

Manolis (Manos) Karpathiotakis

Ph.D Candidate in Database Systems - École Polytechnique Fédérale de Lausanne

manos.karpathiotakis@epfl.ch

<http://karpathiotakis.net>

INTERESTS

data management in the presence of data variety, query processing, just-in-time code generation, (scale-out) data analytics, hardware accelerators

PROFESSIONAL EXPERIENCE

09/2012 – present **École Polytechnique Fédérale de Lausanne** **Lausanne, Switzerland**

- Doctoral Assistant & Member of the Data-Intensive Applications and Systems (DIAS) laboratory, advised by Prof. Anastasia Ailamaki.
- Working on analytics over heterogeneous data and hardware (<http://dias.epfl.ch/vida>).

06/2015 – 09/2015 **IBM Almaden Research Center** **San Jose, California**

- Intern at the Big Data Research Group, mentored by Avrilia Floratou and Fatma Özcan.
- Worked on enhancing the federated querying capabilities of Spark, a big data engine.

10/2013 – 06/2015 **École Polytechnique Fédérale de Lausanne** **Lausanne, Switzerland**

- Research Assistant in the **Human Brain Project** (<https://www.humanbrainproject.eu>).
- Worked on the query engine of a platform used for the diagnosis of brain diseases.

06/2009 – 09/2012 **University of Athens** **Athens, Greece**

- Implementation team coordinator, Research Assistant, Scientific Programmer in the context of the European FP7 Programs **SensorGrid4Env** (<http://semsorgrid4env.eu>) and **TELEIOS** (<http://www.earthobservatory.eu>).
- Worked on a scalable spatiotemporal RDF store (<http://www.strabon.di.uoa.gr>).

HONORS & AWARDS

- EPFL Teaching Assistant Award, 2016
- IBM PhD Fellowship Award, 2015-2016
- EPFL Computer Science Fellowship, 2012-2013
- Awarded 3rd place in Semantic Web Challenge 2012

EDUCATION

• 09/2012 – present: **Doctoral Student** **École Polytechnique Fédérale de Lausanne**
Thesis: Just-in-time Analytics Over Heterogeneous Data and Hardware

• April 2011: **M.Sc. in Advanced Information Systems** **University of Athens, Greece**
Grade: 9.34 / 10
Thesis: Design and Implementation of a Registry for the Semantic Sensor Web

• June 2008: **B.Sc. in Informatics & Telecommunications** **University of Athens, Greece**
Grade: 8.35 / 10 - Graduation Rank: 3rd / 52
Thesis: e-Government with the utilization of GIS technologies

TECHNICAL SKILLS

Programming & Scripting: C, C++, Intel Thread Building Blocks (TBB) C++ libraries, CUDA, (LLVM) Assembly, Java, Scala, Go, Bash Shell, Python

SQL & Database Systems: PostgreSQL, MySQL, Oracle Database, MS SQL Server, IBM DB2, HP Vertica, MonetDB, SciDB

NoSQL: MongoDB, Apache Cassandra, Redis

Big Data: Apache Spark, Apache Hive, IBM Big SQL, Apache Hadoop (MapReduce)
Data Representations: JSON, XML, Parquet, RDF, OWL, GML, KML
Spatial Data Management: PostGIS, ArcGIS, QGIS, GeoServer, OpenLayers
Debugging & Performance Analysis: gdb, perf, Intel VTune, NVIDIA CUDA Profiler (nvprof)
Testing: Google Test, JUnit, ScalaTest, Jenkins
Software Development IDEs: Eclipse, NetBeans, IntelliJ IDEA, Scala IDE, NVIDIA Nsight, Microsoft Visual Studio, Protégé, Altera Modelsim & Quartus

NOTABLE PROJECTS

- A query engine for binary, CSV, and JSON data, which generates its LLVM C++ implementation just-in-time (<http://dias.epfl.ch/ViDa>).
- A Scala-based optimizer extension for Apache Spark SQL.
- Strabon, a Java-based semantic geospatial DBMS (<http://www.strabon.di.uoa.gr/>).
- Multiple database operators in CUDA C++ for GPU execution.
- An adaptor for PostgreSQL, allowing it to retrieve data from a distributed K-V store.
- Implementation of a concurrent B+ Tree index in Java.
- Simulation of a distributed auction system in Java.
- Implementation and comparison of SVD-based and clustering-based recommendation algorithms for the Netflix prize challenge using Apache Hadoop.

SELECTED PUBLICATIONS

Analytics over Heterogeneous Data and Hardware:

- M. Karpathiotakis, A. Floratou, F. Ozcan, A. Ailamaki, “**No Data Left Behind: Real-Time Insights from a Complex Data Ecosystem**”. In *SoCC*, 2017
- R. Appuswamy, M. Karpathiotakis, D. Porobic, A. Ailamaki, “**The Case For Heterogeneous HTAP**”. In *CIDR*, 2017
- M. Karpathiotakis, I. Alagiannis, A. Ailamaki, “**Fast Queries Over Heterogeneous Data Through Engine Customization**”. In *PVLDB*, Vol.9(12), 2016
- A. Dziedzic, M. Karpathiotakis, I. Alagiannis, R. Appuswamy, A. Ailamaki, “**DBMS Data Loading: An Analysis on Modern Hardware**”. In *ADMS*, 2016
- M. Karpathiotakis, I. Alagiannis, T. Heinis, M. Branco, A. Ailamaki, “**Just-In-Time Data Virtualization: Lightweight Data Management with ViDa**”. In *CIDR*, 2015
- M. Karpathiotakis, M. Branco, I. Alagiannis, A. Ailamaki, “**Adaptive Query Processing on RAW Data**”. In *PVLDB*, Vol.7(12), 2014

Spatiotemporal Data Management:

- K. Kyzirakos, M. Karpathiotakis, M. Koubarakis, “**Strabon, a Semantic Geospatial DBMS**”. In *ISWC*, 2012
- K. Kyzirakos, M. Karpathiotakis, G. Garbis, C. Nikolaou, K. Bereta, I. Papoutsis, T. Herekakis, D. Michail, M. Koubarakis, C. Kontoes, “**Wildfire monitoring using satellite images, ontologies and linked geospatial data**”. In *Journal of Web Semantics* 24
- M. Koubarakis, M. Karpathiotakis, K. Kyzirakos, C. Nikolaou, M. Sioutis, “**Data Models and Query Languages for Linked Geospatial Data**”. In *RW*, 2012

LANGUAGES

Greek (Native), English (Fluent), French (Beginner)