RESUMÉ

MICHELANGELO PULIGA PHD

BRIEF CV

Bio

I'm an Italian researcher and data scientist (master degree in Physics and PhD in Civil Engineering) with a long experience on Data Science and analytical methods. With a significant expertise in data manipulation, data mining with a variety of techniques. I have a demonstrated record of publication where the usage of Machine Learning for database creation, model improvement and deploy. I have spent most of my time in academia but I have also worked as a consultant for companies, and as a full time Unix Sysadmin. For few years I was a freelance web applications' developer. I worked in academia both in Italy and abroad (UK and CH). Currently I am Chief Data Scientist of a privately held laboratory Linkalab that I have cofounded in 2008 and it is active in the field of Natural language processing, Machine learning, Knowledge graphs and cloud architectures.

My early education has been as a Physicist involved in simulation of computational materials science, therefore I moved to the ICT Industry (Tiscali SPA) working for 3 years as a Unix System administrator (mostly web Apache, firewalls like NetApp, and database such as Oracle). Then I have worked as a freelance in web application development (creating ad hoc booking systems for hotels, and other web apps in Php).

I came back to academia in 2006 using my knowledge of Unix systems in the High computing facility of Consorzio Cybersar where I realized experiments of virtualization in a Grid environment (initially the INFN grid). During this period I also started to learn the early technologies of the cloud computing and the Amazon AWS pioneering its ecosystem.

In 2010 I obtained a PhD in Civil Engineering working to statistical models of extreme events in Hydrology. I had the possibility, during the Phd to improve my knowledge of GIS tools (Grass,Arcgis,QGis and other related) and my skills in managing large amounts of climatic data for the CLIMB FP7 project. I was visiting Phd student in Reading UK, (dept. of Mathematics).

In 2010 I moved to Zurich CH to work in the ETH Chair of system design for an international project oriented to forecast/understand the economic crisis (FoC project). My work there consisted in investigating systemic risk in financial institutions using tools from complex networks (graph theory). I was also in charge to prepare/acquire data from complex financial databases. In the same project I was also responsible of the data platform that we realized in Drupal with several extensions in Python (Turbogears) and we used for internal computation as well as for public demonstrations (for instance to the EU authorities during the revision phase of the project). Among others we published an important methodology to test the systemic risk of banks interlinked by a chain of assets/liabilities; this method known as "DebtRank" is nowadays, with minor modifications, used by central banks to forecast the risk associated to the banking sector in Europe.

After the Swiss period I moved to Lucca Italy to work as a researcher in the Axis unit of economics and data science of the IMT Alti studi institute. Once in Lucca I started again to follow the activities of the research and consultancy company Linkalab we founded in 2008 in Cagliari with other researchers. With the role of principal data scientist I've participated in several important projects for large media Italian companies such as Sky, Rai, La Stampa. We also worked together with institutions such as Camere di commercio della Calabria, Regione Piemonte, Regione Emilia Romagna in several projects related to urban mobility, banking and security. In those projects I was in charge of the choice of the best technologies, prototyping and in particular data acquisition/analysis. I also had the possibility to coordinate a team of 4/5 developers for the larger projects using methodologies of agile development and extreme programming. We realized an interesting pilot project with SIAE on digital transformation.

Finally in Linkalab we realized several Analytics as a service securing data from Twitter streams and Facebook APIs, we also participated in a round of investors for a new company (BigPress) that is still in the incubator early stage. For this company I was appointed by the colleagues of Linkalab and other partners as the designated BigPress CEO (2014-2015). For the moment these projects are in standby.

After the period in Lucca I'am working mostly with Linkalab with several customers in Milano especially at the ENI datalab and Talent Garden. I was also invited to teach in the class of Luiss Business school for two years in Rome. Actually I am professional fellow of scuola superiore Sant'Anna Pisa, and I am a member of the European team for Business data science: an international master by Talent Garden.

To complete my expertise I have tutored more than 200 students in 5 years of classes of Machine Learning and AI for the Talent Garden company, Italy (and Austria). I was appointed faculty manager of the Business Data Science course in the fall 2021 held for the TAG Vienna, Austria.

Research interests

My research activity and interests are oriented to the study of the economic and social impact of online platforms in particular social networks. Getting data from Twitter and other social networks is the key to evaluate new social phenomena to demonstrate hypothesis and trends. I use a a mix of machine learning methodologies, statistical tests and network analysis to infer robust insights. I spent several years building research databases, from finance, to public health, from Twitter, to Facebook, from world bank to ownership data (Eikon). All of these works ended in peer reviewed journals. I published 28 works in several journals, among others: Nature Scientific Reports, Plos One, Health Policy, Social Science & Medicine, Structural Change and Economic Dynamics. My research is then strongly multidisciplinary but mostly related to economics, economic policy and public health.

Fields of interest in science and technology

- Complex networks
- Machine learning and optimization method
- Data visualization, Data optimization and analytics
- NoSQL databases
- Systemic risk with clustering methods
- Spatial econometrics and long range correlation: mapping tools
- Social networks analysis
- Public health and disease tracking
- Ownership networks, health policy
- Differential privacy and Blockchain
- Python/R Data and computing environment

Skills & Tools:

- Team management (Agile XP programming), Project writing, Project Design.
- Teaching to business classes and university students
- Data Science, Social Networks Analysis, Machine Learning, and Deep learning, Text Analysis, Statistical analysis.
- Scraping technologies and Web Interaction
- GIS science and mapping Tools, Financial Networks analysis
- Cloud computing in the Amazon/Digital Ocean ecosystems, Restful API creation, data optimization and Analytics as a service
- API of the major social networks (Youtube, Twitter, Facebook), Javascript visualization libraries
- Database technologies Postgres/Mysql/MongDB. Python,R,C programming languages

WORKING EXPERIENCE

04/2021–present	Faculty member at Talent Garden for the international master in Data Science.
06/2018–06-2020	Short contracts with Scuola Superiore Sant'Anna and IMT Alti studi Lucca in specific projects
06/2014–06-2017	Postdoc Researcher at IMT Alti Studi Lucca, AXES section of economics and data science
06/2013–06/2014	Web/Tech platform responsible for Crisilab at IMT Alti Studi, Lucca.
12/2010–05/2013	Postdoc Researcher in the Chair of System Design ETH, Zurich. Research in financial stability and financial networks for the FOC (Forecast of Economic Crisis) European project. Development of a visualization platform (accessible at `http://www.focproject.eu`.)
10/2006–10/2010	Junior researcher in informatics and High performance Computing for Cybersar project of the Consorzio Cosmolab (also Consorzio Cybersar) , Cagliari, Italy.
01/2009–12/2009	Project team leader for the e-Works platform and for the Hamatole Electronic Ink development projects, Cagliari.
01/2008–present	Co-founder of the Linkalab Srl (complex networks computational laboratory) Cagliari.
01/2006–present	Free lance activity as a web applications developer and Unix sysadmin (mail servers, vpn-firewall and web applications).
11/2004–10/2005	Oracle DBA and PLSQL developer for Tiscali Srl Italia
05/2002–10/2004	Unix system administrator for the Europan broadband provider Tiscali B2B Italy.
11/2001–04/2002	Research contract: Realizing a FORTRAN numerical code for Molecular Dynamics of linear defects in CU for the Istituto Nazionale di Fisica della Materia, Italy (L. Colombo Supervisor).

EDUCATION & RESEARCH

10/1995–10/2001	Undergraduate Studies in Physics, Università degli Studi di Cagliari. Undergraduate Thesis: *Interaction between grain boundaries and dislocations in CU*; Supervisor: Prof. Luciano Colombo
04/2010–10/2010	PhD Ad Hoc programme in Mathematics, University of Reading, UK

01/2008 / 01-2011	PhD at the Università degli Studi di Cagliari, Dept. Civil Engineering Project title: *Extreme events in hydrology and climate change*; Supervisors: Prof. Roberto Deidda, Dr. Valerio Lucarini (University of Reading, UK)
12/2010–05/2013	Postdoctoral position at Chair of System Design, DMTEC, ETH Zurich, CH
06/2013–06/2014	Web platform supervisor at Crisislab project, IMT Alti Studi, Lucca, IT
06/2014–12/2016	Postdoctoral research fellow at Crisislab project, IMT Alti Studi, Lucca, IT, Axes Unit
09/2017–10/2017	Teacher of Natural language processing at the Luiss Big data business school
09/2018–10/2018	Teacher of Natural language processing at the Luiss Big data business school
11/2019–12/2019	Teacher of Deep learning at Talent Garden DataScience & AI master 2019
11/2020–12/2020	Teacher of Deep learning at Talent Garden DataScience & AI master 2020
11/2021–12/2021	Teacher of Deep learning at Talent Garden DataScience & AI master, Business data science international master

PROGRAMMING LANGUAGES, SERVERS, CMS AND VIRTUALIZATION TECHNOLOGIES

Programming languages	Python, C, R
Frameworks	Flask, Jquery, D3js, scraping Scrapy
Servers	Apache, Openvpn, Iptables
Virtualization technologies	VMware, Vbox, Docker
Cloud technologies	Amazon EC2, S3, EMR, DynamoDB, SNS, SQS, lambda
Scientific computing technologies	Jupyter, Dask, Scikit-learn, Pandas, Keras, PyCaret
Database technologies	Mysql, Postgres, PostGIS, MongoDB
GIS mapping	Grass,Shapely,QGis,Leaflet

LANGUAGE KNOWLEDGE

Italian	native
English	fluent
French	good
Spanish	fair

PUBLICATION LIST (REVERSED CHRONOLOGICAL ORDER)

[1] M. Lopreite, M. Puliga, M. Riccaboni, and S. De Rosis. A social network analysis of the organizations focusing on tuberculosis, malaria and pneumonia. *Social Science & Medicine*, 278:113940, 2021.

[2] M. Lopreite, P. Panzarasa, M. Puliga, and M. Riccaboni. Early warnings of covid-19 outbreaks across europe from social media. *Scientific reports*, 11(1):1–7, 2021.

[3] S. De Rosis, M. Lopreite, M. Puliga, and M. Vainieri. The early weeks of the italian covid-19 outbreak: sentiment insights from a twitter analysis. *Health Policy (Amsterdam, Netherlands)*, 2021.

[4] E. Brancaccio, R. Giammetti, M. Lopreite, and M. Puliga. Monetary policy, crisis and capital centralization in corporate ownership and control networks: A b-var analysis. *Structural Change and Economic Dynamics*, 51:55–66, 2019.

[5] Z. Zhu, G. Morrison, M. Puliga, A. Chessa, and M. Riccaboni. The similarity of global value chains: A network-based measure. *Network Science*, 6(4):607–632, 2018.

[6] M. Lopreite, M. Puliga, and M. Riccaboni. The global health networks: A comparative analysis of tuberculosis, malaria and pneumonia using social media data. 2018.

[7] E. Brancaccio, R. Giammetti, M. Lopreite, and M. Puliga. Centralization of capital and financial crisis: A global network analysis of corporate control. *Structural Change and Economic Dynamics*, 45:94–104, 2018.

[8] A. M. Petersen and M. Puliga. High-skilled labour mobility in europe before and after the 2004 enlargement. *Journal of the Royal society interface*, 14(128):20170030, 2017.

[9] G. Biorci, A. Emina, M. Puliga, L. Sella, and G. Vivaldo. Tweet-tales: moods of socio-economic crisis? In *Data Science and Social Research*, pages 205–213. Springer, 2017.

[10] M. Puliga, A. Flori, G. Pappalardo, A. Chessa, and F. Pammolli. The accounting network: How financial institutions react to systemic crisis. *PLoS One*, 11(10):e0162855, 2016.

[11] A. Langousis, A. Mamalakis, M. Puliga, and R. Deidda. Threshold detection for the generalized pareto distribution: Review of representative methods and application to the noaa ncdc daily rainfall database. *Water Resources Research*, 52(4):2659–2681, 2016.

[12] A. Bessi, F. Zollo, M. Del Vicario, M. Puliga, A. Scala, G. Caldarelli, B. Uzzi, and W. Quattrociocchi. Users polarization on facebook and youtube. *PloS one*, 11(8):e0159641, 2016.

[13] Z. Zhu, M. Puliga, F. Cerina, A. Chessa, and M. Riccaboni. Global value trees. *PloS one*, 10(5):e0126699, 2015.

[14] Y.-H. Eom, M. Puliga, J. Smailović, I. Mozetič, and G. Caldarelli. Twitter-based analysis of the dynamics of collective attention to political parties. *PloS one*, 10(7):e0131184, 2015.

[15] M. Puliga, G. Caldarelli, and S. Battiston. Credit default swaps networks and systemic risk. *Scientific reports*, 4(1):1–8, 2014.

[16] M. Popović, H. Štefančić, B. Sluban, P. Kralj Novak, M. Grčar, I. Mozetič, M. Puliga, and V. Zlatić. Extraction of temporal networks from term co-occurrences in online textual sources. *PLoS One*, 9(12):e99515, 2014.

[17] A. Gabrielli, S. Battiston, G. Caldarelli, N. Musmeci, and M. Puliga. Reconstructing topological properties of complex networks from partial information using the fitness model. In *APS March Meeting Abstracts*, volume 2014, pages P1–265, 2014.

[18] C. Dal Maso, G. Pompa, M. Puliga, G. Riotta, and A. Chessa. Voting behavior, coalitions and government strength through a complex network analysis. *PloS one*, 9(12):e116046, 2014.

[19] G. Cimini, T. Squartini, N. Musmeci, M. Puliga, A. Gabrielli, D. Garlaschelli, S. Battiston, and G. Caldarelli. Reconstructing topological properties of complex networks using the fitness model. In *International Conference on Social Informatics*, pages 323–333. Springer, 2014.

[20] G. Caldarelli, A. Chessa, F. Pammolli, G. Pompa, M. Puliga, M. Riccaboni, and G. Riotta. A multi-level geographical study of italian political elections from twitter data. *PloS one*, 9(5):e95809, 2014.

[21] G. Caldarelli, S. Battiston, M. Puliga, W. Quattrociocchi, and A. Scala. Multilevel complex networks and systems. *Bulletin of the American Physical Society*, 59, 2014.

[22] N. Musmeci, S. Battiston, G. Caldarelli, M. Puliga, and A. Gabrielli. Bootstrapping topological properties and systemic risk of complex networks using the fitness model. *Journal of Statistical Physics*, 151(3):720–734, 2013.

[23] G. Caldarelli, A. Chessa, F. Pammolli, A. Gabrielli, and M. Puliga. Reconstructing a credit network. *Nature Physics*, 9(3):125–126, 2013.

[24] G. Caldarelli, S. Battiston, M. Puliga, R. Kaushik, P. Tasca, C. of System Design Collaboration, I. A. S. L. Collaboration, et al. Debtrank a centrality measure for financial systems and beyond. In *APS March Meeting Abstracts*, volume 2013, pages G28–008, 2013.

[25] S. Battiston, M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli. Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, 2(1):1–6, 2012.

[26] M. Puliga. Extreme events in hydrology: an approach using exploratory statistics and the generalized pareto distribution. performances and properties of the gpd estimators with outliers and rounded-off datasets. 2011.

[27] R. Deidda and M. Puliga. Performances of some parameter estimators of the generalized pareto distribution over rounded-off samples. *Physics and Chemistry of the Earth, Parts A/B/C*, 34(10-12):626–634, 2009.

[28] R. Deidda and M. Puliga. Sensitivity of goodness-of-fit statistics to rainfall data rounding off. *Physics and Chemistry of the Earth, Parts A/B/C*, 31(18):1240–1251, 2006.

Date Signature

February 17, 2022