

# Antonio Mallia

+44 7598 951472/+39 328 7444631 • [me@antoniomallia.it](mailto:me@antoniomallia.it) • [www.antoniomallia.it](http://www.antoniomallia.it)  
[in it.linkedin.com/in/antoniomallia](https://in.linkedin.com/in/antoniomallia) • [S antonio.mallia](https://s.antonio.mallia) • [github.com/amallia](https://github.com/amallia)

---

## Work experience

2018–present **PhD Fellow**, *New York University*, New York, US.  
Teaching Assistant for Web Search Engines (Fall 2018)

Jun–Aug 2019 **Applied Scientist Intern**, *Amazon*, Barcelona, Spain.  
Working in the Search team

2016–2018 **Software Engineer**, *Bloomberg L.P.*, London, United Kingdom.  
Working in the Collaboration Frameworks team on a distributed architecture, and ensuring reliability, scalability and performance. Working in an Agile team, doing software development/testing, code optimization and code review. Learned how to work with large volumes of real-time data, and how to develop software in a large-scale, fast-paced environment.  
Detailed achievements:

- Involved in the design and development of a tagging and sharing framework.
- Designed and refactored code following the microservice paradigm.
- Created a Continuous Integration infrastructure to automate build, tests and deployment.
- Migrated the build system to use the more modern CMake.
- Introduced static code analysis and code coverage to run on the codebase.
- Investigated and debugged production environment issues caused by software defects, system configuration errors, and data inconsistencies.

2013–2016 **Software Engineer**, *NIC .it - National Research Council*, Pisa, Italy.  
Worked for the Italian National Research Council on projects for the ccTLD .it Registry.  
Detailed achievements:

- Developed several features of the main application used by NIC .it domain name registrars. Focused on database query optimization, export of data, third-party web-services integration.
- Co-operated on the development of an internal application, used by administrative staff to manage contracts and billing system. Focused on UX enhancement, data presentation and export.
- Co-operated on the development of an ETL system and presentation of data using interactive charts.
- Developed a UI library for JavaServer Faces (JSF) application and integration with the main NIC applications.
- Planned and developed an e-voting system used for the election of the Registry steering committee.

2011–2016 **Software Engineer**, *Self-Employed*, Remote.  
Worked as a freelancer remotely and on personal projects.  
Detailed achievements:

- Planned and developed a web application providing web privacy and anonymity tools: user management, load balancing, payment integration (Paypal, Bitcoin, credit card gateway, phone). Design of a RESTful API for third-party integration.
- Mail server configuration and administration.
- Google Chrome and Mozilla Firefox extension development to integrate features of a web application.
- Planned and developed a CMS to perform feed aggregation using parsing technics and categorisation.
- Software prototyping of a crawling and parsing system of an airline company to extract flight information and post-analysis.
- Developed a ad-hoc download manager software, including packaging for Windows and OS X.
- Owned an Italian blog network about technology widely-read and realised using Wordpress. Development of several plugins, integration with advertising agency web services. System administration main focused on performance optimisation.

## Education

2018–present **Ph.D. Program in Computer Science**, *New York University*.  
Pursuing a Ph.D. with a focus on advancing efficiency in Information Retrieval for large-scale systems.

2018–2020 **Master's degree in Computer Science, New York University.**

Relevant exams:

- Machine Learning
- Natural Language Processing
- Graphics Processing Units (GPUs): Architecture and Programming
- Big Data
- Deep Learning

2014–2015 **Master's degree in Computer Science, University of Pisa, 110/110 cum laude.**

This Master's Degree in Computer Science provided me a solid, in-depth background, in line with the innovation needs of the informatics field. The courses attended increased my knowledge portfolio, in both the theoretical and practical aspect.

**Thesis:** *Efficient top-k document retrieval with two level indexes.* Supervisor: Prof. Rossano Venturini

Relevant exams:

- Advanced Algorithms
- Advanced Database Systems
- Advanced Programming
- Computational Models
- Distributed Systems: paradigms and models
- Numerical Methods and Optimization
- Principle of Programming Languages

2012–2013 **Erasmus Programme, University of Essex.**

This experience enabled me to understand the value of studying and working in an international environment. Exploring practice-oriented learning methods improved my professional skills.

Relevant exams:

- Programming in Java
- Web Technologies & XML
- Large Scale Software Development & Extreme Programming
- Network Laboratory
- Photonics Networks & Devices
- Mobile Robotics
- Programming Embedded Systems
- Electronic System Design & Integration

2008–2012 **Bachelor's degree in Electronic Engineering, University of Pisa.**

I acquired knowledge of methodological aspects of mathematics and basic disciplines to understand and provide appropriate solutions to engineering problems. During the Curriculum in Electronic Engineering I developed professionally with sound fundamental and multidisciplinary skills ranging from electronics, circuit theory, systems theory, computer science, communications, control theory, economics, electromagnetism, nanoelectronics.

---

## Publications

Antonio Mallia, Michal Siedlaczek, Mengyang Sun, and Torsten Suel. **A Comparison of Top-k Threshold Estimation Techniques for Disjunctive Query Processing.** In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM)*, 2020.

Jimmy Lin, Joel Mackenzie, Chris Kamphuis, Craig Macdonald, Antonio Mallia, Michal Siedlaczek, Andrew Trotman, and Arjen de Vries. **Supporting Interoperability Between Open-Source Search Engines with the Common Index File Format.** In *Proceedings of the Conference on Research and Development in Information Retrieval (SIGIR)*, 2020.

Antonio Mallia, Michal Siedlaczek, Torsten Suel, and Mohamed Zahran. **GPU-Accelerated Decoding of Integer Lists.** In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM)*, 2019.

Antonio Mallia, Michal Siedlaczek, Joel Mackenzie, and Torsten Suel. **PISA: Performant Indexes and Search for Academia.** In *Proceedings of the Open-Source IR Replicability Challenge (OSIRRC) co-located with SIGIR*, 2019.

Antonio Mallia and Elia Porciani. **Faster BlockMax WAND with Longer Skipping**. In *Proceedings of the European Conference on Information Retrieval (ECIR)*, 2019.

Antonio Mallia, Michal Siedlaczek, and Torsten Suel. **An Experimental Study of Index Compression and DAAT Query Processing Methods**. In *Proceedings of the European Conference on Information Retrieval (ECIR)*, 2019.

Joel Mackenzie, Antonio Mallia, Matthias Petri, J. Shane Culpepper, and Torsten Suel. **Compressing Inverted Indexes with Recursive Graph Bisection: A Reproducibility Study**. In *Proceedings of the European Conference on Information Retrieval (ECIR)*, 2019.

Melanie Tosik, Antonio Mallia, and Kedar Gangopadhyay. **Debunking Fake News One Feature at a Time**. *CoRR*, abs/1808.02831, 2018.

Antonio Mallia, Giuseppe Ottaviano, Elia Porciani, Nicola Tonellotto, and Rossano Venturini. **Faster BlockMax WAND with Variable-sized Blocks**. In *Proceedings of the Conference on Research and Development in Information Retrieval (SIGIR)*, 2017.

---

## Other research activities

- PC member for CIKM 2020 Resource Track
- PC member for ECIR 2020
- Reviewer for ACM Transactions on the Web (TWEB)
- External reviewer for KDD 2018, KDD 2019, KDD 2020

---

## Books

Antonio Mallia and Francesco Zoffoli. **C++ Fundamentals**. Packt Publishing. 2019

---

## Languages

Italian **Native**

English **Professional working proficiency**

*TOEFL iBT: 105 (C1)*

---

## Computer skills

OS: Mac OS X, Linux and Windows.

Programming: Java, C++, C, Python, Javascript, PHP.

Database: MySQL, MongoDB, SQLite, Redis.

Web Design: HTML5, CSS3, jQuery.

Development: Eclipse, Sublime Text, PyCharm, IntelliJ IDEA, Clion, Vim.

Environment:

Project: SVN, Git, Trac, Gitlab.

Continuous: Jenkins, Travis CI.

Management:

Integration:

Web Server: Apache, Lighttpd, Jetty, JBoss AS.

Container engine: Docker.

Framework: Seam, Dropwizard, SlimPHP.

Mail: Postfix, Dovecot.

Documentation: Markdown, Latex.

CMS: Wordpress, Ghost.

---

## Invited Talks and Tutorials

04/2016 **Backend for Android Developers**, *Droidcon Italy*, Turin.

10/2017 **C++: unexpected behaviour**, *Topconf*, Düsseldorf.

11/2017 **C++: unexpected behaviour**, *Meeting C++*, Berlin.

---

## Personal skills

Leadership Good experience in team management acquired training and guiding a blog network editorial staff.