Carlo Tosoni

Nationality: Italian **Date of birth:** 16/04/1999 **Place of birth:** Assisi, Italy **Gender:** Male **Phone:** (+39) 3459913191 (Mobile) **Email address:** carlo.tosoni@unive.it **LinkedIn:** https://www.linkedin.com/in/carlo-tosoni-42203b273/ **Website:** https://CarloTosoni.eu **Address:** Marghera, Venice, Italy (Home)

ABOUT MYSELF

Master's degree in computer science

EDUCATION & TRAINING

PhD Student in Computer Science

Ca' Foscari University of Venice [08/09/2023 - Current]

Address: Via Torino 155 30170, Venice (Italy) | **Level in EQF:** 8

Sorting is one of the most powerful techniques for enabling search on a particular data structure. For instance, using binary search is possible to retrieve an item from a sorted list in logarithmic time w.r.t. the list dimension. Similar techniques can be applied also to more convoluted data structures, like for instance finite-state automata. In fact, sorting the states of an automaton based on the strings reaching them enables the development of efficient indexes on the regular language recognized by the automaton itself. This indexing strategy naturally extends the renowned Burrows-Wheeler transform, originally devised for strings, to arbitrary finite-state automata.

During the PhD programme, we aim to significantly improve this indexing strategy by providing efficient algorithms and data structures to sort the states of automaton and to provide time-space efficient techniques to locate patterns with the resulting index. As the current state-of-the-art algorithm to compute such indexes has quadratic time complexity (w.r.t. the number of transitions of the input automaton) this indexing technique is still unfeasible in the realm of the so-called big data. Therefore, to development of a near-linear time algorithm represents a crucial step to make this indexing strategy a viable option for prominent research fields like bioinformatics, where the data dimension has soared exponentially over the past years.

The PhD programme is funded by the following grant of the European Union and it is supervised by professors Nicola Prezza and Ruben Becker of the Ca' Foscari University of Venice.

MSc in Computer Science, curriculum Big Data Technologies

University of Pisa [09/2021 - 07/2023]

Address: Largo Bruno Pontecorvo 356127, Pisa (Italy) | **Final grade:** 110/110 cum laude | **Level in EQF:** 7 | **Thesis:** Compressing the Burrows-Wheeler transform of finite-state automata using run-length encoding

The master's degree in computer science, curriculum Big Data Technologies, is offered by the Department of Computer Science at the University of Pisa. All the courses are provided entirely in English and are focused on the following topics.

• Design, analysis and implementation of advanced algorithms and data structures to efficiently solve combinatorial problems. Analysis of techniques and algorithms for implementing advanced databases. State-of-the-art techniques and paradigms for the analysis of genome sequences in bioinformatics.
• Principles and paradigms of machine learning. Analysis and implementation of algorithms for data mining.
• Megadata analysis, search engines, and information retrieval.
• High performance computing and parallel computing.

Achieved seven times the grade 30/30 cum laude. Weighted average at the end of studies 30.47/32.

BSc in Computer and Electonic Engineering

University of Perugia [31/08/2018 - 31/08/2021]

Address: Via Goffredo Duranti 93, Perugia (Italy) | **Final grade:** 110/110 con lode | **Level in EQF:** 6 | **Thesis:** Development of a decentralized blog using Solid technology

Courses attended

• Feedback control systems, Signal theory, Principles of automatic control, Internet basics.
• Algorithms and data structures, Database management systems, Programming (Java and C).
• Logic design and microcontrollers, Circuit theory, Electronic devices and technologies.

Secondary-School Degree

Liceo Scientifico Statale Annesso al Convitto Nazionale "Principe di Napoli" [09/2013 - 07/2018]

Address: Piazza Giacomo Matteotti 6706081, Assisi (Italy) | **Final grade:** 100/100

PUBLICATIONS

[2025]

New Entropy Measures for Tries with Applications to the XBWT

Authors: Lorenzo Carfagna and Carlo Tosoni

arXiv

[2025]

Analysing New Entropy Measures for Tries

Authors: Lorenzo Carfagna and Carlo Tosoni

International Symposium on String Processing and Information Retrieval (SPIRE)

[2025]

Encoding Co-Lex Orders of Finite-State Automata in Linear Space

Authors: Ruben Becker, Nicola Cotumaccio, Sung-Hwan Kim, Nicola Prezza, and Carlo Tosoni

Annual Symposium on Combinatorial Pattern Matching (CPM)

[2024]

Indexing Finite-State Automata Using Forward-Stable Partitions

Authors: Ruben Becker, Sung-Hwan Kim, Nicola Prezza, and Carlo Tosoni

International Symposium on String Processing and Information Retrieval (SPIRE)

PROGRAMMING LANGUAGES

Programming languages and computer skills

Excellent knowledge of programming languages like Java, C++, C#, and C. Worked also with HTML, Javascript, and CSS and frameworks such as React, Qwik, and Node.js to develop web applications. Good knowledge also with Python 3 and libraries like Pandas, PyTorch, and Keras to train/test machine learning models and to apply data mining techniques. Familiarity with the programs Unity and Blender for designing virtual realities and creating 3D models. Worked occasionally with Haskell, MATLAB, and Rust.

VISITING RESEARCHER

[01/05/2026 - 31/07/2026]

University of Florida (planned)

Plan to visit the research group of Professor Christina Boucher at the University of Florida, in Gainesville, Florida, USA. The purpose of the mobility is to improve state-of-the-art algorithms to sequence/analyse genomic datasets.

[01/04/2026 - 30/06/2026]

University of Helsinki

Visited the department of computer science of the University of Helsinki (Finland) to discuss possible research collaborations with Professor Veli Mäkinen and his research group.

TALKS

[18/02/2026]

New Entropy Measures for Tries with Applications to the XBWT

Workshop Data Structures in Bioinformatics (DSB) 2026.

We introduce new trie entropy measures - worst-case and empirical - exploiting information on the distribution of the edge-labels to achieve compact representations. We show their reachability through an extension of arithmetic coding to tries, as well as the design of efficient trie indexes compressed to their empirical entropy.

[05/02/2025]

Efficient Indexes for Pangenome Graphs through BWT-Based Data Structures

Workshop Data Structures in Bioinformatics (DSB) 2025.

The talk summarized the state-of-the-art techniques to extend the renowned Burrows-Wheeler transform from strings to arbitrary edge-labeled graphs. This transform has enormous potential in bioinformatics, as it can be employed to simultaneously compress and index pangenome graphs representing the DNA of a population.

CERTIFICATIONS

[British Council, 19/03/2022]

IELTS Academic 7.0 Listening 7.5, Reading 7.5, Writing 6.5, Speaking 6.0.

[Huawei Technologies co. ltd, 24/01/2021]

Huawei HCIA Routing and Switching Certification

PEER REVIEWS FOR SCIENTIFIC PUBLICATIONS

Reviews for journals

Reviewed articles for the following journals: Information Systems, Algorithms for Molecular Biology, and PLOS One.

Reviews for conferences/workshops

Appointed sub-reviewer for the following conferences and workshops: Data Compression Conference (DCC), 2024 and 2025. Symposium on Combinatorial Pattern Matching (CPM), 2024. Conference on Wonderful Algorithms in Bioinformatics (WABI), 2024. Symposium on Simplicity in Algorithms (SOSA), 2024. International Workshop on Combinatorial Algorithms (IWOCA), 2025. European Symposium on Algorithms (ESA), 2025. International Symposium on Algorithms and Computation (ISAAC), 2025. International Symposium on Theoretical Aspects of Computer Science (STACS), 2026.

CONFERENCES AND WORKSHOPS ORGANISED

[18/02/2026 - 19/02/2026]

Data Structures in Bioinformatics

Member of the organizing committee for the 11th edition of the workshop on Data Structures in Bioinformatics (DSB 2026).

[21/07/2025 - 25/07/2025]

Symposium on Experimental Algorithms

Member of the organizing committee for the 23rd Symposium on Experimental Algorithms (SEA 2025).

TEACHING ACTIVITIES

[01/03/2026 - 31/05/2026]

Tutor for the course "Algorithms and Learning over Massive Data"

Ca' Foscari University of Venice

[01/07/2024 - 05/07/2024]

Tutor for the summer schools SEAA (Scuola Estiva di Algoritmi Avanzati)

Ca' Foscari University of Venice

ISTITUTIONAL ROLES

[12/2023 - 12/2025]

Representative of the PhD students in computer science

Appointed representative of the PhD students in computer science at the Ca' Foscari University of Venice for the academic years 2023/24 and 2024/25.

DRIVING LICENCE

Driving licence: B

Venezia, 02/01/2025

Carlo Tosoni