

CURRICULUM VITAE

Name and Surname: Fabio Cumbo

Degree: Master of Science Degree in Computer Engineering

Current Positions:

- Ph.D. Candidate at the Department of Engineering – Third University of Rome, Rome, Italy
- Research Associate at the Institute for Systems Analysis and Computer Science “Antonio Ruberti” – National Research Council of Italy (IASI-CNR), Rome, Italy
- Research Lab Assistant at the Department of Biochemistry and Molecular Biology – The Pennsylvania State University, University Park, PA, USA

Research Interests: Bioinformatics, Data Integration, Graph Theory, Big Data

Known Languages:

- Italian Native Speaker
- English B1 Level certificate validated by the Third University of Rome

STUDIES

Attended Academic Institutes:

- Scuola Materna Divino Amore (Rome, Italy)
- Scuola Paritaria San Giovanni Battista (Rome, Italy)
- Liceo Scientifico Statale John Fitzgerald Kennedy (Rome, Italy)
- Istituto d’Istruzione Paritario Media Superiore Tozzi (Rome, Italy)
- Università degli Studi Roma Tre (Rome, Italy)

Degrees:

- Ongoing: Ph.D. Candidate at the department of Engineering – Third University of Rome
- July 24, 2014: Master of Science Degree in Computer Engineering at the Third University of Rome (Score: 102/110)
- October 25, 2012: Bachelor’s Degree in Computer Engineering at the Third University of Rome (Score: 89/110)
- July 8, 2008: Scientific Diploma at “Istituto d’Istruzione Paritario Media Superiore Tozzi” (Score: 62/100)

Date

October 17, 2017

ACADEMIC PROJECTS

October 2008: Joined the faculty of Computer Engineering at the Third University of Rome – Computer Engineering Department.

- Design and development of a library information system with the support of Java technologies (J2EE, JSP, JSF and JPA) and through the integration and interaction with a SQL DBMS (PostgreSQL).
- Design and development of a simple Twitter text classification for the “Artificial Intelligence” course. The classifier tries to recognize positive, negative and neutral messages using a Naïve-Bayes approach integrating the Stanford POS-tagging system (Part-Of-Speech Tagger) to improve the words analysis and tweets classification.
[LINK: <http://github.com/fabio-cumbo/twitter-text-analyzer/>]
- Design and development of a recommendation system for the Instagram social network for the “Intelligent Systems for Internet” course. The software tries to recommend points of interest as geographical coordinates analyzing the interests of a cluster of users in a specific geographical area.
[LINK: <http://github.com/fabio-cumbo/instagram-points-of-interest-analyzer/>]
- Design and development of a platform for the classification and visualization of a real-time Twitter stream for the “Information Visualization” course. The backend section of this project was implemented in Python and the frontend section in JavaScript using the D3.js library for a better visualization.

DEGREES AND CERTIFICATIONS

2014:

- Master of Science Degree in Computer Engineering with the score of 102/110 with a thesis titled “Time Dynamics of Protein Complexes in a Transgenic Mouse Model for Alzheimer’s Disease” on the analysis, design and development of an open source Java library for the analysis of time series microarray data. [July 24, 2014]

2013:

- Certificate of completion of the online course titled “Data Mining with Weka” provided by the Department of Computer Science of the University of Waikato, Hamilton, New Zealand.
[October 20, 2013]

2012:

- Bachelor’s Degree in Computer Engineering with the score of 89/110 with a thesis titled “Selecting relevant nodes and structures in biological networks: a new plugin for Cytoscape” on the analysis and development of a software for network analysis. [October 25, 2012]
- Microsoft 070-640 Certificate: TS: Windows Server 2008 Active Directory, Configuring.
[May 16, 2012]

2011:

- Microsoft 070-646 Certificate: Pro: Windows Server 2008, Server Administrator. [May 9, 2011]
- Microsoft 070-642 Certificate: TS: Windows Server 2008 Network Infrastructure, Configuring.
[April 27, 2011]
- I have attended the eighth edition of Agile Day held in the faculty of Computer Engineering at the Third University of Rome.

2008:

- Scientific Diploma with the score of 62/100 at “Istituto d’Istruzione Paritario Media Superiore Tozzi” with a thesis on staminal cells and their related ethical problems.

2007:

- ECDL certificate (European Computer Driving License).

Date

October 17, 2017

ATTENDED EVENTS

2016:

- Participation to the doctoral school “2nd SyBSyM Lake Como School – Systems Biology and Systems Medicine: Towards a Precision Medicine” at Villa del Grumello, Como, Italy.
- Participation to the conference “13th International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics” at the University of Stirling, UK.
- Participation to the doctoral school “1st SYSBIO.IT School on Computational Systems Biology” at the University of Milano-Bicocca, Milan, Italy.
- Participation to the doctoral school “2nd School on Scientific Data Analytics and Visualization” at the CINECA center of Bologna, Italy.

2015:

- Participation to the conference “Bioinformatics Italian Society 2015” (BITS 2015) with the presentation of a scientific poster titled “The Cancer Genome Atlas Data Querying Tool” at the University of Milano-Bicocca, Milan, Italy.
- Participation to the course “Parallel I/O and Management of Large Scientific Data” organized by CINECA. Topics: “Management of Large Scientific Data”, “Parallel File I/O with MPI2”, and “HDF5 Library” at the CINECA center of Rome, Italy.
- Certificate of participation at the workshop titled “Hands On Big Data: Getting Started With NoSQL And Hadoop” organized by the Codemotion at the Polo Didattico in Rome, Italy. The training objectives are: understanding the typical architecture of a Big Data system; setup, configure and use Hadoop; import data from SQL database into Hadoop; writing simple MapReduce jobs using Java; using Hive and Pig; connecting Hadoop with MongoDB; getting started with Apache Mahout; effectively communicate the results of processing using the techniques of data visualization.

2014:

- Participation to the Maker Faire Rome’s second edition. The Maker Faire is a showcase of innovation, creativity and resourcefulness and a celebration of the Maker movement. It’s a place where innovators show what they are making, and share know-how about technology and craftsmanship.
- Participation the Lazio 2014 edition of InnovAction Lab, the first course that allows you to bring your idea in front of investors, organized by Augusto Coppola, Paolo Merialdo, Carlo Alberto Pratesi and Roberto Magnifico. The presented entrepreneurship project was about a Self Service Social Intelligence B2B platform that helps companies to know what people think about their brand or products using a sentiment analysis engine on data extracted from social networks and the entire World Wide Web.

[STARTED: April, 2014 - FINISHED: June, 2014]

2013:

- Certificate of participation at the third edition of Codemotion held in the faculty of Computer Engineering at the Third University of Rome.
- Participation to the Google Cloud Dev Conf held in the LUISS EnLabs in Rome.
- Certificate of participation at the .NET Campus 2013 event held in the faculty of Computer Engineering at the Third University of Rome.
- Participation to the talk of Eugene Kaspersky on “Modern Cyberthreats Landscape” held in the LUISS EnLabs in Rome.

2012:

- Certificate of participation at the second edition of Codemotion held in the faculty of Computer Engineering at the Third University of Rome.
- Participation in the Hackathon held in the faculty of Computer Engineering at the Third University of Rome submitting a project with LEGO Mindstorm (Robotics LEGO kit that contains software and hardware to create small, customizable and programmable robots).
- Certificate of participation at the .NET Campus 2012 event held in the faculty of Computer Engineering at the Third University of Rome.

Date

October 17, 2017

2011:

- Participation to the first edition of Codemotion held in the faculty of Computer Engineering at the Third University of Rome.
- Certificate of participation at the .NET Campus 2011 event held in the faculty of Computer Engineering at the Third University of Rome.
- Certificate of participation at the .NET Development course held in the faculty of Computer Engineering at the Third University of Rome taught by Microsoft Student Partners attending the same University.

PROFESSIONAL EXPERIENCES

2017:

- Research Lab Assistant at the Department of Biochemistry and Molecular Biology of The Pennsylvania State University, University Park, PA, USA, Eberly College of Science Work Unit under the supervision of Prof. Anton Nekrutenko. The primary responsibility for this position is to work on the development of a portal, which allows querying large databases of genomic data.
[STARTED: September 1, 2017 – FINISHED: March 1, 2018]
- Internship at Wartik Laboratory – Department of Biochemistry and Molecular Biology – The Pennsylvania State University, University Park Campus, 16802 PA, Pennsylvania, USA. The aim of the internship is to collaborate with the team of the Galaxy project to extend and implement new features for the platform. I was involved in other projects concerning the development of new statistical analysis and algorithms, and contributions to Galaxy, Conda, and Bioconda projects. Some of these developed tools are listed below:
 - GDCWebApp: <https://github.com/fabio-cumbo/GDCWebApp4Galaxy>
 - IWTomics for Galaxy: <https://github.com/fabio-cumbo/IWTomics4Galaxy>[STARTED: March 1, 2017 – FINISHED: August 31, 2017]

2016:

- Tutor of two university courses (“Introduction to Big Data”, and “Big Data Analytics and Visualization”) at the Department of Engineering of the International Telematic University UNINETTUNO
[STARTED: September 20, 2016 – Ongoing]
- Professional collaboration with the Marine Technology Research Institute of the National Research Council of Italy – INSEAN-CNR for the development of a database containing information about military and merchant ships in which were used amiantus as a thermal insulator and information about officers and machinists affected by mesothelioma.
[STARTED: September 13, 2016 – Ongoing]

2015:

- Ph.D. Student at the Department of Engineering – Third University of Rome with a research project on the development of an innovative platform for the acquisition, storage, management, integration and analysis of heterogeneous biomedical data.
[STARTED: October 1 – Ongoing]
- Grant holder at IASI-CNR for a project on the “Analysis and development of toolkits for bioinformatics analysis. Extraction, storage and management of genomic data from TCGA”.
[STARTED: March 23, 2015 – FINISHED: September 23, 2015]

Date

October 17, 2017

2014:

- Member of “SYSBIO – Center for Systems Biology”, an Italian “open access” Research Infrastructure, distributed throughout different locations, focused on Systems Biology, the new paradigm for the study of Life Sciences, that integrates molecular analysis, also genome-wide, with math modeling and simulations. It’s funded by the Italian Ministry for Research (MIUR), within the Italian roadmap for Research Infrastructure, and operates under an agreement between the National Research Council (CNR) and the University of Milano-Bicocca.
Active member of the COSYS project, a platform for the interoperability of different software tools for a Systems Biology oriented analysis. The platform has to guarantee also data sharing between researchers of different European research centers. Regarding EU, SysBio aims to become a qualified hub of ISBE (Infrastructure for Systems Biology – Europe), a large-scale European research infrastructure project on the European Strategy Forum on Research Infrastructures (ESFRI) Roadmap. [STARTED: December, 2014 – Ongoing]
- Professional collaborator to research projects in the field of Bioinformatics at IASI-CNR. The performed activity consists in the "Design and implementation of a software for the storage, management and querying of genomic and clinical data. Application of the software to the TCGA. [STARTED: September 21, 2014 - FINISHED: March 21, 2015]
- Participation to the Data-Centric Genomic Computing (GenData 2020) project for the realization of new and innovative technologies for the support of next generation healthcare systems. The contribution to the project consists in the integration of a data extraction and analysis tool of genomic data from TCGA (The Cancer Genome Atlas) database of the USA National Cancer Institute and the USA National Human Genome Research Institute. The project is funded by the Ministry of Education, University and Research of Rome under the PRIN program and it is coordinated by the Politecnico di Milano University under the supervision of the Prof. Stefano Ceri. [STARTED: February, 2014 – Ongoing]

2013:

- Professional collaborator to research projects in the field of Bioinformatics at IASI-CNR continuing the study on relevant changes in the structure of protein complexes in the transgenic Mouse organism with Alzheimer’s disease in collaboration with EBRI. [STARTED: October 1, 2013 - FINISHED: July 31, 2014]
- Professional collaborator to research projects in the field of Bioinformatics at IASI-CNR. The performed activity consists in "Software development for the analysis of Protein-Protein Interaction Networks (PPI)". The main research activity concerned the analysis of significant changes in the structures of protein complexes starting from temporal gene expression microarray data for the transgenic Mouse organism with Alzheimer’s disease. Microarray data, provided by the European Brain Research Institute (EBRI) – Rita Levi Montalcini Foundation, contain gene expression profiles for two different control classes (VH and AD11) for the same organism relative to four time series on three different tissues (cerebral cortex, basal forebrain and hippocampus). Databases used to obtain information on protein complexes are CORUM, Gene Ontology and STRING. [STARTED: February 7, 2013 - FINISHED: July 6, 2013]
- Member of the research group in Bioinformatics, Computational Biology and Systems Biology as Research Assistant at IASI-CNR. [LINK: http://dmb.iasi.cnr.it/about_us.php]

2012:

- Collaborator at ITALWARE S.R.L. for a census of the Information Technology equipment of the Italian Ministry of Economy and Finance in Via Dalmazia, Via Liguria, Via Sicilia and Via Casilina in Rome.

2011:

- Internship at IASI-CNR on “design and development of algorithms for the computation of characteristic parameters in biological networks” under the supervision of Dr. Paola Bertolazzi (Research Director) and Dr. Giovanni Felici (Senior Researcher) and the support of the “Computational and Systems Biology” research group. [STARTED: September 15, 2011 – FINISHED: October 25, 2012]

Date

October 17, 2017

TALKS

2016:

- Participant and speaker to the “13th International Conference on Computational Intelligence Methods for Bioinformatics and Biostatistics” at the University of Stirling, UK, with a talk titled “IRIS-TCGA: An Information Retrieval and Integration System for Cancer Genomic Data” about a web tool able to perform the retrieval of genomic and clinical data from TCGA and then filter and integrate them using built-in integration methods.
[LINK: <http://www.cs.stir.ac.uk/events/cibb2016/>]
- Scientific committee member and speaker at the “1st SYSBIO.IT School on Computational Systems Biology” at the University of Milano-Bicocca with a talk titled “Introduction to COSYS platform” about a web platform able to manage, simulate, analyze, and visualize biochemical models.
[LINK: <http://www.sysbio.it/school-2016/>]

NON-PROFESSIONAL EXPERIENCES

2012:

- I have developed an application named ‘SharingApp’ for the sharing of educational materials on behalf of the faculty of Aerospace Engineering of the La Sapienza University of Rome, Italy.
[LINK: <http://aerosharingapp.altervista.org/>]

2011:

- I have contributed to the Gephi project, an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs. My contributions to the project include Italian translation of entire project and beta-testing activity.
[LINK: <https://gephi.org/about/people/>]
- I have contributed to develop plugins for Cytoscape project (an open source software platform for visualizing complex networks and integrating these with any type of attribute data).
[LINK: http://chianti.ucsd.edu/cyto_web/plugins/]
- I have collaborated as a beta tester for the weekly online magazine Hakin9 (IT Security Magazine). My role was to help in beta testing the received articles, evaluating its substantial content and suggesting corrections. [LINK: <http://hakin9.org/>]

PUBLICATIONS

2017:

- Fabio Cumbo, Anton Nekrutenko, and Giovanni Felici, "GDCWebApp: filtering, extracting, and converting genomic and clinical data from the Genomic Data Commons portal."
[TYPE: Research Article – Submitted to the 2017 Cold Spring Harbor meeting: Genome Informatics – New York City, NY, USA]
- Fabrizio Celli, Fabio Cumbo, and Emanuel Weitschek, "Classification of large DNA methylation data sets for identifying cancer drivers: BIGBIOCL."
[TYPE: Research Article – Ready – to be submitted]
- Marzia A. Cremona, Alessia Pini, Fabio Cumbo, Kateryna D. Makova, Francesca Chiaromonte, and Simone Vantini, "IWTomics: testing high-resolution "Omics" data at multiple locations and scales."
[TYPE: Research Article – Ready – to be submitted]
- Fabio Cumbo, Davide Vergni, and Daniele Santoni, "**Investigating transcription factor synergism in humans.**" DNA Research 2017
[TYPE: Research Article – ACCEPTED – WAITING FOR PUBLISHING]
- Fabio Cumbo, Marco S. Nobile, Chiara Damiani, Riccardo Colombo, Giancarlo Mauri, and Paolo Cazzaniga, "**COSYS: A Computational Infrastructure for Systems Biology.**" Lecture Notes in Bioinformatics vol. 10477, Computational Intelligence Methods for Bioinformatics and Biostatistics, Springer, October, 2017
[TYPE: Research Article – PUBLISHED: October 1, 2017]
- Fabio Cumbo, Emanuel Weitschek, Paola Bertolazzi, and Giovanni Felici, "**IRIS-TCGA: an information retrieval and integration system for genomic data of cancer.**" Lecture Notes in Bioinformatics vol. 10477, Computational Intelligence Methods for Bioinformatics and Biostatistics, Springer, October, 2017
[TYPE: Research Article – PUBLISHED: October 1, 2017]
- Fabio Cumbo, Giulia Fiscon, Marco Masseroli, Stefano Ceri, and Emanuel Weitschek. "**TCGA2BED: Extracting, Extending, Integrating, and Querying The Cancer Genome Atlas.**" BMC Bioinformatics 2016
[TYPE: Research Article – PUBLISHED: January 3, 2017]

2016:

- Emanuel Weitschek, Fabio Cumbo, Eleonora Cappelli, and Giovanni Felici, "Genomic Data Integration: A case study on next generation sequencing of cancer." 2016 27th International Workshop on Database and Expert Systems Applications (DEXA)

2015:

- Fabio Cumbo, Giulia Fiscon, Stefano Ceri, Marco Masseroli, and Emanuel Weitschek. "**The Cancer Genome Atlas Data Querying Tool.**" BITS 2015 Conference – 12th Annual Meeting of the Bioinformatics Italian Society
[TYPE: Poster – EXPOSED: June 3-5, 2015 – University of "Milano-Bicocca", Milan, Italy]
- Ivan Arisi, Mara D'Onofrio, Rossella Brandi, Antonino Cattaneo, Paola Bertolazzi, Fabio Cumbo, Giovanni Felici, and Concettina Guerra. "**Time Dynamics Of Protein Complexes In The AD11 Transgenic Mouse Model For Alzheimer's Disease Like Pathology.**" BMC Neuroscience 16, n. 1 (2015): 28.
[TYPE: Research Article – PUBLISHED: April 29, 2015]

2014:

- Fabio Cumbo, Giovanni Felici, and Paola Bertolazzi. "**Selecting Relevant Nodes And Structures In Biological Networks. BiNAT: A New Plugin For Cytoscape.**" F1000Research 2014, 3:287 (doi: 10.12688/f1000research.5753.1).
[TYPE: Software Tool Article - PUBLISHED: November 21, 2014]
- Fabio Cumbo, Paola Paci, Daniele Santoni, Luisa Di Paola, and Alessandro Giuliani. "**GIANT: A Cytoscape Plugin for Modular Networks.**" PLOS ONE 9, no. 10 (2014): e105001.
[TYPE: Research Article – PUBLISHED: October 2, 2014]

Date

October 17, 2017

IT SKILLS

Known IDE:

- Excellent knowledge of Eclipse IDE platform for Java and Java-EE.
- Excellent knowledge of Bloodshed Dev-C++ platform for C/C++.
- Excellent knowledge of NetBeans IDE platform for Java.
- Good knowledge of Microsoft Visual Studio 2010.

Programming Languages:

- Excellent knowledge of Java, J2EE and SQL languages.
- Good knowledge of C, PYTHON, PHP, HTML, CSS, BASH, XML, XPath and XQuery languages.
- Good knowledge of UML modeling language.
- Good knowledge of R functional language and environment for statistical computing and graphics.
- Good knowledge of Agile software development methodologies.
- Good knowledge of JUnit testing framework for the Java programming language.

Network Emulator:

- Good knowledge of the creation of virtual servers and networks with Hyper-V, VMware Workstation and Netkit.

Server Side Applications:

- Excellent knowledge of PostgreSQL, MySQL and the embedded H2 relational DBMS.
- Good knowledge of Apache Hive data warehouse software for querying and managing large datasets residing in distributed storage.
- Good knowledge of MongoDB, an open-source document database, and the leading NoSQL database.
- Good knowledge of Apache and Tomcat web servers.

Machine Learning, Data Mining and Network tools:

- Excellent knowledge of network analysis and visualization software usage (Cytoscape and Gephi).
- Excellent knowledge of Weka software for Data Mining.
- Good knowledge of the Apache Mahout, a scalable machine learning library.

Parallel and Distributed Computing:

- Excellent knowledge of the Apache Hadoop framework.
- Good knowledge of Pig, a SQL-like high-level platform for creating MapReduce programs used with Hadoop.

Other:

- Excellent knowledge of Android applications development.
- Good knowledge of parallel programming.
- Good knowledge of MapReduce paradigm.
- Excellent knowledge in MATLAB software.
- Good knowledge of version control software (SVN and GIT).
- Basic skills in Statistics.
- Excellent knowledge of the TeX typesetting language.
- Good knowledge of the Apache-POI reporting tool.
- Good knowledge of the Apache Solr framework.

Date

October 17, 2017