Giuseppe M. Mazzeo

Last update on April 3, 2016

gmmazzeo@gmail.com • +(310)480-1268 • www.linkedin.com/in/gmmazzeo
3241 S. Sepulveda Blvd #209 • Los Angeles, CA 90034

Summary

I am a computer scientist/software engineer with experience in both research and industry. My general research areas are databases and artificial intelligence, and my recent research topics are question answering over RDF data and clustering algorithms. I have past experience with the design and development of enterprise information systems, mainly based on Java technology and open source DBMSs. I like working on challenging and practically relevant problems, and I am particularly good at quickly turning ideas into prototypes.

Experience

University of California, Los Angeles — Los Angeles, USA

Postdoctoral Scholar — *Jan '14 – present*

The research activity, conducted under the supervision of Prof. Carlo Zaniolo, focuses on:

- development of a *controlled natural language* question answering system for querying RDF knowledge bases, that improves state-of-the-art CNL systems by providing a better accuracy ($> 90\%$), scalability and usability. A demo is available at yellowstone.cs.ucla.edu/canali
- development of a *parameter-free* algorithm for *clustering* data around-centroids, that outperforms state-of-the art algorithms by a wide margin. A parallel version of the algorithm, based on Apache Spark was also developed. More information is available from yellowstone.cs.ucla.edu/clubs

University of Milano-Bicocca — Milano, Italy

Research associate — *May '12 – Dec '13*

Devised techniques for modeling lawsuits and predicting their outcome (task part of electronic Justice Relationship Management project).

Open Knowledge Technologies Srl — Rende, Italy

CTO — *Nov '10 – Apr '12*

Co-founded this small startup and served as CTO. Main led projects:

- design and development of www.condomani.it, currently the most used SaaS for property management in Italy
- design and implementation of an OLAP system for analyzing the logs of the TETRA network, used by Italian Police (task part of TETRis – Innovative Services over TETRA project).

Zenit Srl — Vibo Valentia, Italy

Software Engineer — *Nov '08 – Apr '12*

Co-founded this small startup and worked as software engineer. Main led projects:

- management software for Pubbliemme Srl, one of the main Italian outdoor advertising companies. Main benefits delivered by the software I developed:
  - improved the overall workflows across functional business units
  - reduced from one day to few seconds the time needed to create the orders for billposters (about 5k billpostings every 2 weeks)
- management software for Banco Metalli Italiano Spa, the main Italian precious medal trading company. Main benefits delivered by the software I developed:
  - quick and error-free generation of the invoices, based on a complex mechanism depending on the continuously changing quotation of precious metals
  - instant overview of the incoming and outgoing goods

National Research Council — Rende, Italy

Research Associate — *Jan '08 – Feb '10*

Continued my research in the field of multi-dimensional data compression, achieving the following main results:

- a framework for building privacy-preserving histograms on multi-dimensional data
- a P2P-based framework supporting the extraction of aggregates from multi-dimensional data

University of Calabria — Rende, Italy

Contract Professor — *Jan '07 – Sep '10*

Instructor for the upper-division course of *Information Systems* (around 100 students per academic year) at the department of CS, where I also served as first-level thesis advisor for 30+ students.

Education

University of Calabria — Rende, Italy

Ph.D in Computer Science — *Feb 2007*

Devised algorithms for efficient and accurate summarization of multi-dimensional data. The algorithms were also adapted for the context of data streams.

Laurea degree (BS+MS) with honors in Computer Engineering — *May 2003*

Selected publications

- G. M. Mazzeo, C. Zaniolo: The Parallelization of a Complex Hierarchical Clustering Algorithm: faster unsupervised learning for larger data sets. UCLA CSD Technical Report No. 160001
- G. M. Mazzeo, C. Zaniolo: Answering Controlled Natural Language Questions on RDF Knowledge Bases. EDBT 2016 demo track
- E. Masciari, G. M. Mazzeo, C. Zaniolo: A New, Fast and Accurate Algorithm for Hierarchical Clustering on Euclidean Distances. PAKDD 2013
- F. Buccafurri, F. Furfaro, G. M. Mazzeo, D. Saccà: A quad-tree based multiresolution approach for two-dimensional summary data. Inf. Syst. 36(7) 2011
- F. Furfaro, G. M. Mazzeo, A. Pugliese: Managing Multidimensional Historical Aggregate Data in Unstructured P2P Networks. TKDE 22(9) 2010
- F. Furfaro, G. M. Mazzeo, D. Saccà: A Probabilistic Framework for Building Privacy-Preserving Synopses of Multi-dimensional Data. SSDBM 2008
- F. Furfaro, G. M. Mazzeo, D. Saccà, C. Sirangelo: Compressed hierarchical binary histograms for summarizing multi-dimensional data. KAIS 15(3) 2008
- F. Furfaro, G. M. Mazzeo, C. Sirangelo: Exploiting Cluster Analysis for Constructing Multi-dimensional Histograms on Both Static and Evolving Data. EDBT 2006

Skills

Technical expertise: Algorithm design, database design, software design and implementation. The languages/technologies/frameworks that I have been using through the years include: Java, JavaScript, jQuery, R, Scala, PHP, SQL, SPARQL, JavaEE, Struts2, HTML, CSS, XML, RDF, REST, JSON-RPC, Lucene, Spark, MySQL, PostgresSQL, Virtuoso, Cassandra, Mondrian.

Natural languages: Italian *(native speaker)*, English *(full professional proficiency)*, Spanish *(beginner)*.