# Francesco Bonchi

*AI Innovator and Scientific Leader*

---

### ▬▬ Overview

Research leader with a double profile – **academic** and **industrial** – active in the international AI and Data Science research community **since 1998**. I have published more than **250 scientific papers** and edited more than **30 scientific volumes**. I have a wide experience in several real-world government and industrial AI and Data Science projects in various application domains. I've also filed **16 USA patents** of which **9 granted**, and received the prestigious **2013 Yahoo Master Inventor Award**. Finally, I also have **several years experience in the management** of research in an industrial lab, and I've been **trained to people management and leadership** in California, by Yahoo, over a 3-year period (2012-2014).

| | |
|---|---|
| Keywords: | *AI, machine learning, algorithms, data mining, big data, web mining, social networks, social media, graph mining, algorithmic bias, fairness by-design, privacy-preserving data mining, anonymity.* |
| H-index: | **62** (as of October 2023, with $> 15400$ citations, according to Google Scholar). |
| More info: | `http://www.francescobonchi.com` |

### ▬▬ Main positions

| | |
|---|---|
| Since 2022 | **Co-Founder** and **Research Director** at CENTAI, Turin, Italy. |
| Since 2022 | Member of the Board of Directors (CdA in Italian) of CENTAI. |
| Since 2022 | CEO at SOM S.r.l. (Sound of Mind): the holding owning the 51% of CENTAI. |
| Since 2022 | Member of the Board of Directors (CdA in Italian) of Anti-Financial Crime Digital Hub. |

### ▬▬ Other scientific positions and roles

| | |
|---|---|
| Since 2023 | **Section Editor** in "Computational Complexity" for PLOS Complex Systems |
| Since 2019 | Member of the Scientific Advisory Board of Competence Center Machine Learning Rhine-Ruhr (ML2R). |
| Since 2019 | Member of the Advisory Board of RAIS (Real-time Analytics for Internet of Sports) Marie Curie European Training Network. |
| Since 2018 | Member of the **Academic Board of the PhD program in Data Science** at La Sapienza University of Rome. |
| Since 2018 | Associate Editor of ACM Transactions on Knowledge Discovery from Data (TKDD). |
| Since 2018 | Member of the IEEE Data Science and Advanced Analytics (DSAA) *Steering Committee*. |
| Since 2015 | (Part-time) **Research Director**, Big Data & Data Science, at Eurecat (Technological Center of Catalonia), Barcelona, Spain. |
| Since 2010 | Member of the ECML PKDD *Steering Committee*. |

### ▬▬ Past positions and education

| | |
|---|---|
| 2020 – 2022 | **Scientific Director** at **ISI Foundation** Turin, Italy. |
| 2020 – 2022 | Coordinator of the "Learning and Algorithms for Data Analytics" Research Area at the ISI Foundation. |

| | |
|---|---|
| 2015 – 2022 | Member of the Editorial Board of Data Mining and Knowledge Discovery (DMKD) Journal. |
| 2013 – 2022 | Associate Editor of ACM Transactions on Intelligent Systems and Technology (TIST). |
| 2018 – 2022 | Associate Editor of International Journal of Data Science and Analytics (JDSA). |
| 2018 – 2020 | Deputy Director for the Industrial Research area at ISI Foundation, Turin, Italy. |
| 2015 – 2020 | Research Leader for "Algorithmic Data Analytics" group at ISI Foundation, Turin, Italy. |
| 2017 – 2019 | Member of the Scientific Advisory Board of Veritone Inc., (NASDAQ: VERI). |
| 2015 – 2019 | Associate Editor of IEEE Transactions on Big Data (TBD). |
| 2017 – 2018 | Member of the Faculty Committee of the Ph.D. School on *"Data Science and Computation"*, University of Bologna (first PhD School on Data Science in Italy). |
| 2013 – 2018 | Associate Editor of IEEE Transactions on Knowledge and Data Engineering (TKDE). |
| 2014 – 2015 | **Director of Research** at Yahoo Labs. |
| 2014 – 2015 | **Deputy head** of Yahoo Labs Barcelona. |
| 2012 – 2015 | Head of the Web Mining research group at Yahoo Labs. |
| 2014 – 2016 | Associate Editor of Knowledge and Information Systems: An International Journal (KAIS). |
| 2010 – 2014 | Senior Research Scientist at Yahoo Labs. |
| 2010 – 2013 | Organizer of the "Yahoo Labs Barcelona Seminars" series. |
| 2008 – 2010 | Research Scientist at Yahoo Labs. |
| 2006 – 2007 | Researcher ("Ricercatore art.23") at ISTI - CNR, Pisa, Italy. |
| 2002 – 2006 | PostDoc ("Assegno di ricerca") at ISTI - CNR, Pisa, Italy. |
| 1999 – 2003 | **Ph.D. in Computer Science** at University of Pisa, Italy. |
| 1998 | Research contract on the project *"Lotta all'evasione fiscale"* (Fiscal Fraud Detection) at CNUCE - CNR, Pisa, Italy, funded by the Italian Ministry of Finance. |
| 1992–1997 | **Laurea *magnum cum laude*** in Computer Science at University of Pisa, Italy. |

---— Awards

| | |
|---|---|
| 2022 | **Best Paper Award at The World Wide Web Conference (WWW 2022)**: one of the most prestigious and competitive international conferences in Computer Science (out of more than 1800 papers submitted, 300 accepted for presentation, only 1 best paper award). |
| 2018 | IEEE Data Science and Advanced Analytics (DSAA) **Service Award**, 2018. |
| 2013 | Yahoo **Master Inventor** award. |
| 2009 | *Best Paper Award* at the 2009 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2009). |
| 2009 | *Best Student Paper Award* at the 18th ACM Conference on Information and Knowledge Management (CIKM 2009). |
| 2009 | *Best Student Paper Runner-up* at the 18th ACM Conference on Information and Knowledge Management (CIKM 2009). |
| 2003 | *Student Travel Awards*, at The Third IEEE International Conference on Data Mining (ICDM 2003). |
| 1999 | *Student Travel Award*, at The ACM SIGKDD international conference on Knowledge discovery and data mining (KDD 1999). |

---— Patents and publications

I have filed **16 USA patents** of which **9 granted**. Further details about my patents can be found at `http://www.francescobonchi.com/patents.html`.

I have published more than **250 scientific papers**, many of which in the most prestigious venues of the Data Mining (e.g., KDD, ICDM), Web (e.g., WWW, SIGIR, WSDM), and Databases (e.g., VLDB, SIGMOD, ICDE) research areas. I have also authored or edited more than **30 scientific volumes**. The complete list of my publications including books, edited volumes, journals, conferences and book chapters can be found at `http://www.francescobonchi.com/publications.html`.

## Scientific service

Besides serving in the editorial board of several journals (see previous page) I regularly serve as Area Chair (or Senior PC member) in the Program Committee of the main conferences in the areas of Data Science, Machine Learning and Databases. The (almost) complete list of my service activities can be found at: `http://francescobonchi.com/service.html`.

In the following I report a selection of the main service activities.

### Conference chairing and organization

| | |
|---|---|
| KDD 2024 | I'll be **General Chair**, together with Ricardo Baeza-Yates, of **ACM SIGKDD 2024** conference, to be held in Barcelona, Spain, in August 2024. |
| ECML PKDD 2023 | **General Chair**,of *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, September 2023, Turin, Italy. We had more than 1400 registrations (more than the double of previous years). |
| WWW 2021 | **Workshop Organizer** of *Workshop on Misinformation and Integrity in Social Networks at the 30th Web Conference (WWW 2021)*, April 2021, Ljubljana, Slovenia. |
| ICMLA 2020 | **PC Chair** of *The 19th IEEE International Conference On Machine Learning And Applications*, December 2020, Miami FL, USA. |
| DSAA 2019 | **PC Chair (Applications Track)** of *The 6th IEEE International Conference on Data Science and Advanced Analytics*, October 2019, Washington DC, USA. |
| ASONAM 2019 | **Tutorial Chair** of *The 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, August 2019, Vancouver, Canada. |
| BSMDMA 2018 | **General Chair** of the *IEEE BigData 2018 Workshop on Big Social Media Data Management and Analysis*, December 2018, Seattle (WA), USA. |
| DSAA 2018 | **General Chair** of *The 5th IEEE International Conference on Data Science and Advanced Analytics*, October 2018, Turin, Italy. |
| CIKM 2018 | **Workshop Chair** of the *2018 ACM International Conference on Information and Knowledge Management*, October 2018, Turin, Italy. |
| ECML PKDD 2018 | **PC Chair** of the *European Conference on Machine Learning and Principle and Practice of Knowledge Discovery in Databases*, September 2018, Dublin, Ireland. |
| WWW 2018 | **PC Chair** of the *Social Network Analysis and Graph algorithms for the Web track* at The International World Wide Web Conference, April 2018, Lyon, France. |
| HT 2017 | **PC Chair** of the *28th ACM Conference on Hypertext and Hypermedia*, July 2017, Prague, Czech Republic. |
| ICDM 2016 | **PC Chair** of *IEEE International conference on Data Mining*, December 2016, Barcelona, Spain. |
| ECML PKDD 2015 | **Demonstration Chair** of *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, September 2015, Porto, Portugal. |
| ECML PKDD 2014 | **Sponsorhip Chair** of *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, September 2014, Nancy, France. |
| ICDM 2012 | **Demo Chair** of *The 2012 IEEE International Conference on Data Mining*, December 2012, Brussels, Belgium. |
| SWM 2011 | **PC Chair** of *The International Workshop on Social Web Mining* , July 2011, Barcelona, Spain. |
| ECML PKDD 2010 | **PC Chair** of *The 21st European Conference on Machine Learning and The 14th European Conference on Principles and Practice of Knowledge Discovery in Databases*, September 2010, Barcelona, Spain. |
| WSDM 2009 | **Promotion Chair** of *The Second ACM International Conference on Web Search and Data Mining*, February 2009, Barcelona, Spain. |
| ICDM 2008 | **Workshop Chair** of *The Eighth IEEE International Conference on Data Mining*, December 2008, Pisa, Italy. |
| PinKDD 2008 | **PC Chair** of *The Second ACM SIGKDD International Workshop on Privacy, Security, and Trust in KDD(PinKDD 2008)*, workshop held in conjunction with *The Fourteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, Nevada, August 2008. |
| PinKDD 2007 | **PC Chair** of *The First ACM SIGKDD International Workshop on Privacy, Security, and Trust in KDD(PinKDD 2007)*, workshop held in conjunction with *The Thirteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Jose, California, August 2007. |

| | |
|---|---|
| PADM 2006 | **PC Chair** of the *IEEE International Workshop on Privacy Aspects of Data Mining (PADM 2006)*, workshop held in conjunction with *The Sixth IEEE International Conference on Data Mining*, December 2006, Hong Kong. |
| KDID 2005 | **PC Chair** of *The Fourth International Workshop on Knowledge Discovery in Inductive Databases (KDID 2005)*, workshop held in conjunction with *The 16th European Conference on Machine Learning (ECML) and the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)*, October 2005, Porto, Portugal. |

### Keynotes, invited talks, tutorials and panels

| | |
|---|---|
| IDA 2023 | **Keynote Speaker** at *The 21st Symposium on Intelligent Data Analysis (IDA 2023)*, Louvain-la-Neuve, Belgium, 12-14, April 2023. |
| XKDD 2020 | **Keynote Speaker** at *The 2nd ECML PKDD International Workshop on eXplainable Knowledge Discovery in Data Mining*, Ghent, Belgium, September 14th, 2020. |
| CSoNet 2019 | **Keynote Speaker** at *The 8th International Conference on Computational Data and Social Networks*, Ho Chi Minh City, Vietnam, November 18-20, 2019. |
| GFAIH 2019 | **Invited Talk and Panel** at The Global Forum on AI for Humanity (GFAIH) in the session *"AI & Human values: inequalities, biases, fairness, representation of minorities, gender balance"* together with Ricardo Baeza-Yates, Kate Crawford, Laurence Devillers, and Eric Salobir. November 2019, Paris, France. |
| CONSOB - Bocconi 2019 | **Keynote Speaker** *"AI and Blockchain: Future Convergence and Synergies"* at the Public Hearing on Initial Coin Offerings organized by CONSOB at Bocconi University, May 21, 2019. |
| SIGMOD 2019 | **Panelist** (title:" The Responsibility Challenge for Data") at the 2019 International Conference on Management of Data, June 30-July 5, 2019, Amsterdam, Netherlands. |
| SOFSEM 2019 | **Keynote Speaker** at The 45th International Conference on Current Trends in Theory and Practice of Computer Science, Nový Smokovec, Slovakia, January 27-30, 2019. |
| NetSciX 2019 | **Keynote Speaker** at The International School and Conference on Network Science, Santiago, Chile, January 3-5, 2019. |
| ICDM 2018 | **Tutorial Instructor** at The 18th IEEE International Conference on Data Mining, November 17-20, 2018, Singapore. |
| Complex Networks 2017 | **Tutorial Instructor** at the 6th International Conference on Complex Networks and Their Applications. November 28, 2017, Lyon, France. |
| DAB 2017 | **Keynote Speaker** at the 1st CIKM Workshop on Data and Algorithmic Bias (DAB 2017), Singapore, November 6, 2017. |
| VLDB 2017 | **Tutorial Instructor** at the 43rd International Conference on Very Large Databases 2017 (VLDB 2017), August 28th - September 1, 2017, Munich, Germany. |
| DATA 2017 | **Keynote Speaker** at the 6th International Conference on Data Science, Technology and Applications (DATA 2017), Madrid, Spain, July 24-26, 2017. |
| ICWE 2017 | **Keynote Speaker** at the 17th International Conference on Web Engineering (ICWE'17), Rome, Italy, June 5-8, 2017. |
| ICWE 2017 | **Keynote Speaker** at the INNS Conference on Big Data, Thessaloniki, Greece, October 23-25, 2016. |
| CCS Warmup 2017 | **Keynote Speaker** at the CCS Warmup - The 3rd School on Complex Systems, Sept 16-18, 2016, Amsterdam. |
| KDD 2016 | **Tutorial Instructor** at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2016), August 13-17, 2016, San Francisco, California. |
| WWW 2016 | **Tutorial Instructor** at the 25th International World Wide Web Conference (WWW 2016), April 11-15, 2016, Montreal, Canada. |
| DSAA 2015 | **Keynote Speaker** at The 2015 IEEE International Conference on Data Science and Advanced Analytics (IEEE DSAA 2015), October 2015, Paris, France. |
| BigMine 2015 | **Keynote Speaker** at Big Data Mining, 4th International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine 2015), satellite event of KDD 2015, August 2015, Sydney, Australia. |
| iSocial | **Invited Speaker** at iSocial Workshop "Online Social Networks: Data Management and Privacy/Security", April 2015, Crete, Greece. |

| | |
|---|---|
| KDD 2014 | **Tutorial Instructor** at The 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2014), August 2014, New York City, USA. |
| WSDM 2014 | **Keynote Speaker** at the *Workhop on Diffusion Networks and Cascade Analytics*, February 2014, New York City, USA. |
| ICDM 2013 | **Keynote Speaker** at *The Third IEEE ICDM Workshop on Data Mining in Networks* (DaMNet 2013), December 2013, Dallas, Texas. |
| BIG DATA day | **Invited Speaker** at Journée de rencontre BIG DATA, September 2013, Toulouse, France. |
| ECML PKDD 2013 | **Keynote Speaker** at *The 4th International Workshop on Mining Ubiquitous and Social Environments* (MUSE), September 2013, Prague, Czech Republic. |
| ECCS 2013 | **Invited Speaker** at Workshop "Complexity in Online Social Networks and Big Data", September 2013, Barcelona, Spain. |
| ICDM 2012 | **Keynote Speaker** at the IEEE ICDM 2012 PhD Student Forum, December 2012, Brussels, Belgium. |
| MIA 2012 | **Invited Speaker** at *Institute for Mathematics and its Applications (IMA) Annual Program Year Workshop - "User-Centered Modeling"*, May 2012, University of Minnesota, USA. |
| WI-IAT 2011 | **Keynote Speaker** at *The 2011 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, August 2011, Lyon, France. |
| ICDM 2010 | **Panelist** (title:"Top-10 Data Mining Case Studies") at *The Tenth IEEE International Conference on Data Mining*, December 2010, Sydney, Australia. |
| ICDM 2008 | **Panelist** (title:"Social Networks and Data Mining: Where's the Beef?") at *The Eighth IEEE International Conference on Data Mining*, December 2008, Pisa, Italy. |

## PhD Supervising and PhD Committees

I've been **supervisor of 7 PhD students which already graduated** in the past, while I'm **currently supervising 9 PhD students.** I've also been member of **18 PhD defense committees**. I've been also supervisor for several undergrad thesis in Italy and Spain (2003-2023) and host for several research internships at Yahoo Labs (2008-2015), ISI Foundations and Eurecat (2015-2021), and CENTAI (2022-2023).

More information can be found at: `http://francescobonchi.com/service.html`.

## Industrial experience

During my time at University of Pisa and C.N.R. (1998 - 2007) I developed **ten years of consulting experience in AI and Data Science**. Starting in 1998 with a project about **fiscal fraud detection** (*Lotta all'evasione fiscale*) which was funded by the **Italian Ministry of Finance**: this was one of the first projects of this type in Italy, in which AI techniques were adopted to analyse data for a concrete and important goal such as fiscal fraud detection. After that, I had a plethora of different experiences in different application domains, such web mining (with one of the first internet company of Italy), promotions analysis, planning and churn prediction in retail domain (with Coop and Carrefour Italia), mobility data analysis (with the municipalities of Milan and Pisa), biomedical and genetical data mining in support of solid organ transplantation, data mining for bioinformatics, geographic knowledge discovery, just to mention a few.

Later, I've been working for eight years (2008-2015) in a purely industrial research lab, at Yahoo Labs in Barcelona, first as Research Scientist, then as Senior Research Scientist, and finally as **Director of Research**.

As an **industrial research lab manager**, I've been responsible for the areas Web and Data Mining, and Social Analysis. I've been the official Yahoo Labs liaison for products such as the **Yahoo Finance** portal and for the **Social Platforms** team. My duties included to find opportunities for my researchers to **impact the business of the company**, manage industrial research projects and maintain the relationships with the various stakeholders and business partners. I was also able to **influence the product roadmap** and set shared goals between the research team and the various product teams. It was my responsibility to **ensure implementation of the strategy**, to perform quarterly performance evaluation for each team member, set team goals for each quarter, as well as help the team members to set their own personal goals. It was also my responsibility to attract and nurture the right talent. Finally, I had the opportunity to interact with Yahoo Labs leaders to influence policies and to negotiate decisions, goals, and objectives for area of responsibility: e.g., for what concerns performance assessment criteria, staffing, and long term development of the areas of responsibility.

On the technical side, my team at Yahoo Labs was developing and applying novel methods to analyze web-scale data, obtaining actionable knowledge about people and content on the web, and using this knowledge to efficiently support products and innovative applications **impacting millions of users worldwide**. We have been close collaborators of

production teams related to search, advertisement, social platforms, machine learning and vertical products such as **Yahoo Finance**, **Tumblr** and IntoNow. In my team we were also developing **big scale Machine Learning**, Data Mining, and Information Retrieval libraries and platforms for internal use, as well as externally visible data analytics and visualization modules.

At ISI Foundation, in 2017 I've helped the creation of ICLAI (Innovation Center Lab for AI) a research lab shared with **Intesa Sanpaolo** (ISP), one of the largest financial institutions in Europe. Essentially, ISP was providing us the scientific challenges (mostly in the area of **machine learning for finance**) and the budget to execute the challenges; we were creating for ISP an industrial research lab embedded within a purely research institute. We were hiring the right research scientists for the given challenges, whose main task was to execute the challenges under my scientific leadership, while being able to maintain an academic profile and research activity. In parallel to the budget to execute each challenge, the funder also provided support, on a philanthropic basis, to the fundamental research activity. After the successful completion of the first 3 years of project, the partnership was renewed for another 3-years period. The budget **was > 6 million Euro for each 3-years period**, and it helped hiring a dynamic team of >15 top-class researchers from all over the world. We were able to attract some of the first senior hires for ICLAI, thanks to my personal professional network (several of them came from my ex-team at Yahoo Labs Barcelona).

Thanks to the success of the ICLAI experience, jointly with Intesa Sanpaolo in 2022 we created **CENTAI**, an **industrial research lab** whose ambition is to push the boundaries of applied and foundational research in domains related to Artificial Intelligence and Complex System sciences. CENTAI is a joint-stock company between SOM s.r.l. (51%) and ISP (49%). SOM s.r.l. is a company of researchers, of which I'm the unique executive officer. At CENTAI I act as **Research Director** and I'm a member of the **Board of Directors**.

In 2022 we also created the Anti-Financial Crime (AFC) Digital Hub, a consortium company between CENTAI, ISP, ISP Innovation Center, University of Turin and Polytechnic of Turin, whose goal is to develop research and innovation projects about the application of AI techniques to fighting financial crime. I'm a member of the Board of Directors of the AFC Digital Hub.

From 2017 to 2019 I've been member of the **Scientific Advisory Board of Veritone Inc.**, (NASDAQ: VERI), a California-based company providing AI and cognitive computing solutions to transform audio, video, and other unstructured data, as well as analyzes it in conjunction with structured data in an orchestrated and automated manner to generate actionable intelligence.

Further details regarding industrial experience can be provided upon request.