

Leonardo Pellegrina

October, 2025

Department of Information Engineering
University of Padova
Via Gradenigo 6a, Padova, Italy, IT 35131
leonardo.pellegrina@unipd.it
<http://www.dei.unipd.it/~pelleagri>

SHORT BIO

Leonardo Pellegrina is a tenure-track Assistant Professor at the Department of Information Engineering of the University of Padova. His research activities focus on efficient and statistically sound Data Mining algorithms for pattern discovery from large data and large graphs, with applications to computational biology. He received an Honorable mention for the 2021 SIGKDD Dissertation Award (worldwide top-3 Ph.D. thesis in Data Mining & Knowledge Discovery), and was selected as one of the Best Program Committee members at ACM The Web Conference 2022 and 2023 (assigned to top 5% PC members). During his Ph.D. he visited the department of Computer Science of Brown University (Providence, RI, USA) as a Visiting Research Fellow. He has presented his research work in several international conferences and workshops (e.g., KDD, VLDB, RECOMB, ISMB, ECCB), and presented tutorials at ACM KDD'19 and SDM'21.

CURRENT POSITION

Assistant Professor (RTT), Department of Information Engineering, University of Padova.
from 3rd March 2025

PREVIOUS ACADEMIC APPOINTMENTS

Guest Researcher, Division of Theoretical Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden, hosted by Prof. Aristides Gionis. June 2025

Assistant Professor (RTDa), Department of Information Engineering, University of Padova. MIUR PNRR CN1 SPOKE 1, Future HPC & Big Data project CN00000013.
2nd March 2023 - 2nd March 2025

Postdoctoral Researcher (Assegno Senior tipo B), Department of Information Engineering, University of Padova. Competitive research grant. PI of the project *Efficient Algorithms to Identify Significant Patterns in Tumors*.
1st December 2022 - 1st March 2023

Postdoctoral Researcher (Assegno Junior tipo B), Department of Information Engineering, University of Padova. Competitive research grant. PI of the project *Rigorous and Efficient Algorithms for Network Analysis and Bioinformatics Applications*.
1st December 2020 - 30th November 2022

Visiting Research Fellow, Department of Computer Science, Brown University, RI (USA), advised by Prof. Eli Upfal. January 2019 - June 2019

EDUCATION

Ph.D. in Information Engineering, University of Padova (XXXIII cycle) 24th March 2021
Thesis: *Rigorous and Efficient Algorithms for Significant and Approximate Pattern Mining*.
http://www.dei.unipd.it/~pelleagri/thesis/leonardo_pellegrina_tesi.pdf
Advisor: Prof. Fabio Vandin.

M.S. in Computer Engineering, University of Padova, *summa cum laude* 12th September 2017
Thesis: *Efficient Incremental Mining of Significant Patterns using Permutation Testing*.
<https://www.dei.unipd.it/~pelleagri/thesis/MS-thesis-LP.pdf>
Advisor: Prof. Fabio Vandin.

B.S. in Information Engineering, University of Padova

24th November 2015

Thesis: *Network Performance Evaluation of Fat-Tree and B-Cube Topologies*.

Advisor: Prof. Leonardo Badia.

TEACHING ACTIVITIES

Lecturer, *Methods for Scalable Graph Analytics*, PhD Doctoral Programme in Information Engineering, University of Padova (responsible for 10 hours) A.A. 2024/2025

Lecturer, *Introduction to Machine Learning and Artificial Intelligence: Methods and Applications*, UNIPhD Doctoral Programme, PhD Training week, University of Padova (responsible for 4 hours) 15th March 2024

Lecturer, *Dati e Algoritmi (Data structures and Algorithms)* 9 cfu, B.S. in Computer Engineering, Department of Information Engineering, University of Padova. A.A. 2024/2025

Lecturer, *Dati e Algoritmi (Data structures and Algorithms)* 9 cfu, B.S. in Computer Engineering, Department of Information Engineering, University of Padova. Students evaluation (Soddisfazione Complessiva): 8,04/10. A.A. 2023/2024

Lecturer (Docente a contratto con responsabilità dell'insegnamento, 6 cfu), *Dati e Algoritmi (Data structures and Algorithms)*, B.S. in Computer Engineering, Department of Information Engineering, University of Padova. Students evaluation (Soddisfazione Complessiva): 7,89/10. A.A. 2022/2023

Lecturer (Docente a contratto con responsabilità dell'insegnamento, 6 cfu), *Dati e Algoritmi 1 (Data structures and Algorithms 1)*, B.S. in Computer Engineering, Department of Information Engineering, University of Padova. Students evaluation (Soddisfazione Complessiva): 8,07/10. A.A. 2021/2022

Teaching Assistant (Tutor Junior), *Machine Learning* (total 80 hours), M.S. in Computer Engineering, Department of Information Engineering, University of Padova 2017 - 2021

Teaching Assistant (Tutor Junior), *Foundations of Computer Science* (total 50 hours), B.S. in Computer Engineering, Department of Information Engineering, University of Padova 2019 - 2020

Teaching Assistant (Didattica Integrativa), *Big Data Computing* (total 50 hours), M.S. in Computer Engineering, Department of Information Engineering, University of Padova 2017 - 2019

Teaching Assistant (Tutor Junior), *Tutorato Formativo* (total 100 hours), B.S. in Information Engineering, Department of Information Engineering, University of Padova 2017 - 2019

PUBLICATIONS

Authors are sorted by contribution (not alphabetical). First and last authorship are denoted with *.

Indicators (as 15th October 2025):

- Scopus: h-index 8, number of citations 166.
- Google Scholar: h-index 8, number of citations 247.

Journals

- [J1] *Fast Approximation of Frequent k-mers and Applications to Metagenomics*, **Leonardo Pellegrina***, Cinzia Pizzi and Fabio Vandin, **Journal of Computational Biology Special Issue of Best Papers from the RECOMB 2019 conference**, 27, 4, 534-549 (2020). <https://doi.org/10.1089/cmb.2019.0314>

- [J2] *Efficient Mining of the Most Significant Patterns with Permutation Testing*, **Leonardo Pellegrina*** and Fabio Vandin, **Data Mining & Knowledge Discovery** 34, 1201–1234 (2020). <https://doi.org/10.1007/s10618-020-00687-8>
- [J3] *SPRISS: Approximating Frequent k -mers by Sampling Reads, and Applications*, Diego Santoro, **Leonardo Pellegrina**, Matteo Comin, and Fabio Vandin, **Bioinformatics**, 38(13), 3343–3350, <https://doi.org/10.1093/bioinformatics/btac180>
- [J4] *MCRapper: Monte-Carlo Rademacher Averages for Poset Families and Approximate Pattern Mining*, **Leonardo Pellegrina***, Cyrus Cousins, Fabio Vandin and Matteo Riondato, **ACM Transactions on Knowledge Discovery and Data Mining** (2022) 16, 6, Article 124. <https://doi.org/10.1145/3532187>
- [J5] *Discovering Significant Evolutionary Trajectories in Cancer Phylogenies*, **Leonardo Pellegrina*** and Fabio Vandin, **Bioinformatics** (2022), Volume 38. <https://doi.org/10.1093/bioinformatics/btac467>
- [J6] *SILVAN: Estimating Betweenness Centralities with Progressive Sampling and Non-uniform Rademacher Bounds*, **Leonardo Pellegrina*** and Fabio Vandin, **ACM Transactions on Knowledge Discovery and Data Mining** (2023) 18, 3, Article 52. <https://dl.acm.org/doi/10.1145/3628601>
- [J7] *Efficient Discovery of Significant Patterns with Few-Shot Resampling*, **Leonardo Pellegrina*** and Fabio Vandin, **Proceedings of the VLDB Endowment** (2024) 17, 10. <https://doi.org/10.14778/3675034.3675055>

Conference Proceedings

- [C1] *An innovative space tether deployer with retrieval capability: Design and test of STAR experiment*, Gilberto Grassi, Alessia Gloder, **Leonardo Pellegrina**, Mattia Pezzato, Alvise Rossi, Francesco Branz, Matteo Duzzi, Riccardo Mantellato, Lorenzo Olivieri, Francesco Sansone, Enrico C. Lorenzini, Alessandro Francesconi, Proceedings of the 68th International Astronautical Congress (**IAC 2017**). <http://hdl.handle.net/11577/3243555>
- [C2] *Design and test in microgravity of a space tether length and length rate measurement device*, Gilberto Grassi, Mattia Pezzato, Alessia Gloder, Riccardo Mantellato, Alessandro Francesconi, Enrico Lorenzini, Alvise Rossi and **Leonardo Pellegrina**, IEEE International Workshop on Metrology for AeroSpace (**MetroAeroSpace 2017**). <https://doi.org/10.1109/MetroAeroSpace.2017.7999602>
- [C3] *Towards open-source shared implementations of keyword-based access systems to relational data*, Alex Badan, Luca Benvegnú, Matteo Basetton, Giovanni Bonato, Alessandro Brighente, Alberto Cenzato, Piergiorgio Ceron, Giovanni Cogato, Stefano Marchesin, Alberto Minetto, **Leonardo Pellegrina**, Alberto Purpura, Riccardo Simionato, Nicolò Soletti, Matteo Tassarotto, Andrea Tonon, Federico Vendramin, Nicola Ferro, Workshop Proceedings of the joint 20th International Conferences on Extending Database Technology and Database Theory (**EDBT/ICDT 2017**). <http://hdl.handle.net/11577/3254670>
- [C4] *Keyword-based access to relational data: To reproduce, or to not reproduce?*, Alex Badan, Luca Benvegnu, Matteo Basetton, Giovanni Bonato, Alessandro Brighente, Stefano Marchesin, Alberto Minetto, **Leonardo Pellegrina**, Alberto Purpura, Riccardo Simionato, Matteo Tassarotto, Andrea Tonon, Nicola Ferro, Proceedings of the Symposium on Advanced Database Systems (**SEBD 2017**). <http://hdl.handle.net/11577/3254675>
- [C5] *Efficient Mining of the Most Significant Patterns with Permutation Testing*, **Leonardo Pellegrina*** and Fabio Vandin, Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2018**). (18.4% acceptance rate) <https://doi.org/10.1145/3219819.3219997>
- [C6] *Fast Approximation of Frequent k -mers and Applications to Metagenomics*, **Leonardo Pellegrina***, Cinzia Pizzi and Fabio Vandin, Proceedings of the 23th Annual International Conference on Research in Computational Molecular Biology (**RECOMB 2019**). https://link.springer.com/chapter/10.1007/978-3-030-17083-7_13

- [C7] *SPuManTE: Significant Pattern Mining with Unconditional Testing*, **Leonardo Pellegrina***, Matteo Riondato and Fabio Vandin, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2019**). (14.2% acceptance rate) <https://dl.acm.org/doi/abs/10.1145/3292500.3330978>
- [C8] *MCRapper: Monte-Carlo Rademacher Averages for Poset Families and Approximate Pattern Mining*, **Leonardo Pellegrina***, Cyrus Cousins, Fabio Vandin and Matteo Riondato, Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2020**). (16.8% acceptance rate) <https://doi.org/10.1145/3394486.3403267>
- [C9] *SPRISS: Approximating Frequent k-mers by Sampling Reads, and Applications*, Diego Santoro*, **Leonardo Pellegrina***, and Fabio Vandin, 25th Annual International Conference on Research in Computational Molecular Biology (**RECOMB 2021**). <https://doi.org/10.48550/arXiv.2101.07117>
- [C10] *Efficient Centrality Maximization with Rademacher Averages*, **Leonardo Pellegrina***, Proceedings of the 29th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2023**). (22.1% acceptance rate) <https://dl.acm.org/doi/abs/10.1145/3580305.3599325>
- [C11] *Scalable Rule Lists Learning with Sampling*, **Leonardo Pellegrina***, Fabio Vandin, Proceedings of the 30th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2024**). ($\approx 20\%$ acceptance rate) <https://doi.org/10.1145/3637528.3671989>
- [C12] *Fast Percolation Centrality Approximation with Importance Sampling*, Antonio Cruciani, **Leonardo Pellegrina***, Proceedings of the IEEE International Conference on Data Mining (**ICDM 2025**). (22% acceptance rate)

Thesis

- [1] *Efficient Incremental Mining of Significant Patterns*, **Leonardo Pellegrina**, M.S. Thesis (2017). <https://www.dei.unipd.it/~pelleagri/thesis/MS-thesis-LP.pdf>
- [2] *Rigorous and Efficient Algorithms for Significant and Approximate Pattern Mining*, **Leonardo Pellegrina**, Ph.D. Thesis (2021). http://www.dei.unipd.it/~pelleagri/thesis/leonardo_pellegrina_tesi.pdf

Preprints

- [1] *Sharper convergence bounds of Monte Carlo Rademacher Averages through Self-Bounding functions*, **Leonardo Pellegrina**, arXiv (2020). <https://arxiv.org/abs/2010.12103>

Patents

Patent no. IT201700039763A1 for the invention *Automatic equipment for the launch and controlled recovery of a tethered mass in orbit*, c/o Italian Office for Patents and Trademarks (2017).

Patent no. WO2018189633A1 for the invention *Automatic apparatus for the controlled launch and recovery of a tethered mass in orbit*, World Intellectual Property Organization (2018). <https://patents.google.com/patent/WO2018189633A1>

PRESENTATIONS AT INTERNATIONAL CONFERENCES

Efficient Mining of the Most Significant Patterns with Permutation Testing, 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2018) (short talk), London (UK). 21th August 2018

Fast Approximation of Frequent k-mers and Applications to Metagenomics, 23th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2019), Washington, DC (USA). 8th May 2019

SPuManTE: Significant Pattern Mining with Unconditional Testing, 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2019) (short talk), Anchorage, AK (USA). 6th August 2019

MCRapper: Monte-Carlo Rademacher Averages for Poset Families and Approximate Pattern Mining, 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020), San Diego, CA (USA), virtual event. 26th August 2020

Discovering Significant Evolutionary Trajectories in Cancer Phylogenies, 26th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2022) (poster), La Jolla, San Diego, USA, 2022. 23th May 2022

Discovering Significant Evolutionary Trajectories in Cancer Phylogenies, 30th Conference on Intelligent Systems for Molecular Biology (ISMB 2022), Madison, Wisconsin, USA, 2022. 12th July 2022

Discovering Significant Evolutionary Trajectories in Cancer Phylogenies, 21th European Conference on Computational Biology (ECCB 2022), Sitges, Barcelona, ES, 2022. 20th September 2022

Efficient Centrality Maximization with Rademacher Averages, 29th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2023), Long Beach, CA (USA). 8th August 2023

SILVAN: Estimating Betweenness Centralities with Progressive Sampling and Non-uniform Rademacher Bounds, 20th International Workshop on Mining and Learning with Graphs (MLG), In conjunction with ECMLPKDD 2023 Conference, (short talk) Turin, Italy, 2023. 22th September 2023

Efficient Discovery of Significant Patterns with Few-Shot Resampling, 50th International Conference on Very Large Data Bases (VLDB 2024) Guangzhou, China. 27th August 2024

INVITED TALKS

Rigorous Mining of the Most Significant Patterns, International Workshop on Foundations of Learning from Data (FouLarD), Bertinoro, Italy. 10th September 2018

Efficient Mining of the Most Significant Patterns with Permutation Testing, Kick off meeting of the PRIN project AHeAd, Luiss University, Rome, Italy. 23th September 2019

SPuManTE: Significant Pattern Mining with Unconditional Testing, International Workshop on Computation and Statistics in Data science (CaStleD), Bertinoro, Italy. 4th October 2019

Fast Approximation of Frequent k -mers and Applications, invited lecture for the *Machine Learning* course, Department of Information Engineering, University of Padova, Italy. 25th November 2019

SPRISS: Approximating Frequent k -mers by Sampling Reads, and Applications, 9th International RECOMB Satellite Workshop on Computational Methods in Genetics (RECOMB-Genetics). 2th September 2021

Identifying Patterns of Tumor Evolution, ARQUS European Alliance Research Focus Forum on Cancer Research, Department of Surgery, Oncology and Gastroenterology, University of Padova. 29th September 2022

Statistical learning techniques to efficiently identify central nodes from large graphs, International Workshop on Scalable Approaches to High Performance Computing (ScalPerf), Bertinoro, Italy.

18th September 2023

Efficient Discovery of Significant Patterns with Few-Shot Resampling, school of Management Sciences and Information Systems, Rutgers Business School (virtual talk). 8th November 2024

FELLOWSHIPS AND AWARDS

Hans von Muldau Team Award for the Best Team Project at the International Astronautical Congress (IAC) 2017, Adelaide, Australia. September 2017

Cariparo Foundation Ph.D. Fellowship, Padova, Italy. 2017 - 2021

Luciano Iglesias Foundation Fellowship, Padova, Italy. May 2018

Ing. Aldo Gini Foundation Fellowship, Padova, Italy. November 2018

ISCB RECOMB 2019 Travel Fellowship. February 2019

Honorable mention for 2021 ACM SIGKDD Ph.D. Dissertation Award (Worldwide top-3 Ph.D. thesis in Data Mining & Knowledge Discovery).
<https://kdd.org/awards/view/2021-sigkdd-dissertation-award-winners> August 2021

Honorable mention for Best Program Committee member (top 5%) at **ACM The Web Conference 2022**. <https://www2022.thewebconf.org/awards/#mention-2> April 2022

ISCB ECCB 2022 Travel Fellowship. June 2022

Honorable mention for Best Program Committee member (top 5%) at **ACM The Web Conference 2023**. <https://dl.acm.org/action/showFmPdf?doi=10.1145%2F3543507> May 2023

Teaching Award of Merit, University of Padova, academic year 23/24.
<https://www.unipd.it/premialita-rtd-didattica> 14th November 2024

SCIENTIFIC ACTIVITIES

Organization roles

RECOMB 2020, International Conference on Research in Computational Molecular Biology (Publicity and Social Media chair).

RECOMB 2021, International Conference on Research in Computational Molecular Biology (Publicity and Social Media chair).

Co-organizer and Co-chair of the special session *Informatics research in bioinformatics: contributions from the CINI-InfoLife network* at **CIBB 2024** (19th conference on Computational Intelligence methods for Bioinformatics and Biostatistics).

Research Groups and Projects

University of Padova, **Algorithms for Inferential Data Mining**. Role: WP leader. Total amount: €140.000. 2018 - 2020

Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), **PRIN AHeAD: efficient Algorithms for HArnessing networked Data**. Role: participant. Total amount: €784.860. 2019 - 2022

Department of Information Engineering, University of Padova. Junior Postdoctoral Researcher competitive grant (call open to research projects proposals from all Information Engineering area), project **Rigorous and Efficient Algorithms for Network Analysis and Bioinformatics Applications**. Role: PI. 1st December 2020 - 30th November 2022

University of Padova, **SID 2020: Resource-Allocation Tradeoffs for Dynamic and Extreme data (RATED-X)**. Role: participant. Total amount: €12.000. 2021 - 2022

European Union, **BRAINTEASER: BRinging Artificial INTelligence home for a better cAre of amyotrophic lateral sclerosis and multiple SclERosis**. Role: participant. Total UniPD amount: €732.250. 2021 - 2024

Department of Information Engineering, University of Padova. Senior Postdoctoral Researcher competitive grant (call open to research projects proposals from all Information Engineering area, ranked 1st out of 7 submissions), project **Efficient Algorithms to Identify Significant Patterns in Tumors**. Role: PI. 1st December 2022 - 1th March 2023

Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), **PRIN EXPAND: scalable algorithms for EXPloratory Analyses of heterogeneous and dynamic Networked Data**. Role: participant. 2023 - 2026

Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), National Center for HPC, Big Data, and Quantum Computing, **PNRR CN1 SPOKE 1, Future HPC & Big Data**, project CN00000013. Role: participant. 2023 - 2026

Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR), Complementary National Plan PNC-I.1 *Research initiatives for innovative technologies and pathways in the health and welfare sector*, D.D. 931 of 06/06/2022, PNC0000002 **DARE - Digital Lifelong Prevention**. Role: participant. 2023 - 2026

Tutorials

- [1] *Multiple Hypothesis Testing and Statistically-sound Pattern Mining*, **Leonardo Pellegrina**, Matteo Riondato, Fabio Vandin, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD 2019**).
<https://doi.org/10.1145/3292500.3332286>
- [2] *Multiple Hypothesis Testing and Statistically-sound Pattern Mining*, **Leonardo Pellegrina**, Matteo Riondato, Fabio Vandin, Proceedings of the 2021 SIAM International Conference on Data Mining (**SDM 2021**).
https://meetings.siam.org/session/dsp_programsess.cfm?SESSIONCODE=71922

Editorial roles

Member of the Editor Board of Big Data Research journal. July 2025 - today

Member of the Editor Board of Data Science and Engineering journal. July 2025 - today

Member of the Editor Board of Frontiers in Big Data journal (Review Editor of the Data Mining and Management section). April 2024 - today

Action Editor for the Conference Editorial Board of the 29th International Conference on System Theory, Control and Computing (**ICSTCC 2025**). December 2024 - today

Program Committees

RECOMB Satellite Workshop on Massively Parallel Sequencing (**RECOMB-Seq**) 2021.
ACM International Conference on Web Search and Data Mining (**WSDM**) 2022-2026.
SIAM International Conference on Data Mining (**SDM**) 2022-2025.

ACM The Web Conference (**TheWebConf**) 2022-2026.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**) 2022-2026.
Conference on Computational Intelligence Methods for Bioinformatics & Biostatistics (**CIBB**) 2023-2025.
The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**) 2023, 2024.
ACM International Conference on Information and Knowledge Management (**CIKM**) 2024, 2025.

Conference Reviewing

Workshop on Algorithms in Bioinformatics (**WABI**) 2018.
International Conference on Research in Computational Molecular Biology (**RECOMB**) 2019, 2020, 2022, 2024.
ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (**KDD**) 2019, 2020, 2021.
The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**) 2019, 2020, 2021.
ACM International Conference on Information and Knowledge Management (**CIKM**) 2019.
IEEE International Conference on Data Mining (**ICDM**) 2019, 2020, 2022.
ACM International WSDM Conference on Web Search and Data Mining (**WSDM**) 2020.
AAAI Conference on Artificial Intelligence 2020.
International Conference on Intelligent Systems for Molecular Biology (**ISMB**) 2020, 2024.
ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (**BCB**) 2021.
ACM The Web Conference (**TheWebConf**) 2021.
SIAM Conference on Applied and Computational Discrete Algorithms (**ACDA**) 2023.

Journal Reviewing

Journal of Machine Learning Research (**JMLR**)
Machine Learning
Data Mining & Knowledge Discovery (**DAMI**)
Scientific Reports
Journal of Graph Algorithms and Applications (**JGAA**)
Bioinformatics
Genome research
BMC Bioinformatics
IEEE Journal of Biomedical and Health Informatics (**JBHI**)

Supervision

Bachelor Students: Federico Berton (2021, co-supervised with Diego Santoro and Fabio Vandin), Mattia Bastianello, Giulia Beraldo, Francesco Biscaccia Carrara, Simone Corrà, Riccardo Modolo, Martina Naldoni, Alessandro Viespoli (from 2023 co-supervised with Fabio Vandin), Alberto Penzo, Eddy Frighetto, Matteo Mannuzzato (2023), Alessandra D'Apolito (2024), Nicolò Fioranzato, Sirio Trentin, (2025).

Master Students: Samuele Benfatti, Daniel Carlesso, Michele Russo, Davide Seghetto (from 2022, co-supervised with Fabio Vandin), Simone D'Antimo (2024).

Ph.D. Students: Giorgio Venturin (from 2023, co-supervised with Fabio Vandin).

Visiting Ph.D. Students: Sijing Tu, KTH Royal Institute of Technology (June-July 2024), Antonio Cruciani, GSSI (July 2024).

Professional Memberships

Member of the International Society for Computational Biology (**ISCB**).
Member of the Association for Computing Machinery (**ACM**).
Member of the CINI Data Science Laboratory.
Member of the Young-InfoLife CINI group.

SOFTWARE PACKAGES

- TopKWY, an efficient algorithm for mining the most significant patterns with permutation testing. <https://github.com/VandinLab/TopKWY>
- SAKEIMA, a fast algorithm to compute accurate approximations of frequent k -mers from large sequencing data sets based on random sampling of k -mers. <https://github.com/VandinLab/SAKEIMA>
- SPuManTE, an efficient algorithm for mining significant pattern with unconditional testing. <https://github.com/VandinLab/SPuManTE>
- MCRapper, a practical algorithm to compute rigorous approximations of pattern languages with a poset structure using Monte-Carlo Empirical Rademacher Averages. <https://github.com/VandinLab/MCRapper>
- SPRISS, a flexible and fast algorithm to compute sound approximations of frequent k -mers by sampling reads from large sequencing data sets. <https://github.com/VandinLab/SPRISS>
- SILVAN, a fast sampling-based algorithm based on non-uniform Monte-Carlo Empirical Rademacher Averages to compute rigorous approximations of the betweenness centrality of nodes in large graphs. <https://github.com/VandinLab/SILVAN>
- MASTRO, our algorithm to discover statistically significant conserved evolutionary trajectories of alterations from phylogenetic tumor trees. <https://github.com/VandinLab/MASTRO>
- CentRA, a scalable algorithm for centrality maximization using Rademacher Averages. <https://github.com/leonardopellegrina/CentRA>
- FSR, our algorithm to efficiently discover statistically significant patterns with few-shot resampling. <https://github.com/VandinLab/FSR>
- SamRuLe, our sampling-based algorithm to efficiently learn almost optimal interpretable prediction models (rule lists) from large datasets with sampling. <https://github.com/VandinLab/SamRuLe>