018.

14. T. Theodosiou, D. Karapiperis, V.S. Verykios, "Using Wavelets for Matching Records Privately", **PCI**, 2017.
15. D. Karapiperis, A. Gkoulalas-Divanis, and V. S. Verykios "Distance-Aware Encoding of Numerical Values for Privacy-Preserving Record Linkage", **ICDE,** p. 135- 138, 2017.
16. D. Karapiperis, A. Gkoulalas-Divanis, and V. S. Verykios "LSHDB: A Parallel and Distributed Engine for Record Linkage and Similarity Search", **ICDM Demo,** p. 1 - 4, 2016.
17. D. Karapiperis, V.S. Verykios **"**A Fast and Efficient Hamming LSH-Based Scheme for Accurate Linkage", **KAIS**, 49(3), p. 861-884, 2016.
18. D. Karapiperis, D. Vatsalan, V.S. Verykios, and P. Christen "Efficient Record Linkage using a Compact Hamming Space", **EDBT**, p. 209 - 220, 2016.
19. D. Karapiperis, V.S. Verykios "A Tutorial On Blocking Techniques for Privacy-Preserving Record Linkage", **AlgoCloud Workshop**, 2015.
20. D. Karapiperis, D. Vatsalan, V.S. Verykios, and P. Christen "Large-Scale Multi-Party Counting Set Intersection using a Space Efficient Global Synopsis", **DASFAA**, p. 329-345, 2015.
21. D. Karapiperis, V.S. Verykios "Load-Balancing the Distance Computations in Record Linkage"**, SIGKDD Explorations**, 17(1), p. 1-7, 2015.
22. D. Karapiperis and V. S. Verykios "An LSH-based Blocking Approach with a Homomorphic Matching Technique for Privacy-Preserving Record Linkage", **TKDE**, 27(4), p. 909-921, 2015**.**
23. D. Karapiperis and V. S. Verykios "A distributed near-optimal LSH-based framework for Privacy-Preserving Record Linkage", **COMSIS**, 11 (2), p. 745-763, 2014.
24. D. Karapiperis and V. S. Verykios "A distributed framework for scaling Up LSH-based computations in Privacy Preserving Record Linkage**", BCI**, p. 102-109, 2013.

Google scholar profile:
https://scholar.google.gr/citations?user=3KkH_YQAAAAJ&hl=en
DBLP profile:
https://dblp.org/pid/132/9541.html

PROGRAMMING SKILLS

- Java (including parallel programming using threads)
- Python
- Hadoop Map/Reduce
- PHP, Javascript, HTML
- Bash shell

A sample of my work is publicly available at Github:
LSHDB (https://github.com/dimkar121/LSHDB) is an open-source framework for large-scale record linkage and similarity search tasks using Locality-Sensitive Hashing and noSQL data stores. Parallelism lies at the core of its mechanism, since queries are executed in parallel using a pool of threads.

WORK EXPERIENCE

2/2006 – present. Local Government, National Education Services, IT department **(Senior Software Engineer and System Designer)**.

- Design and development of a Java application, which integrates and stores data from multiple data sources (e.g., Oracle, MySQL, SQL server, and text files) into a central repository, on which a REST API is built for cross-platform interoperability.
- Design and development of a Java application for linking text files, which consist of citizens' records, to identify duplicate records.
- Design and development of an authentication component using PHP that grants access to privileged users to certain web pages, using the Active Directory and an MS-SQL database as the credentials' repositories.
- Installation and parameterization of corporate mail services using Postfix, Dovecot and GNU Mailman on CentOS Linux. Configuration of Dovecot to authenticate POP3 users against the Active Directory (http://wiki2.dovecot.org/HowTo/ActiveDirectoryNtlm).
- Design and development of a Java application for importing financial data from the corporate database server (SQL Server), by applying a "group-by" policy, and then migrating the data to a MySQL server, which plays the role of a data warehouse for OLAP applications.
- Design and development of a PHP application which consumes web services offered by the Finance Ministry to Local Government agencies.
- Design and development of a PHP application using the REST API of the DIAVGEIA (TRANSPARENCY) system, which is a data repository containing the financial transactions conducted by the various departments of the Greek public sector.
- Design and development of web applications using PHP, MySQL hosted by Linux operating systems.

3/2003 – 12/2005. Metropolis Informatics SA **(Senior Software Engineer, Team-leading Software Developer, and System Designer)**.

Design and development of the company-wide software infrastructure, which established the core of the front-layer for all J2EE applications.
The projects, in which I had active role were:
- INMIS project (OTE Internet MIS), based on the J2EE framework.
- Scada / Telerise project based on the J2EE framework.

9/2001-12/2001. High Integrity Systems Engineering Research Group, University of York, UK **(Research Intern).**

- Development of Java components (servlets) converting VisioXML into GSML and vice versa, using Xerces.

1/1997-6/2000. University of Macedonia, Greece **(Research Assistant).**

- Development of web and database applications.

6/2007- present. National Centre of Public Administration **(Lecturer)**.

I prepare and give lectures in the following topics:
- "Client-side internet programming using Javascript and HTML".
- "Developing applications using the Java programming language".

- “Developing open-source applications using PHP and MySQL”.
- "The Linux shell".

REFERENCES

- Professor Vassilios S. Verykios, Hellenic Open University, Patras, Greece (verykios@eap.gr).
- Professor Peter Christen, Australian National University, Canberra, Australia (peter.christen@anu.edu.au).
- Dr. Aris Gkoulalas-Divanis, Senior Privacy Research Scientist, IBM Watson Health, Cambridge, MA, USA (gkoulala@us.ibm.com).