Curriculum Vitae et Studiorum

Name	**Andrea Borghesi**
Date of Birth	January 24 1988
Citizenship	Italian
Affiliation	DISI, University of Bologna
Address	Viale del Risorgimento, 2, 40136 Bologna (BO) - Italy
Phone	+39 051 20 93274
E-mail	andrea.borghesi3@unibo.it
Web Page	https://www.unibo.it/sitoweb/andrea.borghesi3

Short Bio

Andrea Borghesi received the M.S. degree in Computer Engineering and the Ph.D. in Computer Engineering from the University of Bologna (Italy), in 2013 and 2017 respectively. He is currently an *Assistant Professor ("RTD-A")* at the Department of Computer Science and Engineering (DISI) of the University of Bologna. He received the *Italian National Scientific Qualification for the role of Associate Professor in Computer Engineering* (ASN 2018-2020 settore 09/H1 II Fascia, VI Quadrimestre).

His research interests are focused on **optimization techniques and Machine Learning approaches for complex systems and novel computer paradigms**, with a special attention towards power and energy related aspects and predictive and prescriptive maintenance. His expertise spans different areas, with the general target of minimizing power consumption and automatize and improve the maintenance of the target systems: optimization and implementation of novel scheduling policies for supercomputers; creation of Machine/Deep Learning predictive models for automated maintenance; design and deployment of fine-grained monitoring infrastructure for heterogeneous large computing machines; studying the interaction of methods from different Artificial Intelligence areas to solve problems over complex, real-world systems. A big focus of his research has been the exploration of areas at the intersection of different branches of Artificial Intelligence, such as combining Machine Learning, optimization and simulation models.

Contents

Position & Education

Record of Employment

- 03/2020 – *Present*
  Executive Scientific Representative (HPC scientific unit) at the Inter-departments Center for Artificial Intelligence at University of Bologna, Italy (ALMA-AI, in Italian: "Centro Interdipartimentale di Intelligenza Artificiale")

- 12/2019 – *Present*
  Assistant Professor (RTDA) at the Department of Information Engineering and Computer Science (DISI), University of Bologna, Italy

- 2018 – *2019*
  Postdoctoral fellow at the Department of Information Engineering and Computer Science (DISI), University of Bologna, Italy
  Supervisor: Prof. Michela Milano

- 2017 – 2018
  Research intern at the Department of Electrical, Electronic and Information Engineering (DEI), University of Bologna, Italy
  Supervisor: Prof. Luca Benini

- 02/2016 – 07/2016
  Research Fellow at University College of Cork (UCC), Ireland
  Supervisor: Prof. Barry O'Sullivan

- *June* 2013 – *December* 2013
  Research intern at the Department of Information Engineering and Computer Science (DISI), University of Bologna, Italy
  Supervisor: Prof. Michela Milano

Education

- **Ph.D. in Computer Science and Engineering**, University of Bologna, 2017
  Thesis Title: Power-aware Job Dispatching in High Performance Computing Systems
  Advisors: Prof. Michela Milano, Prof. Luca Benini

- **Master Degree in Computer Engineering**, University of Bologna, 2013
  Score: 110/110 cum laude
  Thesis Title: Integrazione di ottimizzazione e simulazioni per il piano energetico regionale dell'Emilia-Romagna (Eng: Integrating optimization and simulation approaches to devise the energy plan of the Emilia-Romagna Region)
  Advisor: Prof. Michela Milano
  Topic: Applications of Artificial Intelligence

Professional Activities

Contribution to National & International Academic Research Projects

International Projects:

- **HORIZON-CL4-2021-HUMAN-01-24-101070363-AEQUITAS: ASSESSMENT AND ENGINEERING OF EQUITABLE, UNBIASED, IMPARTIAL AND TRUSTWORTHY AI SYSTEMS**
  Period: 11/2022 – present
  Role: Risk Manager (Involvement: MEDIUM/HIGH)
  Description:
  AI-based decision support systems are increasingly deployed in industry, in the public and private sectors, and in policy-making. As our society is facing a dramatic increase in inequalities and intersectional discrimination, we need to prevent AI systems to amplify this phenomenon but rather mitigate it. To trust these systems, domain experts and stakeholders need to trust the decisions. How the principles underpinning fairness and trustworthiness, translate into technical, social, and lawful requirements in the AI system design is still an open question. Similarly we don't know how to test if a system is compliant with these principles and repair it in case it is not.

  AEQUITAS proposes the design of a controlled experimentation environment for developers and users to create controlled experiments for 1) assessing the bias in AI systems, e.g., identifying potential causes of bias in data, algorithms, and interpretation of results, 2) providing effective methods and engineering guidelines to repair, remove, and mitigate bias, 3) provide fairness-by-design guidelines, and software engineering techniques to design new bias-free systems.

- **H2020-ICT-2020-101017142-StairwAI: Stairway to AI: Ease the Engagement of Low-Tech users to the AI-on-Demand platform through AI**
  Period: 01/2021 – present
  Role: Technical Contributor (Involvement: HIGH)
  Website: https://cordis.europa.eu/project/id/101017142
  Description:
  The StairwAI project targets low-tech users with the goal of facilitating their engagement on the AI4EU on-demand Platform. This will be achieved through a new service layer enriching the functionalities of the on-demand platform and containing: (1) a multi-lingual interaction layer enabling conversations with the Platform in the user's own language, (2) a horizontal matchmaking service for the automatic discovery of AI assets (tools, data sets, AI experts, consultants, papers, courses etc.) meeting the user business needs and, (3) a vertical matchmaking service that will dimension and provision hardware resources through a proper hardware provider (HPC, Cloud and Edge infrastructures).

  Andrea Borghesi is directly involved in different areas of the project, with a special focus on the work package 5 and 6, relative to the horizontal and vertical matchmaking; in particular, he is leading T6.3. In this area, his contributions revolves around the combination of Machine Learning and optimization to guide the matching process, e.g. by integrating data-driven models in constraint programming models for hardware dimensioning and algorithm configuration. He also contributed to writing the project proposal.

- **H2020-ICT-2019-952215-TAILOR: Foundations of Trustworthy AI - Integrating Reasoning, Learning and Optimization**
  Period: 09/2020 – present
  Role: Technical Contributor (Involvement: MEDIUM)
  Website: https://tailor-network.eu/
  Description:
  Maximising opportunities and minimising risks associated with artificial intelligence (AI) requires a focus on human-centred trustworthy AI; this can be achieved by collaborations between research excellence centres Currently, this work is carried out by an isolated scientific community where research groups are working individually or in smaller networks; TAILOR aims to bring these groups together in a single scientific network on the Foundations of Trustworthy AI.

  Andrea Borghesi research concerns the integration of two AI paradigms, namely Machine Learning and optimization, for instance by embedding empirical models (derived from data) into combinatorial optimization models. In the context of TAILOR, the research activity concerns the development of methods to integrate these paradigms with a special focus towards trustworthy AI (especially Task 4.2).

- **H2020-ICT-2018-3-857191-IoTwins: Distributed Digital Twins for industrial SMEs: a big-data platform**
  Period: 09/2019 – present
  Role: WP Leader (Involvement: HIGH)
  Website: https://www.iotwins.eu/
  Description:
  The concept of digital twins has been around but the Internet of Things has enabled its cost-effective implementation. Digital twins refer to a virtual representation of a physical product or process. The EU-funded IoTwins project plans to build testbeds for digital twins in the manufacturing and facility management sectors. The digital models will integrate data from various sources such as data APIs, historical data, embedded sensors and open data.

  Andrea Borghesi is involved in IoTwins as leader of the work package 3. His research activity focuses on Machine Learning services for digital twins (both in large data centers and the manufacturing area), with the task of collecting requirements, developing and integrating ML methods and approaches, deploying them and evaluating their implementation – within work package 3. He is also involved in the development of a vertical, holistic, monitoring infrastructure to collect and store heterogeneous data from large data centers (big data storing and processing), in addition to analysing the collected data and developing predictive models for optimized maintenance.

- **H2020-FETPROACT-732631-OPRECOMP: Open transprecision computing**
  Period: 02/2019 – present
  Role: Technical Contributor (Involvement: MEDIUM/HIGH)
  Website: http://oprecomp.eu/
  Description:
  This project explores a novel computing paradigm (transprecision computing) which aims at improving the energy efficiency of computing systems by reducing the precision of single variables and intermediate operations to provide the minimum accuracy required for the final results (in contrast with the "conservative" precision provided by standard type systems).

  The research activity of Andrea Borghesi is focused on approaches for the fine tuning of the precision

of Floating Point variables, exploiting techniques from the artificial intelligence area, from machine learning to linear/integer optimization (mathematical/constraint programming). He collaborates especially in work packages 6.1 and 6.2 and is part of the UNIBO representatives at official review meetings.

- **FP7-RI-312763-H2020-653838-PRACE-3IP: OpenPower cluster**
  Period: 01/2017 – 12/2018
  Role: Technical Contributor (Involvement: HIGH)
  Website: http://www.prace-ri.eu/prace-3ip/
  Description:
  PRACE (Partnership for Advanced Computing in Europe) started the process for a pre-commercial procurement on Whole-System Design for Energy Efficient High Performance Computing (HPC) in 2012 with a workshop on the usage of Pre-Commercial Procurement (PCP) in HPC, to obtain R&D services which should result in future PRACE HPC systems to become more energy efficient, especially exploiting OpenPower technologies.

  Andrea Borghesi participated to the design of a novel High Performance Computing system, the D.A.V.I.D.E. supercomputer, in collaboration with industrial and academic partners. His activity was mainly focused on: 1) developing real-time job dispatching policies to limit the system power consumption under a given threshold (*power capping*); 2) building an extensive data collection infrastructure to gather performance counters and job information on a heterogeneous supercomputer with a very high granularity and with big data capabilities; 3) characterizing the HPC system behaviour through machine learning models built with the data collected from the monitoring infrastructure, in particular with the purpose of anomaly detection and power consumption prediction.

- **FP7-IDEAS-ERC-291125-MULTITHERMAN: Multi-Scale Thermal Management of Computing Systems**
  Period: 01/2014 – 03/2018
  Role: Technical Contributor (Involvement: HIGH)
  Website: http://projects.eees.dei.unibo.it/multitherman/
  Description:
  This project studies novel strategies for thermal planning/management which are alternative to worst-case design; the main aim is to pave the way towards methodologies enabling lower power/energy consumption and thermal optimization. The development of a synergistic performance, power and thermal management strategy requires major breakthrough in several areas, namely architectures, run-time systems, resource management middleware, code optimization tools and programming models.

  The major contributions of Andrea Borghesi in the context of the project fall in three areas: 1) the development of novel scheduling policies aimed at curtailing power consumption in large-scale systems and at improving system performance (w.r.t. both users and owners/administrators satisfaction), ranging from the algorithm design to the actual implementation/deployment of the proposed solution; 2) the creation of thermal/power models of complex systems (HPC machines and data centers) to be used for predictive tasks (mainly machine learning models trained in supervised fashion); 3) the creation of data collection frameworks to gather information from heterogeneous sources (for instance, in the context of HPC machines, measures coming form the computing nodes and related components).

- **FP7-ICT-2011-7-288147-ePolicy: Engineering the POlicy-making LIfe CYcle**

Period: 01/2013 – 09/2014
Role: Technical Contributor (Involvement: LOW/MEDIUM)
Website: http://www.epolicy-project.eu/
Description:
The project main aim is to support policy makers in their decision process across a multi-disciplinary effort aimed at engineering the policy making life-cycle. For the first time, global and individual perspectives on the decision process are merged and integrated. The project focuses on regional planning and promotes the assessment of economic, social and environmental impacts during the policy making process (at both the global and individual levels)

Andrea Borghesi followed two main research avenues: 1) understanding the impact of public economic incentives on the decision process of private citizens using agent-based models, with an explicit focus on the adoption of photovoltaic panels in the Italian Emilia-Romagna region; 2) devising novel incentive schemes to foster the fulfillment of macro-level objectives acting only on micro-level stakeholders (through optimization techniques).

National Projects:

- **CINECA-ISCRAS-C-HP10CVV2VY-vertOpt: Vertical Match Making of Artificial Intelligence Applications**
  Period: 09/2021 – 06/2022
  Role: Principal Investigator (Involvement: HIGH)
  Website: http://www.hpc.cineca.it/services/iscra/
  Description:
  This project main objective is the creation of a data set which will be used to develop an optimization engine for vertical matchmaking of Artificial Intelligence algorithms. Given a set of requirements from the end users and a problem to be solved vertical matchmaking consists in finding the appropriate set of hardware (HW) resources resources and algorithm (together with its optimal configuration) for the target task.

- **CINECA-ISCRAS-C-HP10CVDID8-V-Match: Vertical Match Making of Artificial Intelligence Applications**
  Period: 07/2020 – 04/2021
  Role: Principal Investigator (Involvement: HIGH)
  Website: http://www.hpc.cineca.it/services/iscra/
  Description:
  The project main goal is the design and development of a vertical matchmaking service for Artificial Intelligence applications. Given a set of requirements from the end users – time constraints, costs, privacy, etc. – vertical matchmaking consists in finding the appropriate set of hardware (HW) resources resources that satisfy a given algorithm and tool. The aim of the proposed activity is to perform a preliminary exploration of the vertical matchmaking.

- **CINECA-ISCRAS-C-HP10C0V1Z5-EXADATA: Exascale Monitoring and Big Data Analytics**
  Period: 10/2019 – 07/2021
  Role: Principal Investigator (Involvement: HIGH)
  Website: http://www.hpc.cineca.it/services/iscra/
  Description:

The project main aim is the extraction and analysis of the wealth of data generated by a monitoring infrastructure in a supercomputer, with the goal of improved system performance and maintenance and towards data center automation. The target HPC system is the Marconi supercomputer, hosted by the CINECA consortium and monitored via a holistic infrastructure called Examon. The objective of the research activity were threefold: 1) develop of the digital twin of a supercomputer in-production; 2) devise a data-drive, approach for anomaly detection in a HPC system; 3) create power/thermal/energy/performance predictive models.

Research Projects in Collaboration with Private Institutions and Companies :

- **Job-scheduling optimization for High-Performance Computing systems, commissioned research by ENI**
  Period: Set 2022 - March 2023
  Role: Principal Investigator (Involvement: HIGH)
  Description:
  The research consists in collaborating with ENI system administrators and another consulting company (DataReply) for the development of optimized algorithms for job scheduling and allocation in High-Performance Computing (HPC) systems. ENI has a great interest in this domain as they own the largest HPC systems in Italy. The activity focuses on creating the optimization model for the HPC job dispatching solution (based on Quadratic Unconstrained Binary Optimization) and assisting in the creation of a prototype to be adopted in ENI's machines.

- **Advanced diagnostic in complex industrial systems, with YANMAR Italy S.P.A.** *Period*: Jun 2019 - Feb 2020
  Role: Scientific Responsible (Involvement: HIGH)
  Description:
  The research consisted consisted in the development of a diagnostic framework for power generation and electric/thermal (both hot and cold) energy consumption device, monitored through suitable sensors. These devices can undergo unexpected changes in their behaviour. The goals of the activity concerned 1) the definition of general (e.g. non device-specific) Machine Learning (ML) models to support the diagnosis of anomalies and faulty states and 2) the experimental evaluation of such models.

International Collaborations:

- **Energy Efficient High Performance Computing Group (EEHPCWG)**
  Period: 06/2016 – present
  Role: Member *Website*: https://eehpcwg.llnl.gov/
  Description:
  The work group mission is to encourage implementation of energy conservation measures, energy efficient design in high performance computing (HPC), and share ideas.

  Andrea Borghesi participated to the group meetings and discussion, collaborating in the writing of white papers, international conference papers, and technical reports.

Program Chair & Organization Committees

Andrea Borghesi has been **chair/co-chair** of:

- HiPEAC 2020, the International Conference on High Performance and Embedded Architectures and Compilers, Jan 2020, Bologna, Italy

Andrea Borghesi has served as **guest editor** for:

- Applied Sciences Journal, ISSN 2076-3417, (MDPI), special issue "Machine Learning for Industry 4.0: From Manufacturing and Embedded Systems to Cloud Computing and Data Centers", May 2020 - Feb 2021 (https://www.mdpi.com/journal/applsci/special_issues/Machine_Learning_Industry_4)

Andrea Borghesi has been a member of the **program committee** of:

- International Joint Conference on Artificial Intelligence (IJCAI), since 2017
- European Conference on Artificial Intelligence (ECAI), since 2019
- Association for the Advancement of Artificial Intelligence Conference (AAAI), since 2018
- International Conference on Principles and Practice of Constraint Programming (CP), since 2018
- ACM International Conference on Computing Frontiers (CF), since 2021
- European Congress Embedded Real Time Systems (ERTS), since 2019

Andrea Borghesi has been a member of the following **international research collaborations** of:

- Energy Efficient High Performance Computing Working Group, since 2017

Presentations at Workshops & Conferences

Andrea Borghesi has been **invited speaker** at the following workshops and conferences (national and international):

- Openday IoTwins, workshop organized by CINECA, title of the intervention “Examon e la manutenzione predittiva in ambito HPC”, Jun 28 2022, Bologna (Italy)
- Data Week 2022, Big Data Value Association (BDVA), title of the intervention “IoTwins - Big Data Platform for optimized and replicable Industrial and Facility Management models ”, May 26 2022, Online
- Workshop “Evento Gemelli Digitali - Applicazioni Pratiche e Case History” (English translation: “Digital twins Event - Practical Applications and Case History”), May 25 2022, Bologna (Italy)
- CINECA School “Digital twins per l’industria” (English translation: “Digital twins for the industry”), Dec 14 2021, Bologna (Italy)
- European Big Data Value Forum (EBDV), title of the intervention “IoTwins - Distributed Digital Twins for industrial SMEs: a big-data platform”, Nov 29 2021, Ljubljana (Slovenia) & Online
- Workshop “Intelligenza Artificiale per il settore Manifatturiero” (English translation: “Artificial Intelligence for the manufacturing sector”) hosted at MECSPE-2019, Mar 30 2019, Parma (Italy)
- CINECA Workshop “HPC for Industry 4.0”, May 21-23 2019, Milan (Italy)

- PRACE School “Energy Efficiency in HPC”, organized at CINECA, Bologna (Italy), Feb 22 2019
- Workshop “Presentation of D.A.V.I.D.E. supercomputer”, organized at CINECA, Bologna (Italy), Nov 29 2018
- Workshop “Deep Learning and Machine Learning for HPC”, organized at CINECA, Bologna (Italy), Oct 23 2018

Andrea Borghesi **presented his research** at the following workshops and conferences (national and international):

- 27th European Conference on Modeling and Simulation 2013 (ECMS13), May 28-31, 2013, Alesund (Norway)
- 19th International Conference on Principles and Practices of Constraint Programming (CP13), Sep 16-20, 2013, Uppsala (Sweden)
- XIII Conference of the Italian Association for Artificial Intelligence (AI*IA2013), Dec 4-6, 2013, Turin (Italy)
- 20th International Conference on Principles and Practices of Constraint Programming (CP14), Sep 8-12, 2014, Lyon (France)
- International Conference on Renewable Energy Research and Application (ICRERA2014), Oct 18-22, 2014, Milwaukee (US)
- International Conference on High Performance Computing and Simulation (HPCS15), Jul 19-23, 2015, Amsterdam (Netherlands)
- 21st International Conference on Principles and Practices of Constraint Programming (CP15), Aug 31 - Sep 4, 2015, Cork (Ireland)
- XIV Conference of the Italian Association for Artificial Intelligence (AI*IA2015), Sep 24-26, 2015, Ferrara (Italy)
- ISC High Performance (ISC16), Jun 18-22, 2016, Frankfurt (Germany)
- ACM International Conference on Computing Frontiers (CF18), May 8-10, 2018, Ischia (Italy)
- CINECA Workshop “Deep Learning and Machine Learning for HPC”, Oct 23, 2018, Bologna (Italy)
- Thirsty-First Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-19), Jan 27 - Feb 1, 2019, Honolulu (US)
- 1st IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS2019), Mar 17-20, 2019, Hsinchu (Taiwan)
- CINECA workshop “Energy Efficiency in HPC”, Feb 22, 2019, Bologna (Italy)
- XVIII Conference of the Italian Association for Artificial Intelligence (AI*IA2019), Nov 19-23, 2019, Rende (Italy)
- ACM International Conference on Computing Frontiers (CF20), May 11-13, 2020, Catania (Italy) – The conference took place remotely due to Covid19 pandemic emergency

Referee Services in Journals & Conferences

Andrea Borghesi is a **reviewer** for the following journals (J) and conferences (C):

- (J) IEEE Transactions on Dependable and Secure Computing (TDSC)
- (J) IEEE Transactions on Parallel and Distributed Systems (TPDS)
- (J) Elsevier Environmental Science & Policy
- (J) Springer Wireless Networks
- (J) Springer Constraints
- (J) Intelligenza Artificiale
- (C) IJCAI International Joint Conference on Artificial Intelligence
- (C) AAAI Conference on Artificial Intelligence
- (C) ECAI European Conference on Artificial Intelligence
- (C) International Conference on Principles and Practice of Constraint Programming (CP)
- (C) International Conference on the Integration of Constraint Programming, Artificial Intelligence, and Operations Research (CPAIOR)
- (C) European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)
- (C) IEEE/ACM International Symposium in Cluster, Cloud, and Grid Computing (CCGrid)
- (C) European Congress Embedded Real Time Systems (ERTS)

Prizes and Awards

Andrea Borghesi has received the following prizes & awards:

- Winner of GAUSS AWARD 2016 (Research Paper Award), obtained with the paper “Predictive Modeling for Job Power Consumption in HPC Systems”
- Winner of the prize "IJCAI 2019 – Distinguished Program Committee Member"

Teaching Activities

In the academia (bachelor & master degree)

2021 – *present*
Adjunct Professor
Master degree in Engineering Management, University of Bologna, Italy
Metodi e Modelli di Data Analytics, 3CFUs (30 hours)
Lectures, exercises, final project supervision
Academic years: 2021-2022 (Spring), 2022-2023 (Fall)

2020 – *present*
Adjunct Professor
Bachelor degree Computer Engineering, University of Bologna, Italy
Fondamenti di Informatica, 6CFUs (60 hours)
Lectures, exercises, final project supervision
Academic years: 2020-2021 (Fall), 2021-2022 (Fall), 2022-2023 (Fall)

2017 – 2020
Adjunct Professor
Bachelor degree Automation/Electrical Engineering, University of Bologna, Italy
Fondamenti di Informatica, 3CFUs (30 hours)
Lectures, exercises, final project supervision
Academic years: 2017-2018 (Spring), 2018-2019 (Spring), 2019-2020 (Spring)

In the academia (PhD courses)

2020 – *present*
Course title: From embedded devices to the cloud: a vertical perspective on IoT systems (20 hours)
PhD in Data science and computation, University of Bologna, Italy
Courses held: June 2020, July 2022

2021 – *present*
Course title: From embedded devices to the cloud: a vertical perspective on IoT systems (20 hours)
PhD in Computer Science and Engineering, University of Bologna, Italy
Courses held: July 2021

Others

- Course held at OPTIT SRL (Cesena, Italy), titled "Advanced techniques and applications of Deep Learning", 20 hours, April 2021
- Course held at HPE-COXA (Modena Italy), titled "ML and AI for predictive analytics", 16 hours, 8 and 16 October 2020
- Course held at OPTIT SRL (Cesena, Italy), titled "Introduction to Machine and Deep Learning", 4 hours, 29 November 2019

In the academia (Teaching Assistant/Tutor/Mentor)

2014 – *present*
Teaching Assistant
Master degree Computer/Electronics Engineering, University of Bologna, Italy
Intelligent Systems
Seminars, exercises, final project supervision

2013 – 2017
Teaching Assistant
Bachelor degree Automation/Electrical Engineering, University of Bologna, Italy
Fondamenti di Informatica
Lectures, exercises, final project supervision

2019 – *present*
Teaching Assistant
Master degree Electronics Engineering, University of Bologna, Italy
Lab of Big Data Architectures M
Lectures, projects, final project supervision

2017 – 2018
Teaching Assistant
Master degree Advanced Automotive Engineering, University of Bologna, Italy
Algorithms and Systems for Big Data Processing
Lectures

Student Supervision Activity

Andrea Borghesi has supervised the following **post-graduate students and temporary research fellows**:

- Carmine Di Santi, temporary research fellow (AdR) May 2022 - ongoing, research area: "Development and Deployment of AI-based Models for Industrial Digital Twins ", University of Bologna
- Matteo Francobaldi, temporary research fellow (AdR) Jan 2022 - ongoing, research area: "Vertical optimization methods for Hardware dimensioning and AI algorithms configuration", University of Bologna
- Andrea Boscarino, temporary research fellow Nov (AdR) 2020 - Dec 2021, research area: "Learning models and optimization algorithms for static variable precision tuning in the transprecision computing area", University of Bologna
- Marco Di Felice, post-graduate research scholarship Jun 2019 - Feb 2020, research area: "Advanced diagnostic in complex industrial systems", in collaboration with YANMAR Italy, University of Bologna

Andrea Borghesi has supervised (or co-supervised) the following students during their **final graduation project** (bachelor/master thesis):

- Marco Pazzaglia, master degree in Computer Engineering, topic “Intelligent Systems”, 2021/2022, University of Bologna
- Carmine De Santi, master degree in Computer Engineering, topic “Intelligent Systems”, 2020/2021, University of Bologna
- Daniele Verì, master degree in Artificial Intelligence, topic “AI for Industry”, 2020/2021 (main advisor: prof. Michela Lombardi), University of Bologna
- Silvia Brescia, master degree in Computer Engineering, topic “Intelligent Systems”, 2020/2021 (main advisor: prof. Michela Milano), University of Bologna
- Emanuele Di Giacomo, master degree in Computer Engineering, topic “Intelligent Systems”, 2019/2020 (main advisor: prof. Michela Milano), University of Bologna
- Michela Minerva, master degree in Computer Engineering, topic “Intelligent Systems”, 2019/2020 (main advisor: prof. Michela Milano), University of Bologna
- Andrea Piretti, master degree in Computer Engineering, topic “Intelligent Systems”, 2019/2020 (main advisor: prof. Michela Milano), University of Bologna
- Matteo Berti, master degree in Computer Science, 2019/2020 (main advisor: prof. Ozalp Babaoglu), University of Bologna
- Federico Cocchi, bachelor degree in Automation Engineering, topic “Fondamenti di Informatica”, 2019/2020 (main advisor: prof. Michela Milano), University of Bologna
- Andrea Boscarino, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Lorenzo Pizzigati, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Federico Livi, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Kevin Leto, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Giovanni Baratta, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Enzo Pio Palmisano, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Alberto Bambini, master degree in Computer Engineering, topic “Intelligent Systems”, 2018/2019 (main advisor: prof. Michela Milano), University of Bologna
- Marco Di Felice, master degree in Computer Engineering, topic “Intelligent Systems”, 2017/2018 (main advisor: prof. Michela Milano), University of Bologna
- Isabella Viarchi, master degree in Computer Engineering, topic “Intelligent Systems”, 2017/2018 (main advisor: prof. Michela Milano), University of Bologna

- Francesco Baruzzi, bachelor degree in Automation Engineering, topic "Fondamenti di Informatica", 2017/2018 (main advisor: prof. Michela Milano), University of Bologna
- Valerio Iachini, master degree in Computer Engineering, topic "Sistemi Intelligenti", 2013/2014 (main advisor: prof. Michela Milano), University of Bologna
- Francesca Collina, master degree in Computer Engineering, topic "Sistemi Intelligenti", 2013/2014 (main advisor: prof. Michela Milano), University of Bologna

Andrea Borghesi has supervised the following students during their **internships**:

- Massimo Schembri, September-November 2020, during his master degree in Computer Engineering, University of Bologna
- Federico Cocchi, February-June 2020, during his bachelor degree in Automation Engineering, University of Bologna

Andrea Borghesi has supervised the following students during their **"Summer of Code HPC" activity**, performed in conjunction between University of Bologna and CINECA consortium:

- Martin Molan, July-August 2019
- Petr Stehlik, July-August 2017

Research Interests

My main research interests cover different areas from the field of Artificial Intelligence (AI), prevalently Machine Learning (ML) and optimization and simulation with Agent-Based Models (ABM), with a focus on applied solutions and especially in the areas of management of large computing systems and power/energy aware topics. A core research direction that I have explored consists in the exploration of the areas at the intersection of multiple branches of AI research, with a particular interest towards mingling the sub-symbolic layers with higher-level, human-like, reasoning-based approaches. For instance, I have tackled the challenge of integrating data-driven models for predictive and classification tasks, such as ML models, into optimization and simulation models, striving to combine the benefits of the two paradigms; in the opposite direction, I have also studied the impact of using constraints derived from the optimization area in order to boost the performance of ML models (e.g., [C21]).

Hardware dimensioning The problem of determining what is the right hardware (HW on premise or on the cloud) architecture and its dimensioning for AI algorithms is still crucial. Searching for the optimal solution is often challenging, as it is not trivial to anticipate the behaviour of an algorithm on diverse architectures. I tackled this problem proposing an automated decision support tool called HADA and described in [J8, C16]. I have exploited the Empirical Model Learning paradigm, based on the integration of ML models into an optimization problem. The key idea is to integrate domain knowledge held by experts with data-driven models that learn the relationships between HW requirements and AI algorithm performances. In particular, the approach starts with benchmarking multiple AI algorithms on different HW resources, generating data used to train ML models; then, optimization is used to find the best HW configuration that respects user-defined constraints (e.g., budget, time, solution quality).

Prescriptive and predictive maintenance in industrial contexts and large scale data centers Another research direction stemming from my interest in complex systems (such as HPC systems, data centers, factories, etc) explores the possibility to improve and automatize the maintenance of these systems, both from a predictive (e.g. forecast when a fault is going to happen) and a prescriptive (decide a set of actions to prevent/mitigate the fault) point of view, for example, as I have studied in the context of Industry 4.0[C10]. The key in this area is building accurate models capable of representing the behaviour of the complex systems. The first step to create useful models is having good data describing the system in its entirety. For this purpose, I have participated to the design and deployment of a supercomputer with an innovative, scalable, and fine-grained monitoring infrastructure, capable to recreate an extremely accurate image of the monitored system[C1, C4, C2] (in the context of the *PRACE-3IP* project). Thanks to the detailed data, I have then developed an approach for building models describing the state of the system. In large systems it is not easy to identify malfunctions in real time; as part of my research I have developed automated tools for anomaly detection, based on ML/DL models and using HPC machines as case study. I experimented mainly with two broad families of ML models: I) supervised ones[C28, J9], where the ML models need to be trained with a balanced data set containing examples belonging to all status-classes to detect (thus anomalies need to be injected on the system to obtain the training set[C27]); II) semi-supervised models, where the ML model learns the normal behaviour of a supercomputer and can then detect anomalies since they diverge from the normal distribution [C7, J2] – the last tool is also publicly available as a prototype [1]. As a follow-up of the semi-supervised approach, I also deployed an anomaly detection model on a real supercomputer to perform online anomaly detection in real time, with minimal overhead[C11]. Additionally, I have also dealt with the problem of *explaining* the output of fault detection models in the HPC domain[C23, C24].

Moving beyond pure anomaly detection, the natural follow-up is to study the possibility of *anticipating* the insurgence of faults, an information which could be used in predictive maintenance solutions. In this direction, I proposed a DL-based approach combining a semi-supervised model and a supervised one to anticipate the anomalies in a production tier-0 supercomputer [J6, C25, C26]. Recently, I have also collaborated to the creation of a suite of AI-based services for optimizing the management of supercomputer facility, in a series of tasks ranging from anomaly detection in computing nodes and failure predictions of hard drives to power and thermal models of the computing nodes and the server rooms; this works culminated in a tool called *ExaMon-X*[J4]. Again in the context of fault prediction but in a different industrial domain, namely failure prediction of ball bearings in an industrial machine producing bottle caps, I have devised and validated a DL model for remaining useful life estimation[C18].

Transprecision computing Transprecision computing is a novel computing paradigm aimed at designing systems to deliver just the required precision for intermediate computations rather than tolerating errors implied by imprecise HW or SW computations. Adopting this principle, the user-defined constraints on the precision of final results are always met but intermediate operations can be deployed to custom, lower precision compute units to save energy. Within the scope of the *OPRECOMP* project, I contributed with an exploration of approaches for mixed precision tuning of Floating Point (FP) variables, i.e. fine-tune the precision by specifying the number of bits to be used for the mantissa and the exponent of the FP variable (using a smaller number of bits decreases the precision, potentially saving energy)[C15]. In this context I developed a prototype tool for static precision tuning, *StaticFPTuner*[2], which combines ML models (to represent and learn the relation between variable precision and overall accuracy of the computation) and optimization via Mathematical Programming (to find the optimal precision given a

[1] https://github.com/AndreaBorghesi/anomaly_detection_HPC
[2] https://github.com/oprecomp/StaticFPTuner

target accuracy). In the context of transprecision computing I also studied the benefits attainable by injecting prior domain knowledge in a DL models, obtaining significant increase in prediction accuracy[C5].

Optimization and power-aware job dispatchers for High Performance Computing systems High Performance Computing (HPC) systems are large machines composed by thousands of heterogeneous components. Improving their performance and facilitating their maintenance has been a very interesting research challenge over the years. For instance, deciding the optimal scheduling order of submitted jobs is a very complex task that I have tackled with Constraint Programming and ad-hoc heuristic algorithms[C3, C17]. If we consider also the power aspect (the context of the *MULTITHERMAN* project) and we aim for a power-aware job dispatcher, the difficulty of the problem increases: first we need to develop predictive models capable of estimating a job power consumption before its execution[C6], then AI-inspired approaches can be combined with heuristic techniques to devise optimal scheduling policies[C9, C8, J1] while respecting tight time limits arising from the real-time requirements of job dispatching in HPC systems – as a corollary problem, I studied the optimal computing frequency assignment for a set of jobs already running in a supercomputer in the event of a sudden reduction in available power[C13]. Dealing with a related aspect, I've also been involved in the development of runtime framework for parallel applications capable of dynamically reduce the frequency of computing nodes in HPC systems during communication phases which are not impacted by the frequency reduction[J7]. An additional related issue I addressed is the design of pricing schemes to foster the adoption of power-aware strategies, schemes capable to benefit both system owners and users[J3]. Remaining in the area of energy efficiency for HPC and data center, along the years I have participated to the Energy Efficient High Performance Computing Working Group (EEHPCWG[3]), whose effort I have helped with surveys and detailed analysis documenting the state of the art and state of practice in both academia and industry in terms of energy efficient solutions[C22, C20].

Agent-Based Models for energy policies Designing and evaluating energy policies is a difficult challenge because the energy sector is a complex system that cannot be adequately understood without using models merging economic, social and individual perspectives. Policy makers need precise models to estimate the impact of their plans. In my research within the *ePolicy* project I focused on policies for promoting energy production from renewable energy sources and, in particular, on photovoltaic (PV) power generation in the Emilia-Romagna region[C14, C12]. In this context, I worked on ABMs devised merging self-reported behaviour (interviews to PV installers) and observed behaviour (the historical installed PV power trend). In this way the agent decision process is grounded on realistic assumptions and the historical data can be used to fine-tune the models, obtaining a method capable of predicting the PV power adoption [J5, C19].

Publication List

Journals

[J1] Andrea Borghesi, Andrea Bartolini, Michele Lombardi, Michela Milano, and Luca Benini. Scheduling-based power capping in high performance computing systems. *Sustainable Computing: Informatics and Systems*, 19:1–13, 2018.

[3] https://eehpcwg.llnl.gov/

[J2] Andrea Borghesi, Andrea Bartolini, Michele Lombardi, Michela Milano, and Luca Benini. A semisupervised autoencoder-based approach for anomaly detection in high performance computing systems. *Engineering Applications of Artificial Intelligence*, 85:634–644, 2019.

[J3] Andrea Borghesi, Andrea Bartolini, Michela Milano, and Luca Benini. Pricing schemes for energy-efficient hpc systems: Design and exploration. *The International Journal of High Performance Computing Applications*, 33(4):716–734, 2019.

[J4] Andrea Borghesi, Alessio Burrello, and Andrea Bartolini. Examon-x: a predictive maintenance framework for automatic monitoring in industrial iot systems. *IEEE Internet of Things Journal*, 2021.

[J5] Andrea Borghesi and Michela Milano. Merging observed and self-reported behaviour in agent-based simulation: A case study on photovoltaic adoption. *Applied Sciences*, 9(10):2098, 2019.

[J6] Andrea Borghesi, Martin Molan, Michela Milano, and Andrea Bartolini. Anomaly detection and anticipation in high performance computing systems. *IEEE Transactions on Parallel and Distributed Systems*, 2021.

[J7] Daniele Cesarini, Andrea Bartolini, Andrea Borghesi, Carlo Cavazzoni, Mathieu Luisier, and Luca Benini. Countdown slack: a run-time library to reduce energy footprint in large-scale mpi applications. *IEEE Transactions on Parallel and Distributed Systems*, 31(11):2696–2709, 2020.

[J8] Allegra De Filippo, Andrea Borghesi, Andrea Boscarino, and Michela Milano. Hada: An automated tool for hardware dimensioning of ai applications. *Knowledge-Based Systems*, page 109199, 2022.

[J9] Alessio Netti, Zeynep Kiziltan, Ozalp Babaoglu, Alina Sîrbu, Andrea Bartolini, and Andrea Borghesi. A machine learning approach to online fault classification in hpc systems. *Future Generation Computer Systems*, 2019.

Conferences & Workshops

[C1] Wissam Abu Ahmad, Andrea Bartolini, Francesco Beneventi, Luca Benini, Andrea Borghesi, Marco Cicala, Privato Forestieri, Cosimo Gianfreda, Daniele Gregori, Antonio Libri, et al. Design of an energy aware petaflops class high performance cluster based on power architecture. In *2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 964–973. IEEE, 2017.

[C2] Andrea Bartolini, Francesco Beneventi, Andrea Borghesi, Daniele Cesarini, Antonio Libri, Luca Benini, and Carlo Cavazzoni. Paving the way toward energy-aware and automated datacentre. In *Proceedings of the 48th International Conference on Parallel Processing: Workshops*, page 8. ACM, 2019.

[C3] Andrea Bartolini, Andrea Borghesi, Thomas Bridi, Michele Lombardi, and Michela Milano. Proactive workload dispatching on the eurora supercomputer. In *International Conference on Principles and Practice of Constraint Programming*, pages 765–780. Springer, 2014.

[C4] Andrea Bartolini, Andrea Borghesi, Antonio Libri, Francesco Beneventi, Daniele Gregori, Simone Tinti, Cosimo Gianfreda, and Piero Altoè. The davide big-data-powered fine-grain power and performance monitoring support. In *Proceedings of the 15th ACM International Conference on Computing Frontiers*, pages 303–308. ACM, 2018.

[C5] Andrea Borghesi, Federico Baldo, Michele Lombardi, and Michela Milano. Injective domain knowledge in neural networks for transprecision computing. In *International Conference on Machine Learning, Optimization, and Data Science*, pages 587–600. Springer, 2020.

[C6] Andrea Borghesi, Andrea Bartolini, Michele Lombardi, Michela Milano, and Luca Benini. Predictive modeling for job power consumption in hpc systems. In *International conference on high performance computing*, pages 181–199. Springer, 2016.

[C7] Andrea Borghesi, Andrea Bartolini, Michele Lombardi, Michela Milano, and Luca Benini. Anomaly detection using autoencoders in high performance computing systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9428–9433, 2019.

[C8] Andrea Borghesi, Francesca Collina, Michele Lombardi, Michela Milano, and Luca Benini. Power capping in high performance computing systems. In *International Conference on Principles and Practice of Constraint Programming*, pages 524–540. Springer, 2015.

[C9] Andrea Borghesi, Christian Conficoni, Michele Lombardi, and Andrea Bartolini. Ms3: A mediterranean-stile job scheduler for supercomputers-do less when it's too hot! In *2015 International Conference on High Performance Computing & Simulation (HPCS)*, pages 88–95. IEEE, 2015.

[C10] Andrea Borghesi, Giuseppe Di Modica, Paolo Bellavista, Varun Gowtham, Alexander Willner, Daniel Nehls, Florian Kintzler, Stephan Cejka, Simone Rossi Tisbeni, Alessandro Costantini, et al. Iotwins: Design and implementation of a platform for the management of digital twins in industrial scenarios. In *2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pages 625–633. IEEE, 2021.

[C11] Andrea Borghesi, Antonio Libri, Luca Benini, and Andrea Bartolini. Online anomaly detection in hpc systems. In *Proceedings of the 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS 2019)*, pages 229–233. IEEE, 2019.

[C12] Andrea Borghesi and Michela Milano. Multi-agent simulator of incentive influence on pv adoption. In *2014 International Conference on Renewable Energy Research and Application (ICRERA)*, pages 556–560. IEEE, 2014.

[C13] Andrea Borghesi, Michela Milano, and Luca Benini. Frequency assignment in high performance computing systems. In *International Conference of the Italian Association for Artificial Intelligence*, pages 151–164. Springer, 2019.

[C14] Andrea Borghesi, Michela Milano, Marco Gavanelli, and Tony Woods. Simulation of incentive mechanisms for renewable energy policies. In *ECMS*, pages 32–38, 2013.

[C15] Andrea Borghesi, Giuseppe Tagliavini, Michele Lombardi, Luca Benini, and Michela Milano. Combining learning and optimization for transprecision computing. In *Proceedings of the 17th ACM International Conference on Computing Frontiers*, pages 10–18, 2020.

[C16] Allegra De Filippo and Andrea Borghesi. Constrained hardware dimensioning for ai algorithms. In *PAIS 2022*, pages 145–148. IOS Press, 2022.

[C17] Cristian Galleguillos, Alina Sîrbu, Zeynep Kiziltan, Ozalp Babaoglu, Andrea Borghesi, and Thomas Bridi. Data-driven job dispatching in hpc systems. In *International Workshop on Machine Learning, Optimization, and Big Data*, pages 449–461. Springer, 2017.

[C18] Bassem Hichri, Anass Driate, Andrea Borghesi, and Francesco Giovannini. Predictive maintenance based on machine learning model. In Ilias Maglogiannis, Lazaros Iliadis, John Macintyre, and Paulo Cortez, editors, *Artificial Intelligence Applications and Innovations*, pages 250–261, Cham, 2022. Springer International Publishing.

[C19] Valerio Iachini, Andrea Borghesi, and Michela Milano. Agent based simulation of incentive mechanisms on photovoltaic adoption. In *Congress of the Italian Association for Artificial Intelligence*, pages 136–148. Springer, 2015.

[C20] Gregory A Koenig, Matthias Maiterth, Siddhartha Jana, Natalie Bates, Kevin Pedretti, Milos Puzovic, Andrea Borghesi, Andrea Bartolini, and Dave Montoya. Energy and power aware job scheduling and resource management: Global survey—an in-depth analysis. In *The 2nd International Industry/University Workshop on Data-center Automation, Analytics, and Control (DAAC'18), in conjunction with Supercomputing'18*, pages 2018–11.

[C21] M. Lombardi, F. Baldo, A. Borghesi, and M. Milano. An analysis of regularized approaches for constrained machine learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12641 LNAI:112–119, 2021.

[C22] Matthias Maiterth, Gregory Koenig, Kevin Pedretti, Siddhartha Jana, Natalie Bates, Andrea Borghesi, Dave Montoya, Andrea Bartolini, and Milos Puzovic. Energy and power aware job scheduling and resource management: Global survey—initial analysis. In *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 685–693. IEEE, 2018.

[C23] Martin Molan, Andrea Borghesi, Francesco Beneventi, Massimiliano Guarrasi, and Andrea Bartolini. An explainable model for fault detection in hpc systems. In *International Conference on High Performance Computing*, pages 378–391. Springer, 2021.

[C24] Martin Molan, Andrea Borghesi, Luca Benini, and Andrea Bartolini. Analysing supercomputer nodes behaviour with the latent representation of deep learning models. In *European Conference on Parallel Processing*, pages 171–185. Springer, 2022.

[C25] Martin Molan, Andrea Borghesi, Luca Benini, and Andrea Bartolini. Semi-supervised anomaly detection on a tier-0 hpc system. In *Proceedings of the 19th ACM International Conference on Computing Frontiers*, CF '22, page 203–204, New York, NY, USA, 2022. Association for Computing Machinery.

[C26] Martin Molan, Andrea Borghesi, Luca Benini, and Andrea Bartolini. Semi-supervised anomaly detection on a tier-0 hpc system. In *Proceedings of the 19th ACM International Conference on Computing Frontiers*, pages 203–204, 2022.

[C27] Alessio Netti, Zeynep Kiziltan, Ozalp Babaoglu, Alina Sîrbu, Andrea Bartolini, and Andrea Borghesi. Finj: A fault injection tool for hpc systems. In *European Conference on Parallel Processing*, pages 800–812. Springer, 2018.

[C28] Alessio Netti, Zeynep Kiziltan, Ozalp Babaoglu, Alina Sîrbu, Andrea Bartolini, and Andrea Borghesi. Online fault classification in hpc systems through machine learning. In *European Conference on Parallel Processing*, pages 3–16. Springer, 2019.

[C29] Stefano Teso, Laurens Bliek, Andrea Borghesi, Michele Lombardi, Neil Yorke-Smith, Tias Guns, and Andrea Passerini. Machine learning for combinatorial optimisation of partially-specified problems: Regret minimisation as a unifying lens. *arXiv preprint arXiv:2205.10157*, 2022.