



Europass Curriculum Vitae

Personal information

First name(s) / Surname(s) **Gianmarco De Francisci Morales**
Telephone(s) (+39) 0933 823163 Mobile | (+39) 328 0137307
E-mail(s) gianmarco.dfm@gmail.com
Nationality Italian
Date of birth 07/07/1983

Work experience

Dates Aug 11 →
Occupation or position held Apache Pig Committer
Main activities and responsibilities Committer for the Apache Pig project from the Hadoop ecosystem. <http://pig.apache.org>
Member of the open source community that develops Pig, a framework for large scale data analysis (Java).
Name and address of employer Apache Software Foundation

Dates Nov 09 →
Occupation or position held Research collaborator
Main activities and responsibilities Research on cloud computing for large scale data analysis.
Participation to international conferences to present research results.
Name and address of employer ISTI-CNR (National Research Council)
Pisa (Italy)

Dates Oct 10 - Sep 11
Occupation or position held Research intern
Main activities and responsibilities Research on Data Intensive Scalable Computing.
- Social Content Matching in MapReduce.
- From Chatter to Headlines: Harnessing the Real-Time Web for Personalized News Recommendation.
- Scaling out influence estimation on social networks.
- SafeKeeper: fault tolerance for parallel stream processing systems (Yahoo! S4).
Name and address of employer Yahoo! Research
Barcelona (Spain)

Dates May 11 - Aug 11
Occupation or position held Google Summer of Code student
Main activities and responsibilities Contribution to Pig, an open source platform for analyzing large data sets.
Pig is a Hadoop ecosystem project from the Apache Software Foundation. <http://pig.apache.org>
Sugar for Pig: syntactic sugar features to simplify the use of Pig. (Java, ANTLR)
<https://issues.apache.org/jira/browse/PIG-1926>
<https://issues.apache.org/jira/browse/PIG-1904>
Name and address of employer Google / Apache Software Foundation

Dates May 10 - Aug 10
Occupation or position held Google Summer of Code student
Main activities and responsibilities Contribution to Pig, an open source platform for analyzing large data sets.

	<p>Pig is a Hadoop ecosystem project from the Apache Software Foundation. http://pig.apache.org Design and implementation of a binary comparator for secondary sort. (Java) The new comparator is 10x faster than the original one, and produced a 20% end-to-end performance increase for relevant queries. https://issues.apache.org/jira/browse/PIG-1295</p>
Name and address of employer	Google / Apache Software Foundation
Dates	May 07 - Feb 09
Occupation or position held	System administrator
Main activities and responsibilities	Administration of Linux clusters for High Performance Computing (HPC). Management of Grid middleware (gLite). Porting, integration and development of Grid-enabled and parallel (MPI) applications.
Name and address of employer	Consorzio COMETA Catania (Italy)
Dates	Apr 04 - Jul 04
Occupation or position held	Programmer intern
Main activities and responsibilities	Development of a Java framework for booking and allocation of resources in a Grid environment.
Name and address of employer	INFN (National Institute for Nuclear Physics) Catania (Italy)
Dates	Jun 02 - Sep 02
Occupation or position held	Personal teacher
Main activities and responsibilities	One to one teaching of mathematics, algebra and geometry.

Education and training

Dates	Mar 09 →
Title of qualification awarded	Computer Science and Engineering PhD Student
Principal subjects / occupational skills covered	First in admission ranking, 3 years scholarship sponsored by the European Union. Thesis topic: Cloud computing for large scale data analysis. Courses: Complex networks. Fundamentals of computer science and engineering. Formal methods for concurrent systems. Advanced networking architectures. Game theory for computer science. Data integration and semistructured databases. Web algorithmics. Techniques for high availability in distributed systems. Anonymity and trust. Data fusion. Developed projects: - Formal verification of the 2PC and 3PC distributed consensus protocols. - Simulator of the deficit round robin scheduling algorithm. (Python)
Name and type of organisation providing education and training	IMT Institute for Advanced Studies Lucca (Italy)
Level in national or international classification	ISCED 6
Dates	3 Jul 10 - 10 Jul 10
Title of qualification awarded	Lipari School on Computational Complex Systems attendance certificate

Principal subjects / occupational skills covered	Social Networks. Graph mining. Social network analysis. Dynamical processes on complex networks. Rise of cooperation in society with greed.
Name and type of organisation providing education and training	Jacob T. Schwartz International School for Scientific Research Lipari (Italy)
Dates	6 Jul 09 - 10 Jul 09
Title of qualification awarded	CASE Summer School attendance certificate
Principal subjects / occupational skills covered	Agile software process management. Test Driven Development (TDD). Continuous testing of configurable systems. Software and process measurement.
Name and type of organisation providing education and training	Free University of Bozen Bolzano (Italy)
Dates	9 Dec 08 - 18 Dec 08
Title of qualification awarded	Matlab course attendance certificate
Principal subjects / occupational skills covered	Matlab: Fundamentals. GUI design. Integration with external languages. Distributed and parallel computing. (ML01 - ML04 - ML05 - DC01)
Name and type of organisation providing education and training	The MathWorks Palermo (Italy)
Level in national or international classification	
Dates	Oct 04 - Apr 08
Title of qualification awarded	Master's degree in Computer Engineering, with full marks and honors (110/110 cum laude)
Principal subjects / occupational skills covered	Scholarship for 1 year from ERSU (Regional Agency for University Studies). Additional award for outstanding academic career from University of Catania. Developed projects: - Module for user authentication using digital certificates on the J2EE platform. (Java, Tomcat) - Linux kernel module (for versions 2.4 and 2.6) for dynamic smoothing of IP packets by a fuzzy controller in a soft real-time environment. (C) - Hardware description of a 2 levels cache memory hierarchy. (VHDL) - Hardware description of a communication channel (bus) between processor and memory. (C++, SystemC) - Compiler from "Fuzzy Language" to Java bytecode. (Java) - Image processing web service (2 step filter). (Java, Glassfish) - Web Service enabled three-level information system for university exams management. (C#, MS SQL Server) - Industrial CAN fieldbus simulator. (C++, Omnet) - Distributed system for on-line catalogue management. (Java, RMI, AXIS, MySQL)
Name and type of organisation providing education and training	Thesis: "Multi-objective design space exploration of embedded systems in a Grid environment" (developed in C++, written in English) University of Catania, College of Engineering Catania (Italy)
Level in national or international classification	ISCED 5
Dates	Mar 06 - Aug 06
Title of qualification awarded	Erasmus student
Principal subjects / occupational skills covered	Artificial Intelligence. Design Patterns in Java. CNN (Cellular Neural Networks). Erasmus scholarship from ERSU (Regional Agency for University Studies). Socrates scholarship from Univeristy of Catania.
Name and type of organisation	Pazmany Peter Catholic University

providing education and training	Budapest (Hungary)
Dates	Nov 04 - Nov 04
Title of qualification awarded	PET Certificate (ESOL)
Principal subjects / occupational skills covered	Written and spoken English. Passed with merit.
Name and type of organisation providing education and training	University of Cambridge
Level in national or international classification	European level B1
Dates	Oct 01 - Oct 04
Title of qualification awarded	Bachelor's degree in Computer Engineering, with full marks and honors (110/110 cum laude)
Principal subjects / occupational skills covered	Scholarship for 2 years from ERSU (Regional Agency for University Studies). Additional scholarship award for achieving full marks and honors in minimum time from ERSU. Developed projects: - Residence management tool. (MS Access) - Digital modulation tool. (Matlab) - Routing algorithm simulator. (Java) - Distributed newsreader (client-server). (Java, RMI, XML) - Spell checker. (C++) - E-Commerce Web site. (PHP, MySQL)
Name and type of organisation providing education and training	Thesis: "Servizi di allocazione e monitoraggio dei job e delle risorse di una Grid" (Allocation and monitoring services for jobs and resources in a Grid environment, developed in Java) University of Catania, College of Engineering Catania (Italy)
Level in national or international classification	ISCED 4
Dates	Sep 96 - Jul 01
Title of qualification awarded	High school degree focusing on sciences, with marks of 100/100
Principal subjects / occupational skills covered	Mathematics, Physics, Italian, Latin, Philosophy, History, Biology, Astronomy, Chemistry, English.
Name and type of organisation providing education and training	Scientific Lyceum "E. Vittorini" Gela (Italy)
Level in national or international classification	ISCED 3

Personal skills and competences

Mother tongue(s) **Italian**

Other language(s)

Self-assessment
European level ()*

English
Spanish / Castilian
Japanese

Understanding				Speaking				Writing	
Listening		Reading		Spoken interaction		Spoken production			
C2	Proficient user	C2	Proficient user	C2	Proficient user	C2	Proficient user	C2	Proficient user
C2	Proficient user	C2	Proficient user	C2	Proficient user	C2	Proficient user	C2	Proficient user
A2	Basic User	A2	Basic User	A2	Basic User	A2	Basic User	A2	Basic User

(*) [Common European Framework of Reference \(CEF\) level](http://www.cedefop.europa.eu)

Social skills and competences	<p>I can easily adapt to different cultural environments. I honed this skill thanks to various trips abroad to study English language, five months spent in Hungary with the Erasmus exchange program, and one year in Spain during my PhD studies.</p> <p>I can effectively cooperate with colleagues in a team. I have successfully worked in a team during my bachelor's internship and while working as a system administrator. I also led different teams for the development of projects for academic courses.</p>
Organisational skills and competences	<p>Effective time management capability: the short exam periods have taught me to focus on one target at a time to be able to get the most from my abilities.</p> <p>I went through the university with ease while working and cultivating personal interests thanks to self-motivation. I learned how to effectively manage my priorities and adapt myself to the pace needed. I also learned to manage at best my spare time and to appreciate traveling.</p>
Computer skills and competences	<p>Known languages: Java, C/C++, Python, Bash, C#, PHP, SQL, XML, LaTeX.</p> <p>Cloud computing technologies (Hadoop ecosystem, Pig, Giraph, Zookeeper).</p> <p>Good knowledge of Matlab (scientific computing tool).</p> <p>Good knowledge of ANTLR (parser generator).</p> <p>Version control systems (subversion, git) and build systems (ant+ivy, maven).</p> <p>Optimal knowledge of the most widespread operating systems: Linux, Mac OS X, Windows XP.</p> <p>Databases: MySQL, MS SQL Server.</p> <p>Cluster and computing farm management. Virtualization technologies (VMWare, KVM).</p> <p>Network management (wired and wireless), security (firewall, cryptography, VPN). Design and administration of dynamic web sites.</p>
Artistic skills and competences	Digital photography.
Other skills and competences	Hobbies: cooking, reading, chess, archery, basket, role playing.
Additional information	<p>Publications:</p> <p>G. De Francisci Morales, A. Gionis, C. Lucchese “From Chatter to Headlines: Harnessing the Real-Time Web for Personalized News Recommendation” WSDM’12, 5th ACM International Conference on Web Search and Data Mining, Seattle, 2012, accepted</p> <p>G. De Francisci Morales, A. Gionis, M. Sozio “Social Content Matching in MapReduce” PVLDB, Proceedings of the VLDB Endowment, 4(7):460-469, 2011.</p> <p>R. Baraglia, G. De Francisci Morales, C. Lucchese “Document Similarity Self-Join with MapReduce” ICDM’10, 10th IEEE International Conference on Data Mining, Sydney, 2010, pp. 731-736.</p> <p>G. De Francisci Morales, C. Lucchese, R. Baraglia “Scaling Out All Pairs Similarity Search with MapReduce” LSDS-IR’10, 8th Workshop on Large-Scale Distributed Systems for Information Retrieval, @SIGIR’10, Geneva, 2010, pp. 25-30.</p> <p>G. De Francisci Morales, C. Lucchese, R. Baraglia “Large-scale Data Analysis on the Cloud” XXIV Convegno Annuale del CMG-Italia, Rome, 2010 (Best Paper Award).</p>

M. Iacono-Manno, M. Fargetta, R. Barbera, A. Falzone, G. Andronico, S. Monforte, A. Muoio, R. Bruno, P. Di Primo, S. Orlando, E. Leggio, A. Lombardo, G. Passaro, G. De Francisci Morales, S. Blandino

"The Sicilian Grid Infrastructure for High Performance Computing"

IJDST, International Journal of Distributed Systems and Technologies, 1(1):40-54, 2010.

V. Catania, G. De Francisci Morales, A. Di Nuovo, M. Palesi, D. Patti

"An Effective Methodology to Multi-objective Design of Application Domain-specific Embedded Architectures"

DSD'09, Digital System Design Architectures, Methods and Tools, Patras, 2009, pp. 643-650.

V. Catania, G. De Francisci Morales, A. Di Nuovo, M. Palesi, D. Patti

"Multi-objective Design Space Exploration of Embedded Systems in a Grid Environment"

Workshop finale dei Progetti Grid del PON Ricerca 2000-2006, Catania, 2009, pp. 46.

V. Catania, G. De Francisci Morales, A. Di Nuovo, M. Palesi, D. Patti

"High Performance Computing for Embedded System Design: A Case Study"

DSD'08, Digital System Design Architectures, Methods and Tools, Parma, 2008, pp. 656-659.

Additional references supplied on request.