Federica Filippini

federica.filippini@unimib.it | ffilippini | FFede0
57222342275 | 0000-0002-6549-924X | F63c78AAAAAJ

Highlights

- 22 peer-reviewed publications, including 4 IEEE Transactions and 1 ACM Transaction, and 1 invited contribution.
- Co-inventor of one patent.
- Best Paper Award for "ML-Based Performance Modeling in Edge FaaS Systems" at the 11th European Conference on Service-Oriented and Cloud Computing (ESOCC) (2025).
- "Premio Giovani Talenti" (Young Talents Award) organized by University of Milano-Bicocca (2024, 1st prize).
- Springer Award for the Ph.D. Dissertation organized by Politecnico di Milano (2024).
- Best project award; 3rd call for Value Chain Technology Transfer H2020 TETRAMAX - ANDREAS (2022).
- Visiting researcher at the Center for Intelligent Information Technologies and their Applications (CETINIA), University Rey Juan Carlos, Madrid, Spain.
- Visiting researcher and intern at IBM Research Lab Dublin, Ireland.
- 2 invited talks/seminars/panels
- Based on Google Scholar (02/10/2025), my h-index is 8 and total number of citations is 120, with first published paper in 2020 (Scopus h-index 6, citations 72).
- My research interests include: optimization problems applied to resource selection and scheduling in Cloud and distributed Computing Continuum environments, Reinforcement Learning for the automatic runtime adaptation of Artificial Intelligence workloads and automatic load management of distributed FaaS environments, Machine Learning-based performance modeling, Agentic AI.

Academic Positions

2024/03 – present — **Post-Doc Researcher,** University of Milano-Bicocca (DISCo), Multi-Agent Reinforcement Learning for Component Placement and Offloading in the Computing Continuum.

2023/11 – 2024/02 — **Research Fellow,** Politecnico di Milano (DEIB), Resource allocation and scheduling for Artificial Intelligence applications in Edge Computing.

2020/06 – 2021/05 — **Research Fellow,** Politecnico di Milano (DEIB), Scheduling techniques for Deep Learning training jobs on GPUs; development of heuristic solutions for DL training jobs scheduling on GPU-based systems (between 2020/11 and 2021/05 the contract covered the Ph.D. activities).

International Scientific Collaborations

2025/04 – 2025/05 — 5-weeks **research visit** at the Center for Intelligent Information Technologies and their Applications (CETINIA), University Rey Juan Carlos, Madrid, Spain. Topic: Decentralized workload management for Function as a Service systems at the edge.

2022/06 – 2022/12 — 6-months **research internship** at the IBM Dublin Research Lab. Topic: Performance analysis and modeling of Deep Learning distributed training applications on Ray clusters.

Invited Talks and Seminars

- "Resource Allocation and Scheduling Problems in Computing Continua (for Artificial Intelligence and Function-as-a-Service Workflows)", Centro de Investigación para las Tecnologías Inteligentes de la Información y sus Aplicaciones (CETINIA), University Rey Juan Carlos, Madrid, Spain (https://www.cetinia.es/actividades/seminarios#h.5g6qqvi386d8)

International Scientific Collaborations (continued)

- Panelist at the Career Development Support Session for Early Career Researchers, Professional Development Workshop for Early Career Researchers, European Informatics Leaders Summit (ECSS) 2025, Rennes, France (https://www.informatics-europe.org/ecss/program/career-development-workshop.html)

Teaching Activity

2024 – present
- **Lecturer.** Politecnico di Milano, B.S. Management and Production Engineering, *098474 – Mathematical Methods for Operations Research* course (35 hours).

2020 – present
- **Lecturer.** Politecnico di Milano, M.Sc. Mathematical Engineering, *052496 – Algorithms and Parallel Computing* course.
  - a.a. 2020–2021: 33 hours (8 hours lectures, 23 hours exercise sessions, 2 hours laboratory)
  - a.a. 2021–2022: 30 hours (12 hours lectures, 18 hours exercise sessions)
  - a.a. 2022–2023: 22 hours (10 hours lectures, 10 hours exercise sessions, 2 hours laboratory)
  - a.a. 2023–2024: 32 hours (16 hours lectures, 14 hours exercise sessions, 2 hours laboratory)
  - a.a. 2024–2025: 18 hours (14 hours lectures, 4 hours exercise sessions)

2024
- **Laboratory Tutor.** University of Milano-Bicocca, M.Sc. Computer Science, *Cloud Computing* course (12 hours).
- **Tutor.** University of Milano-Bicocca, B.S. Computer Science, *Distributed Systems* course (20 hours).

2023
- **Lecturer.** Graduate School of Management, Politecnico di Milano, *GENERALI - Data Scientist School*, Ed. 6 and 7, "Hadoop Ecosystem introduction-Spark, from RDDs to Dataframes" module (4 hours in ed. 6, 4 hours in ed. 7).

2023 – 2024
- **Lecturer.** Graduate School of Management, Politecnico di Milano, *PE Data Science 23*, "Big Data Technologies" module (3 hours).

2022 – 2023
- **Lecturer.** Graduate School of Management, Politecnico di Milano, *PE Data Science & Business Analitycs 22*, "Big Data Technologies & Non Relational Databases" module (2 hours).

2022 – 2024
- **Lecturer.** Graduate School of Management, Politecnico di Milano, *International Master in Business Analysis and Big Data (BABD)*, "Cloud and Cognitive services" module (6 hours in 2022, 10 hours in 2023, 6 hours in 2024).

2018 – 2019
- **Laboratory Tutor.** Politecnico di Milano, M.Sc. Mathematical Engineering, *052496 – Algorithms and Parallel Computing* course (10 hours in 2018, 10 hours in 2019).

Theses and Ph.D. Students Supervision

2020 – present
- **Co-supervisor** of **15 master theses**: 10 for the M.Sc. Mathematical Engineering at Politecnico di Milano, 2 for the M.Sc. Computer Science Engineering at Politecnico di Milano, 3 for the M.Sc. Computer Science at University of Milano-Bicocca.

2024 – present
- **Co-supervisor** of **1 Ph.D. student** enrolled in the PhD program in Computer Science at University of Milano-Bicocca.

Education

2020 – 2024
- **Ph.D., Politecnico di Milano** Information Technology, Computer Science, **cum laude**.
  Thesis title: *Resource Allocation and Scheduling Problems in Computing Continua for Artificial Intelligence applications*; available at: https://hdl.handle.net/10589/216813.

2017 – 2020
- **M.Sc. Mathematical Engineering, Politecnico di Milano** in Computational Science and Engineering (110/110).
  Thesis title: *Job Scheduling and Optimal Capacity Allocation Problems for Deep Learning Training Jobs*; abstract available at: https://hdl.handle.net/10589/165216.

2013 – 2017
- **B.S. Mathematical Engineering, Politecnico di Milano**.

Research Publications and Patents

Journal Articles

1. A. W. Kambale, H. Sedghani, **F. Filippini**, V. Giacomo, and D. Ardagna, "Tabular Reinforcement Learning Methods for Artificial Intelligence Tasks Offloading in Smart Eye-Wears," *ACM Transactions on Autonomous and Adaptive Systems. (Scimago Q2)*, vol. to appear, pp. 1–30, 2025.
2. R. Sala, B. Guindani, E. Galimberti, **F. Filippini**, H. Sedghani, D. Ardagna, S. Risco, G. Moltó, and M. Caballer, "OSCAR-P and aMLLibrary: Profiling and predicting the performance of FaaS-based applications in computing continua," *Journal of Systems and Software. (Scimago Q1)*, vol. 221, p. 112 282, 2025, ISSN: 0164-1212. DOI: 10.1016/j.jss.2024.112282.
3. A. Tundo, **F. Filippini**, F. Regonesi, M. Ciavotta, and M. Savi, "Decentralized Edge Workload Forecasting With Gossip Learning," *IEEE Transactions on Network and Service Management. (Scimago Q1)*, pp. 1–16, 2025. DOI: 10.1109/TNSM.2025.3570450.
4. **F. Filippini**, J. Anselmi, D. Ardagna, and B. Gaujal, "A Stochastic Approach for Scheduling AI Training Jobs in GPU-Based Systems," *IEEE Transactions on Cloud Computing. (Scimago Q1)*, vol. 12, no. 01, pp. 53–69, Jan. 2024, ISSN: 2168-7161. DOI: 10.1109/TCC.2023.3336540.
5. H. Sedghani, **F. Filippini**, and D. Ardagna, "SPACE4AI-D: A Design-Time Tool for AI Applications Resource Selection in Computing Continua," *IEEE Transactions on Services Computing. (Scimago Q1)*, vol. 17, no. 6, pp. 4324–4339, 2024, (Appendix available: https://doi.org/10.1109/TSC.2024.3479935/mm1). DOI: 10.1109/TSC.2024.3479935.
6. **F. Filippini**, M. Lattuada, M. Ciavotta, A. Jahani, D. Ardagna, and E. Amaldi, "A Path Relinking Method for the Joint Online Scheduling and Capacity Allocation of DL Training Workloads in GPU as a Service Systems," *IEEE Transactions on Services Computing. (Scimago Q1)*, vol. 16, no. 3, pp. 1630–1646, 2023, (Appendix available: https://doi.org/10.1109/TSC.2022.3188440/mm1). DOI: 10.1109/TSC.2022.3188440.

Conference Proceedings

1. **F. Filippini**, L. Cavenaghi, N. Calmi, M. Savi, and M. Ciavotta, "ML-Based Performance Modeling in Edge FaaS Systems," in *Service-Oriented and Cloud Computing*, C. Pahl, A. Janes, T. Cerny, V. Lenarduzzi, and M. Esposito, Eds., ser. ESOCC 2025, **(Best Paper Award)**, Cham: Springer Nature Switzerland, 2025, pp. 112–127, ISBN: 978-3-031-84617-5. DOI: 10.1007/978-3-031-84617-5_10.
2. E. Petriglia, **F. Filippini**, M. Ciavotta, and M. Savi, "Multi-Agent Reinforcement Learning for Workload Distribution in FaaS-Edge Computing Systems," in *7th Workshop on Parallel AI and Systems for the Edge*, ser. PAISE '25, Springer, 2025.
3. R. Cavadini, H. Sedghani, **F. Filippini**, and D. Ardagna, "Runtime Management of Artificial Intelligence Applications Through Hierarchical Reinforcement Learning," in *17th EAI International Conference on Performance Evaluation Methodologies and Tools, to appear*, ser. EAI VALUETOOLS, 2024, pp. 1–22.
4. **F. Filippini**, N. Calmi, L. Cavenaghi, E. Petriglia, M. Savi, and M. Ciavotta, "Analysis and Evaluation of Load Management Strategies in a Decentralized FaaS Environment: A Simulation-Based Framework," in *Workshop on Serverless at the Edge*, ser. SEATED '24, Pisa, Italy, Association for Computing Machinery, 2024, ISBN: 9798400706479. DOI: 10.1145/3660319.3660329.
5. E. Petriglia, **F. Filippini**, G. Pracucci, M. Savi, and M. Ciavotta, "Comparing Actor-Critic and Neuroevolution Approaches for Traffic Offloading in FaaS-powered Edge Systems," in *Workshop on Serverless at the Edge*, ser. SEATED '24, Pisa, Italy, Association for Computing Machinery, 2024, ISBN: 9798400706479. DOI: 10.1145/3660319.3660331.
6. R. Sala, **F. Filippini**, D. Ardagna, D. Lezzi, F. Lordan, and P. Thiem, "Greening AI: A Framework for Energy-Aware Resource Allocation of ML Training Jobs with Performance Guarantees," in *International Conference on Advanced Information Networking and Applications*, Springer, 2024, pp. 110–121. DOI: 10.1007/978-3-031-57931-8_11.
7. **F. Filippini**, R. Cavadini, D. Ardagna, R. Lancellotti, G. Russo Russo, V. Cardellini, and A. Lo Presti, "FIGARO: reinForcement learnInG mAnagement acRoss computing cOntinua," in *Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing*, ser. UCC '23, Taormina (Messina), Italy, Association for Computing Machinery, 2023, ISBN: 9798400702341. DOI: 10.1145/3603166.3632565.

8. **F. Filippini**, B. Lublinsky, M. de Bayser, and D. Ardagna, "Performance Models for Distributed Deep Learning Training Jobs on Ray," in *2023 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, Los Alamitos, CA, USA: IEEE Computer Society, Sep. 2023, pp. 30–35. DOI: 10.1109/SEAA60479.2023.00014.
9. **F. Filippini**, H. Sedghani, and D. Ardagna, "SPACE4AI-R: Runtime Management Tool for AI Applications Component Placement and Resource Selection in Computing Continua," in *Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing*, ser. UCC '23, Taormina (Messina), Italy, Association for Computing Machinery, 2023, ISBN: 9798400702341. DOI: 10.1145/3603166.3632560.
10. E. Galimberti, B. Guindani, **F. Filippini**, H. Sedghani, D. Ardagna, G. Moltó, and M. Caballer, "OSCAR-P and AMLLibrary: Performance Profiling and Prediction of Computing Continua Applications," in *Companion of the 2023 ACM/SPEC International Conference on Performance Engineering*, ser. ICPE '23 Companion, Coimbra, Portugal: Association for Computing Machinery, 2023, pp. 139–146, ISBN: 9798400700729. DOI: 10.1145/3578245.3584941.
11. A. W. Kambale, H. Sedghani, **F. Filippini**, G. Verticale, and D. Ardagna, "Runtime Management of Artificial Intelligence Applications for Smart Eyewears," in *Proceedings of the IEEE/ACM 16th International Conference on Utility and Cloud Computing*, ser. UCC '23, Taormina (Messina), Italy, Association for Computing Machinery, 2023, ISBN: 9798400702341. DOI: 10.1145/3603166.3632562.
12. M. Precuzzi, **F. Filippini**, and D. Ardagna, "Scheduling Deep Learning Jobs Training in the Cloud: Comparing Multiple Approaches," in *Scheduling and Planning Allocation woRKshop, SPARK 2022, Singapore, Republic of Singapore, June 7-12, 2022*, 2022, pp. 1–9. URL: https://icaps22.icaps-conference.org/workshops/SPARK/papers/spark2022_paper_4.pdf.
13. **F. Filippini**, D. Ardagna, M. Lattuada, E. Amaldi, M. Riedl, K. Materka, P. Skrzypek, M. Ciavotta, F. Magugliani, and M. Cicala, "ANDREAS: Artificial intelligence traiNing scheDuler foR accElerAted resource clusterS," in *8th International Conference on Future Internet of Things and Cloud, FiCloud 2021, Rome, Italy, August 23-25, 2021*, IEEE, 2021, pp. 388–393. DOI: 10.1109/FiCloud49777.2021.00063.
14. H. Sedghani, **F. Filippini**, and D. Ardagna, "A Random Greedy based Design Time Tool for AI Applications Component Placement and Resource Selection in Computing Continua," in *IEEE International Conference on Edge Computing, EDGE 2021, Chicago, IL, USA, September 5-10, 2021*, IEEE, 2021, pp. 32–40. DOI: 10.1109/EDGE53862.2021.00014.
15. H. Sedghani, **F. Filippini**, and D. Ardagna, "A Randomized Greedy Method for AI Applications Component Placement and Resource Selection in Computing Continua," in *IEEE International Conference on Joint Cloud Computing, JCC 2021, Oxford, United Kingdom, August 23-26, 2021*, 2021, pp. 65–70. DOI: 10.1109/JCC53141.2021.00022.
16. **F. Filippini**, M. Lattuada, A. Jahani, M. Ciavotta, D. Ardagna, and E. Amaldi, "Hierarchical Scheduling in on-demand GPU-as-a-Service Systems," in *22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2020, Timisoara, Romania, September 1-4, 2020*, IEEE, 2020, pp. 125–132. DOI: 10.1109/SYNASC51798.2020.00030.

Books and Chapters

1. **F. Filippini**, "Resource Allocation and Scheduling Problems in Computing Continua for Artificial Intelligence Applications," in *Special Topics in Information Technology, PoliMI SpringerBriefs*, S. Garatti, Ed., Cham: Springer International Publishing, 2025 (to appear), pp. 1–10. DOI: 10.1007/978-3-031-80268-3_4.

Patents

1. D. Ardagna, **F. Filippini**, A. W. Kambale, T. Ongarello, F. Palermo, H. Sedghani, and D. Trojaniello, *Processing system for executing an artificial intelligence application*, Politecnico di Milano - EssilorLuxottica. Submitted to the European Patent Office.

Other Publications, resulted from statistical analyses performed in support to the Rheumatology Unit at the ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy:

1. M. Di Cicco, E. Verduci, M. Muscara, L. Belloli, A. Adinolfi, C. Casu, D. Filippini, **F. Filippini**, M. C. Gerardi, M. Schettino, *et al.*, "Virtual Rheumatology in the Real World: Use of Telemonitoring to Optimize the Management of Patients with Rheumatoid Arthritis," *ARTHRITIS & RHEUMATOLOGY*, vol. 76, no. 9, pp. 3953–3954, Sep. 2024, ISSN: 2326-5191.
2. M. Di Cicco, O. Epis, E. Verduci, C. Casu, G. Segatto, **F. Filippini**, C. Rossetti, and D. Filippini, "AB0724 THE VALUE OF [18F]FDG PET/CT FOR MONITORING AND TREATMENT DECISION GUIDANCE IN LARGE VESSELL VASCULITIS: IS IT TELLING TOO LITTLE, OR MAYBE TOO MUCH?" *Annals of the Rheumatic Diseases*, vol. 82, pp. 1567–1568, Jun. 2023, ISSN: 0003-4967. DOI: 10.1136/annrheumdis-2023-eular.4714.
3. M. Di Cicco, O. M. Epis, **F. Filippini**, L. Alvaro, L. Belloli, V. Campanella, C. Casu, M. Chevallard, M. Muscarà, E. Schito, *et al.*, "La Crioglobulinemia HCV-Correlata nell'era dell'eradicazione virale completa: qual è l'impatto degli agenti antivirali diretti (DAA) sulle manifestazioni extra-epatiche, e quale il ruolo della terapia sequenziale con Rituximab?" *REUMATISMO, Giornale ufficiale della Società Italiana di Reumatologia*, vol. 71, no. 4, p. 77, 2019.
4. M. Di Cicco, D. A. Filippini, **F. Filippini**, L. Belloli, V. Campanella, C. Casu, M. Chevallard, M. Muscarà, M. Romano, E. Schito, *et al.*, "SAT0460 HCV-RELATED MIXED CRYOGLOBULINAEMIA IN THE DIRECT-ANTIVIRAL AGENTS ERA: IS THERE ADVANTAGE IN SEQUENTIAL THERAPY WITH RITUXIMAB?" *Annals of the Rheumatic Diseases*, vol. 78, pp. 1319–1320, Jun. 2019, ISSN: 0003-4967. DOI: 10.1136/annrheumdis-2019-eular.7781.

Awards and Achievements

2025 **Best Paper Award** for "ML-Based Performance Modeling in Edge FaaS Systems" at the 11th European Conference on Service-Oriented and Cloud Computing (ESOCC 2025)

2024 **Springer Award** for the Ph.D. Dissertation: one among the 12 best theses of the year, selected among the theses defended in the Ph.D. IT programme (XXXVI cycle) at Politecnico di Milano. A short contribution is published in a volume of the PoliMi Springer Briefs (doi: 10.1007/978-3-031-80268-3_4).

"Premio Giovani Talenti" (Young Talents Award) organized by University of Milano-Bicocca, under the Patronage of Accademia Nazionale dei Lincei: **first prize** among competing research fellows and researchers (RTDA, RTDB, RTT, RU) below 37 y.o., in the disciplines *AREA 01 (Mathematics and informatics)* and *AREA 09 (Industrial and information engineering)*.

2021 **Best Project Award** among the third call for Value Chain Technology Transfer Projects of H2020 TETRAMAX. ANDREAS selected for its exceptional contribution and innovation in power and cost management of deep learning training workloads.

Projects

Research Projects

2021 – 2023 **AI-SPRINT H2020**. I worked on the development of design-time and runtime frameworks supporting the deployment and execution of Artificial Intelligence applications in the computing continuum, under performance, security and privacy constraints. Open-source tools:

- SPACE4AI-D: System PerformAnce and Cost Evaluation on Cloud for AI applications Design. A novel open-source tool to support design space exploration, component placement and resource selection of AI applications running on cloud-edge systems (https://github.com/ai-sprint-eu-project/space4ai-d).
- SPACE4AI-R: System PerformAnce and Cost Evaluation on Cloud for AI applications Runtime. A novel open-source tool to support the runtime management of AI applications running on cloud-edge systems (https://github.com/ai-sprint-eu-project/space4ai-r-optimizer).

Projects (continued)

2020 – 2021 **TETRAMAX ANDREAS**. I worked on the development of heuristic policies for the resource selection and scheduling of Deep Learning jobs on GPUs. The ANDREAS (Artificial intelligence traiNing scheDuler foR accElerAted resource clusterS) scheduler achieved TRL 8 and is currently commercialized (https://www.e4company.com/en/andreas-artificial-intelligence-training-scheduler-for-accelerated-resource-clusters/).

Other Projects (Conto Terzi)

2024 – present **Ottimizzazione dei percorsi di visita per l'ispezione di navi** (En. Optimization of inspection routes for ships), led by Martec Marine (https://www.martecmarine.it). I am working on the development of mathematical optimization models and intelligent heuristic solutions to minimize the inspection time of ship rooms, guaranteeing the full coverage or favoring rooms with highest priority under system constraints.

Services

2025 **Program Committee co-chair**, *FastContinuum* workshop, jointly sponsored by the IEEE Symposium on Quantum Software: Quantum Software Engineering (QSW), the IEEE International Conference on Cloud Computing (CLOUD), and the IEEE International Conference on Edge Computing & Communications (EDGE), part of the 2025 IEEE World Congress on SERVICES (https://services.conferences.computer.org/2025/fastcontinuum-2025/).

Program Committee co-chair First International Workshop on Agentic AI, LLMs and Inference in the Computing Continuum (AI-CC), held in conjunction with the 18th IEEE/ACM International Conference on Utility and Cloud Computing (UCC2025) (https://sites.google.com/view/ai-cc25/).

2022 – 2025 **Reviewer**: CCGrid'22, ICA3PP'22, ICDCS'22, IEEE ICWS'22, BDS'23, BigData'24, SoftwareX journal, IEEE Transactions on Network and Service Management, IEEE Transactions on Services Computing, Ad Hoc Networks journal, Cluster Computing journal, Journal of Cloud Computing, Future Generation Computer Systems journal.

2021 – 2023 **Publicity Chair**, Scalable Deep Learning over Parallel And Distributed Infrastructures (ScaDL) Workshop.

2020 – 2021 **Webmaster**, International Symposium on Computer Performance, Modeling, Measurements and Evaluation (IFIP WG 7.3 Performance).