

## Vasilis Efthymiou, Ph.D.

Incoming Assistant Professor at Harokopio University of Athens, Greece

website: <http://sites.google.com/site/vefthym>

e-mail: vefthym AT ics DOT forth DOT gr

---

<b>Research interests</b>	Entity Resolution, Big Data, Knowledge Management and Representation, Semantic Web, Databases	
<b>Research and Work Experience</b>	<b>Assistant Professor at Harokopio University of Athens, Athens, Greece</b> Elected Assistant Professor at HUA on “Big Data Database Systems”	<i>start (exp): Jan. 2024</i>
	<b>Postdoctoral Researcher at FORTH-ICS, Heraklion, Crete, Greece</b> PI of the project <a href="#">ResponsibleER</a> on fair and explainable entity resolution. Information Systems Laboratory of FORTH-ICS. Funding: HFRI	<i>Sep. 2020 - Dec. 2023</i>
	<b>Visiting Instructor at University of Crete, Heraklion, Crete, Greece</b> Teaching the courses “Advanced Database Topics - Issues and Research Challenges in Big Data Management”, “Knowledge Representation and Reasoning”, and “Information Systems Analysis and Design”	<i>Sep. 2020 – July 2023</i>
	<b>Postdoctoral Researcher at IBM Research - Almaden, San Jose, California, USA</b> Research on natural language query processing over Knowledge Bases. Database group of IBM Almaden, Manager: Fatma Özcan	<i>Sep. 2018 - Aug. 2020</i>
	<b>Identity Matching Consultant at Trust Science, Edmonton, Alberta, Canada</b> Working as an external consultant in the field of Identity Matching.	<i>Nov. 2017 - July 2018</i>
	<b>Research Intern at IBM Research – T. J. Watson Research Center, NY, USA</b> Research internship on Web Table annotation using a Knowledge Base, working with Oktie Hassanzadeh and Mariano Rodriguez-Muro.	<i>June 2016 - Sep. 2016</i>
	<b>Visiting Researcher at RMIT, Melbourne, Victoria, Australia</b> A secondment of the SemData project, an International Research Staff Exchange Scheme (IRSES). Hosted by Timos Sellis.	<i>Feb. 2016</i>
<b>Research Assistant at FORTH-ICS, Heraklion, Crete, Greece</b> Work on several research projects throughout my BSc, MSc, and PhD, in the areas of Knowledge Management, Ambient Intelligence, Web of Data and Entity Resolution.	<i>Jun. 2008- July 2017</i>	
<b>Education</b>	<b>PhD in Computer Science</b> <i>Computer Science Department, University of Crete, Heraklion, Greece</i> Thesis: “Entity Resolution in the Web of Data”	<i>Mar. 2013- Sep. 2017</i>
	<b>MSc in Information Systems, and in Biomedical Informatics and Technology</b> <i>Computer Science Department, University of Crete, Heraklion, Greece</i> Thesis: “A real-time semantics-aware activity recognition system”	<i>Mar. 2010- Feb. 2012</i>
	<b>Bachelor in Computer Science</b> <i>Computer Science Department, University of Crete, Heraklion, Greece</i>	<i>Sep. 2005- Feb. 2010</i>
<b>Keynotes and Invited Talks</b>	<ul style="list-style-type: none"><li>• “Fairness in Algorithmic Decision Making”. <b>Invited Talk</b> at Responsible Data Management Course, taught by Renée Miller, Northeastern University, Boston, MA, USA, November 2023</li><li>• “Guided Data Analysis for Conversational Business Intelligence”. <b>Invited Talk</b> at <a href="#">RecSys Talks</a>, Tampere University, Finland, June 2022</li><li>• “Academia vs Industry Career Paths”. <b>Panel member</b>, 1<sup>st</sup> Data Management and Applications Workshop (DMAW) held at University of Athens and ATHENA Research Center, Athens, Greece, September 2021. Moderator: Georgia Koutrika. Co-panelists: Eirini Ntoutsis, Anna Kyriakidou, Nassos Katsamanis, Neoklis Polyzotis</li><li>• “Entity Resolution in the Web of Data”. <b>Invited Talk</b> at Tampere University, Tampere, Finland, February 2021</li><li>• “Ontology-based Natural Language Query Answering”. <b>Keynote Talk</b>, 2<sup>nd</sup> workshop on large-scale RDF analytics (LASCAR) at ESWC 2020, Heraklion, Greece, June 2020</li></ul>	

---

---

**Tutorials given**

- **State of the Art and Open Challenges in Natural Language Interfaces to Data**  
SIGMOD 2020. Portland, Oregon, USA  
Presenters: Fatma Özcan, Abdul Quamar, Jaydeep Sen, Chuan Lei, Vasilis Efstymiou
  - **Web-scale Blocking, Iterative and Progressive Entity Resolution**  
ICDE 2017. San Diego, California, USA  
Presenters: Kostas Stefanidis, Vassilis Christophides, Vasilis Efstymiou
  - **Improving Quality and Scalability in Semantic Data Management**  
ESWC 2016, Heraklion, Crete, Greece  
Presenters: Vassilis Christophides, Oscar Corcho, Vasilis Efstymiou, Vadim Ermolayev, Andreas Harth, Luis- Daniel Ibáñez, Natalya Keberle, Kostas Stefanidis, Ilias Tachmazidis
  - **Entity Resolution in the Web of Data**  
WWW 2014, Seoul, Korea  
Presenters: Kostas Stefanidis, Vasilis Efstymiou, Melanie Herschel, Vassilis Christophides
  - **Entity Resolution in the Web of Data**  
CIKM 2013, Burlingame, CA, USA  
Presenters: Kostas Stefanidis, Vasilis Efstymiou, Melanie Herschel, Vassilis Christophides
- 

**Publications & Patents****Books/Monographs**

- A. Quamar, V. Efstymiou, C. Lei, F. Özcan. Natural Language Interfaces to Data. Foundations and Trends in Databases 11(4): 319-414 (2022)
- V. Christophides, V. Efstymiou, K. Stefanidis. Entity Resolution in the Web of Data. Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan and Claypool (2015)

**Journal Articles**

- G. Papadakis, V. Efstymiou, E. Thanos, O. Hassanzadeh, P. Christen. An Analysis of One-to-One Matching Algorithms for Entity Resolution. VLDBJ (to appear)
  - N. Fanourakis, V. Efstymiou, D. Kotzinos, V. Christophides. Knowledge Graph Embedding Methods for Entity Alignment: An Experimental Review. Data Mining and Knowledge Discovery 37(5): 2070-2137 (2023)
  - C. Lei, A. Quamar, V. Efstymiou, F. Özcan, R. Alotaibi. HERMES: Data Placement and Schema Optimization for Enterprise Knowledge Bases. VLDBJ 32(3): 549-574 (2023)
  - V. Christophides, V. Efstymiou, T. Palpanas, G. Papadakis, K. Stefanidis. An Overview of End-to-End Entity Resolution for Big Data. ACM Computing Surveys 53(6): 127:1-127:42 (2021)
  - J. Sen, C. Lei, A. Quamar, F. Özcan, V. Efstymiou, A. Dalmia, G. Stager, A. Mittal, D. Saha, K. Sankaranarayanan. ATHENA++: Natural Language Querying for Complex Nested SQL Queries. Proceedings of the VLDB Endowment (PVLDB) 13(11): 2747-2759 (2020)
  - V. Efstymiou, K. Stefanidis, V. Christophides. Benchmarking Blocking Algorithms for Web Entities. IEEE Transactions on Big Data 6(2): 382-395 (2020)
  - V. Efstymiou, G. Papadakis, G. Papastefanatos, K. Stefanidis, T. Palpanas. Parallel Meta-blocking for Scaling Entity Resolution Over Big Heterogeneous Data. Information Systems 65: 137-157 (2017)
  - P. Caire, A. Moawad, V. Efstymiou, A. Bikakis, Y. Le Traon: Privacy Challenges in Ambient Intelligence Systems. Journal of Ambient Intelligence and Smart Environments (JAISE) 8(6): 619-644 (2016)
  - V. Efstymiou, M. Koutraki, G. Antoniou. Real-Time Activity Recognition and Assistance in Smart Classrooms. In the Advances in Distributed Computing and Artificial Intelligence Journal (ADCAIJ), 1(1): 9-22 (2012)
-

---

### Conference and Workshop papers

- V. Efhymiou, E. Ioannou, M. Karvounis, M. Koubarakis, J. Maciejewski, K. Nikoletos, G. Papadakis, D. Skoutas, Y. Velegarakis, A. Zeakis. Self-configured Entity Resolution with pyJedAI. IEEE Big Data 2023 (vision paper)
- N. Fanourakis, C. Kontousias, V. Efhymiou, V. Christophides, D. Plexousakis. FairER demo: Fairness-Aware and Explainable Entity Resolution. ISWC 2023 (demo)
- N. Fanourakis, V. Efhymiou, V. Christophides, D. Kotzinos, E. Pitoura, K. Stefanidis. Structural Bias in Knowledge Graphs for the Entity Alignment Task. ESWC 2023
- K. Papantoniou, V. Efhymiou, D. Plexousakis. Automating Benchmark Generation for Named Entity Recognition and Entity Linking. ESWC 2023 (poster)
- E. Chrysostomaki, M. Stratigi, V. Efhymiou, K. Stefanidis, D. Plexousakis. Fair Sequential Group Recommendations in SQUIRREL Movies. TaDA@VLDB 2023 (poster)
- E. Ymeralli, G. Flouris, V. Efhymiou, K. Papantoniou, T. Patkos, G. Petasis, N. Pittaras, G. Roussakis, E. Tzortzakakis. Representing Online Debates in the Context of E-Journalism. SEMAPRO 2022 [best paper award]
- V. Efhymiou. CosmOntology: Creating an Ontology of the Cosmos. DL4KG@ISWC 2022 (position paper)
- G. Papadakis, V. Efhymiou, E. Thanos, O. Hassanzadeh. Bipartite Graph Matching Algorithms for Clean-Clean Entity Resolution: An Empirical Evaluation. EDBT 2022
- V. Efhymiou, K. Stefanidis, E. Pitoura, V. Christophides. FairER: Entity Resolution with Fairness Constraints. CIKM 2021 (short)
- J. Hao, C. Lei, V. Efhymiou, A. Quamar, F. Özcan, Y. Sun, W. Wang. MEDTO: Medical Data to Ontology Matching using Hybrid Graph Neural Networks. KDD 2021
- K. Papantoniou, V. Efhymiou, G. Flouris. EL-NEL: Entity Linking for Greek News Articles. ISWC 2021 (poster) [best poster candidate]
- F. Özcan, C. Lei, A. Quamar, V. Efhymiou. Semantic Enrichment of Data for AI Applications. DEEM@SIGMOD 2021
- A. Vassiliades, T. Patkos, V. Efhymiou, A. Bikakis, N. Bassiliades D. Plexousakis. Object-Action Association Extraction from Knowledge Graphs. SEMANTICS 2021
- A. Vretinaris, C. Lei, V. Efhymiou, X. Qin, F. Özcan. Medical Entity Disambiguation using Graph Neural Networks. SIGMOD 2021
- S. Ahmetaj, V. Efhymiou, R. Fagin, P. Kolaitis, C. Lei, F. Özcan, L. Popa. Ontology-Enriched Query Answering on Relational Databases. IAAI 2021
- R. Alotaibi, C. Lei, A. Quamar, V. Efhymiou, F. Özcan. Property Graph Schema Optimization for Domain-Specific Knowledge Graphs. ICDE 2021
- X. Han, L. Hu, J. Sen, Y.B. Dang, B.Y. Gao, V. Isahagian, C. Lei, V. Efhymiou, F. Özcan, A. Quamar, Z. Huang, V. Muthusamy. Bootstrapping Natural Language Querying on Process Automation Data. IEEE SCC 2020
- A. Quamar, C. Lei, D. Miller, F. Özcan, J. Kreulen, R. Moore, V. Efhymiou. An Ontology-Based Conversation System for Knowledge Bases. SIGMOD 2020

- 
- E. Jiménez-Ruiz, O. Hassanzadeh, V. Efthymiou, J. Chen, K. Srinivas: SemTab 2019: Tabular Data to Knowledge Graph Matching Challenge. ESWC 2020
  - C. Lei, V. Efthymiou, R. Geis, F. Özcan: Expanding Query Answers on Medical Knowledge Bases. EDBT 2020
  - V. Efthymiou, G. Papadakis, K. Stefanidis, V. Christophides: MinoanER: Schema-Agnostic, Non-Iterative, Massively Parallel Resolution of Web Entities. EDBT 2019
  - V. Efthymiou, G. Papadakis, K. Stefanidis, V. Christophides: Simplifying Entity Resolution on Web Data with Schema-agnostic, Non-iterative Matching. ICDE 2018
  - V. Efthymiou, O. Hassanzadeh, M. Rodriguez-Muro, V. Christophides: Matching Web Tables with Knowledge Base Entities: From Entity Lookups to Entity Embeddings. ISWC 2017
  - V. Efthymiou, O. Hassanzadeh, M. Sadoghi, M. Rodriguez-Muro: Annotating Web tables through ontology matching. OM @ ISWC 2016
  - V. Efthymiou, P. Zervoudakis, K. Stefanidis, D. Plexousakis: Group Recommendations in MapReduce. IC-ININFO 2016
  - V. Efthymiou, K. Stefanidis, V. Christophides: Minoan ER: Progressive Entity Resolution in the Web of Data. EDBT 2016 (poster)
  - V. Efthymiou, K. Stefanidis, V. Christophides: Big Data Entity Resolution: From Highly to Somehow Similar Entity Descriptions in the Web. IEEE Big Data 2015
  - V. Efthymiou, G. Papadakis, G. Papastefanatos, K. Stefanidis, T. Palpanas: Parallel Meta-blocking: Realizing Scalable Entity Resolution Over Large, Heterogeneous Data. IEEE Big Data 2015
  - V. Efthymiou, K. Stefanidis, E. Ntoutsis: Top-k Computations in MapReduce: A Case Study on Recommendations. IEEE Big Data 2015 (poster)
  - V. Efthymiou, P. Caire, A. Bikakis: Modeling and Evaluating Cooperation in Multi-Context Systems using Conviviality. BNAIC 2012
  - A. Bikakis, V. Efthymiou, P. Caire, Y. Le Traon.: Introducing Conviviality as a property of Multi-Context Systems. ARCOE 2012
  - A. Moawad, V. Efthymiou, P. Caire, G. Nain, Y. Le Traon: Introducing Conviviality as a New Paradigm for Interactions among IT Objects. AI @ IE 2012
  - P. Caire, A. Bikakis, V. Efthymiou: Conviviality by Design. Social Turn – SNAMAS 2012
  - M. Koutraki, V. Efthymiou, G. Antoniou: S-CRETA: Smart Classroom Real-Time Assistance ISAmI 2012
- Patents Filed**
- A. Quamar, F. Özcan, C. Lei, V. Efthymiou, V. Meduri: Guided Exploration for Conversational Business Intelligence. US 17/188405 (filed on March 1, 2021)
  - C. Lei, J. Hao, V. Efthymiou, F. Özcan, A. Quamar: Hybrid Graph Neural Network. US 17/161093, 17/161152 (filed on January 28, 2021)
  - C. Lei, A. Vretinaris, V. Efthymiou, F. Özcan: Entity Disambiguation Using Graph Neural Networks. US 17/130107 (filed on December 22, 2020)
  - C. Lei, A. Quamar, V. Efthymiou, F. Özcan, R. Alotaibi: Reducing Response Time for Queries Directed to Domain-Specific Knowledge Graph Using Property Graph Schema Optimization. US
-

---

16/800392 (filed on February 25, 2020)

### Technical Reports

- V. Efstathiou, P. Caire: Diagram Analysis Report: Use Cases for Conviviality and Privacy in Ambient Intelligent Systems. ISBN: 978-2-87971-114-0

---

### Awards & Grants

- Teaching grants, awarded at University of Crete, Greece (2021-2022 and 2022-2023)
- Post-doctoral grant from HFRI to lead 3-year project ResponsibleER as PI (2021-2023)
- IBM Manager's Choice Award. Awarded by Fatma Özcan on Nov 26, 2019
- 3<sup>rd</sup> prize at 2019 IBM Almaden AI Hackathon
- IBM Manager's Choice Award. Awarded by Fatma Özcan on Apr 1, 2019
- ISWC 2017 Student Travel Award
- IEEE Big Data 2015 Student Travel Award
- European Data Forum 2015 Travel Grant
- PhD scholarship provided by FORTH (2013-2017)
- MSc scholarship provided by FORTH (2010-2012)
- Undergraduate scholarship provided by FORTH (2008-2010)

---

### Projects

#### Principal Investigator

- ResponsibleER - Responsible by Design Entity Resolution (<https://isl.ics.forth.gr/ResponsibleER>)  
Funding: Hellenic Foundation for Research and Innovation (HFRI), Duration: 1/2021-12/2023

#### Project Member

- DebateLab - From Linked Data to Linked Arguments (<https://debatelab.ics.forth.gr/>)
- SemData - Semantic Data Management (<http://www.semdata-project.eu/>)
- PARTHENOS - Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (<http://www.parthenos-project.eu>)
- IdeaGarden - An Interactive Learning Environment Fostering Creativity (<http://idea-garden.org/>)
- CoPAInS - Conviviality and Privacy in Ambient Intelligent Systems (<http://www.fnr.lu/en/content/view/full/7528>)
- Research and Promotion of Renaissance Cultural heritage in Crete and Cyprus (<http://www.medieval-etia-pafos.gr>)

---

### Service & Membership

#### Scientific Communities

Association for Computing Machinery (ACM), Professional Member

#### Organizing Committee

- Tabular Data Analysis (TaDA) Workshop @ VLDB 2023. Co-Chair; main contact person
- VLDB 2023 PhD Workshop Co-Chair
- ESWC 2023 Posters & Demos Track Co-Chair
- SemTab @ ISWC 2019-2023: Semantic Web Challenge on Tabular Data to Kn. Graph Matching

#### Guest Editor

- Journal of Web Semantics, Special Issue on Automating Knowledge Graph Construction (2021).

#### Senior Program Committee Member

- ESWC 2024 (Resources Track)
- ESWC 2023 (Research Track)

#### Program Committee Member

- EDBT 2024 (Demo Track), 2023 (Demo Track)
- IJCAI 2023
- SIGMOD 2023 (Demo Track), 2019 (Demo Track)
- WWW 2023, 2022 [**honorable mention as PC member**], 2021
- ICDE 2023 (Industry & Applications Track), 2020
- ISWC 2023 (Posters & Demos), 2022 (Posters & Demos)
- DL4KG @ ISWC 2023 (Workshop on Deep Learning for Knowledge Graphs)
- NeSy @ IJCLR 2023, 2022, 2021 (Workshop on Neural-Symbolic Learning and Reasoning)
- KnOD @ WWW 2023, 2022, 2021 (Workshop on Knowledge Graphs for Online Discourse Analysis)
- CIKM 2022
- SEMANTiCS 2022
- TRL @ NeurIPS 2022 (Table Representation Learning Workshop)
- DOLAP @ EDBT 2022, 2021 (Workshop on Design, Optimization, Languages and Analytical Processing of Big Data)

- 
- VLDB 2021 (Demo Track)
  - KiML @ KDD 2020 (Workshop on Knowledge-infused Mining and Learning)
  - LASCAR @ ESWC 2020 (Workshop on Large Scale RDF Analytics)
  - IEEE Big Data 2020, 2019 (Industry & Government Track)

#### Journal Referee

- Neurocomputing
- ACM Journal of Data and Information Quality (JDIQ) – SI: Deep Learning for Data Quality
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- Information Systems
- Semantic Web Journal
- ACM Journal of Data and Information Quality (JDIQ)
- Knowledge and Information Systems (KAIS)
- Big Data Research
- Sustainability
- International Journal of Cooperative Information Systems (IJCIS)
- Journal of Survey Statistics and Methodology (JSSAM)

#### External Reviewer

SIGMOD 2023, EDBT (2021, 2019, 2016), WISE (2021, 2020), KEOD (2021), TPDJ (2021), WWW (2020, 2019), VLDB (demo track 2022, 2020, demo track 2016), CIKM (2020, 2019, 2015), ICDE (2019, 2018), ACM Computing Surveys, TIME 2018, ICDM 2015, DSAA 2015, VLDBJ

#### Evaluation Committee

- ISWC 2022 jury member for best poster and demo
- Award evaluation committee at Data Management and Applications Workshop (DMAW), organized by the National & Kapodistrian University of Athens (NKUA), 2021.

#### Book summary (slide presentations)

“A Semantic Web Primer” 2<sup>nd</sup> edition. G. Antoniou, F. van Harmelen

- Available at <http://www.ics.forth.gr/isl/swprimer/>.

#### Conference volunteer

- 3<sup>rd</sup> IEEE International Conference on Big Data, Santa Clara, CA, USA. Oct 29 – Nov 1, 2015
- 8<sup>th</sup> Extended Semantic Web Conference (ESWC), Heraklion, Greece. May 19 - June 2, 2011

---

## Teaching

#### Visiting Instructor (at University of Crete, Greece) [2020 – 2023]

- Information Systems Analysis and Design. Spring 2022, Spring 2023
- Knowledge Representation and Reasoning. Spring 2021, Fall 2021, Fall 2022
- Advanced Topics in Databases – Big Data Processing & Analytics. Fall 2020

#### Teaching Assistant (at University of Crete, Greece) [2010 – 2017]

- Information Retrieval Systems – Spring 2017
- Advanced Topics in Databases – Issues and Research Challenges in Big Data Management. Spring 2016, Fall 2016
- Advanced Topics in Databases – Big Data Processing & Analytics. Spring 2015
- Computer Organization. Spring 2014
- Object-oriented programming. Fall 2013, Fall 2014, Fall 2015
- Theory of computation. Fall 2011
- Data structures. Fall 2010
- C Programming. Spring 2010, Spring 2011

---

## Student Supervision / Mentoring

- Nikolaos Fanourakis, graduate student at UoC, Spring 2021 (MSc student) – today (PhD student)
  - Sotiris Koumourou, graduate student at UoC, Fall 2023
  - Tarso Keletzi, undergraduate thesis at UoC, Fall 2023
  - Eva Chrysostomaki, undergraduate thesis at UoC, Fall 2022
  - Christos Kontousias, undergraduate thesis at UoC, Spring 2022
  - Neofytos Dionysiou, undergraduate thesis at UoC, Fall 2021
  - Venkata Vamsikrishna Meduri, graduate intern at IBM Research-Almaden, Summer 2020 (partial)
  - Alina Vretinaris, undergraduate intern at IBM Research-Almaden, Summer 2019
  - Dimitris Sbonias, undergraduate thesis at UoC, Fall 2013
  - Giorgos Varvates, undergraduate thesis at UoC, Spring 2013
-

---

**Technical Skills**

**Preferred Programming Languages:** Python, Java, Scala

**Knowledge repr. & querying:** RDF, OWL, XML, JSON, SQL, SPARQL, Cypher, Gremlin

**Frameworks:** Spark, Hadoop, Elasticsearch, JanusGraph, Neo4j, Protégé, IBM Watson Assistant

Some of the open software produced: <https://github.com/vefthym>

---

**Languages**

**Greek** (native), **English** (C2), **German** (B1 - a long time ago...)

---