Giulio Ermanno **Pibiri**

Curriculum Vitae et Studiorum

▆▆▆ Contact Information

National Research Council of Italy (CNR)

Institute of Science and Information Technologies "A. Faedo" (ISTI)

Via G. Moruzzi 1, 56124 Pisa, Italy

Email	giulio.pibiri@di.unipi.it
Email	giulio.ermanno.pibiri@isti.cnr.it
Personal page	http://pages.di.unipi.it/pibiri

▆▆▆ Personal Information

Place of birth	Florence, Italy
Date of birth	13 July 1990

▆▆▆ Research Interests

Indexing, Efficiency, Data Compression, Algorithm Engineering

▆▆▆ Education

2015 – 2018   **PhD in Computer Science (INF/01)**.
- University of Pisa, Pisa, Italy
- Thesis: *Space- and Time-Efficient Data Structures for Massive Datasets*.
  Defended on 08/03/2019.
- Grade: Excellent.
- Supervisor: Rossano Venturini (`http://pages.di.unipi.it/rossano`)

2012 – 2014   **Master Degree in Computer Science & Networking (class LM18)**.
- University of Pisa and Scuola Superiore Sant'Anna, Pisa, Italy
- Thesis: *Dynamic Elias-Fano Encoding*.
  Defended on 06/03/2015.
- Grade: 110/110 *summa cum laude*.
- Supervisor: Rossano Venturini (`http://pages.di.unipi.it/rossano`)

2009 – 2012   **Bachelor Degree in Computer Engineering (class L08)**.
- University of Florence, Florence, Italy
- Thesis: *Quantum Computation & Grover's Algorithm*.
  Defended on 09/10/2012.
- Grade: 110/110 *summa cum laude*.
- Supervisor: Gabriele Vezzosi (`http://www.dma.unifi.it/~vezzosi`)

2004 – 2009   **High School Diploma**.
- Liceo Scientifico Statale Guido Castelnuovo, Florence, Italy
- Grade: 100/100.

Professional Employment

11/2018 – present   **Postdoctoral Researcher in Computer Science**.
- Institute of Science and Information Technologies "A. Faedo" (ISTI), National Research Council of Italy (CNR), Pisa, Italy
- HPC-Lab headed by Raffaele Perego (`http://raffaele.isti.cnr.it`)

06/2017 – present   **Software Developer at ISTI-CNR**.
- European project *Big Data Grapes*, Indexing and Distributed Components, Pisa, Italy
- European project *Large-scale Indie Gaming Analytics*, Pisa, Italy

11/2015 – 11/2018   **PhD Student in Computer Science**.
- University of Pisa, Pisa, Italy
- Thesis: *Space- and Time-Efficient Data Structures for Massive Datasets*.
- Supervisor: Rossano Venturini (`http://pages.di.unipi.it/rossano`)

05/2018 – 10/2018   **Visiting PhD Student**.
- The University of Melbourne, School of Computing and Information Systems, Melbourne, Australia
- Supervisor: Alistair Moffat (`https://people.eng.unimelb.edu.au/ammoffat`)

04/2018 – 05/2018   **Visiting PhD Student**.
- RIKEN Advanced Intelligence Project (AIP), Tokyo, Japan
- Supervisor: Yasuo Tabei (`https://sites.google.com/site/yasuotabei`)

04/2015 – 07/2015   **Software Engineering Intern at IBM**.
Rome, Italy

Awards and Grants

2020   *Young Researcher Award* issued by ISTI-CNR.

2017   *SIGIR Student Travel Grant* issued by ACM SIGIR.

2015   *PhD Scholarship* issued by the University of Pisa, Department of Computer Science.

2015   *Master Degree Award: Best Performance a.y. 2013/2014* issued by Scuola Superiore Sant'Anna.

2015   *Best Master Thesis Award in Theoretical Computer Science*, issued by the Italian chapter of the European Association for Theoretical Computer Science (EATCS).

Publications

2021   Giulio Ermanno Pibiri, *Fast and Compact Set Intersection through Recursive Universe Partitioning*. IEEE Data Compression Conference, pages 10.

2020   Giulio Ermanno Pibiri and Shunsuke Kanda, *Rank/Select Queries over Mutable Bitmaps*. CoRR, `http://arxiv.org/abs/2009.12809`, pages 21.

2020   Giulio Ermanno Pibiri and Rossano Venturini, *Techniques for Inverted Index Compression*. ACM Computing Surveys (CSUR), pages 36.

2020   Giulio Ermanno Pibiri and Rossano Venturini, *Practical Trade-Offs for the Prefix-Sum Problem*. Software: Practice and Experience (SPE), pages 29.

2020	Simon Gog, Giulio Ermanno Pibiri and Rossano Venturini, *Efficient and Effective Query Auto-Completion*. ACM Conference on Research and Development in Information Retrieval (SIGIR), pages 10.
2020	Giulio Ermanno Pibiri and Rossano Venturini, *Succinct Dynamic Ordered Sets with Random Access*. CoRR, `https://arxiv.org/abs/2003.11835`, pages 15.
2020	Raffaele Perego, Giulio Ermanno Pibiri and Rossano Venturini, *Compressed Indexes for Fast Search of Semantic Data*. IEEE Transactions on Knowledge and Data Engineering (TKDE), pages 12.
2019	Giulio Ermanno Pibiri and Rossano Venturini, *On Optimally Partitioning Variable-Byte Codes*. IEEE Transactions on Knowledge and Data Engineering (TKDE), pages 12.
2019	Giulio Ermanno Pibiri and Rossano Venturini, *Handling Massive N-Gram Datasets Efficiently*. ACM Transactions on Information Systems (TOIS), pages 41.
2019	Giulio Ermanno Pibiri, Matthias Petri, Alistair Moffat, *Fast Dictionary-based Compression for Inverted Indexes*. ACM Conference on Web Search and Data Mining (WSDM), pages 9.
2018	Giulio Ermanno Pibiri and Rossano Venturini, *Variable-Byte Encoding is Now Space-Efficient Too*. CoRR, `https://arxiv.org/abs/1804.10949`, pages 14.
2018	Giulio Ermanno Pibiri and Rossano Venturini, *Inverted Index Compression*. Encyclopedia of Big Data Technologies, pages 8.
2017	Giulio Ermanno Pibiri and Rossano Venturini, *Efficient Data Structures for Massive N-Gram Datasets*. ACM Conference on Research and Development in Information Retrieval (SIGIR), pages 10.
2017	Giulio Ermanno Pibiri and Rossano Venturini, *Dynamic Elias-Fano Representation*. Annual Symposium on Combinatorial Pattern Matching (CPM), pages 14.
2017	Giulio Ermanno Pibiri and Rossano Venturini, *Clustered Elias-Fano Indexes*. ACM Transactions on Information Systems (TOIS), volume 2, pages 33.

Software

GitHub page	https://github.com/jermp
rank_select	The C++ library used for the experiments in the paper *Rank/Select Queries over Mutable Bitmaps*.
psds	The C++ library used for the experiments in the paper *Practical Trade-Offs for the Prefix-Sum Problem*.
essentials	A C++ library providing essential core utilities for data structure design and benchmarking.
autocomplete	The C++ library used for the experiments in the paper *Efficient and Effective Query Auto-Completion*.
2i_bench	The C++ library used for the experiments in the paper *Techniques for Inverted Index Compression*.

s_indexes	The C++ library used for the experiments in the paper *Fast and Compact Set Intersection thorough Recursive Universe Partitioning*.
rdf_indexes	The C++ library used for the experiments in the paper *Compressed Indexes for Fast Search of Semantic Data*.
dint	The C++ library used for the experiments in the paper *Fast Dictionary-based Compression for Inverted Indexes*.
opt_vbyte	The C++ library used for the experiments in the paper *On Optimally Partitioning Variable-Byte Codes*.
tongrams	The C++ library implementing the compressed data structures and algorithms described in the papers *Efficient Data Structures for Massive N-Gram Datasets* and *Handling Massive N-Gram Datasets Efficiently*.
clustered_indexes	The C++ library used for the experiments in the paper *Clustered Elias-Fano Indexes*.

Talks

27/07/2020	*Efficient and Effective Query Auto-Completion*. SIGIR conference presentation. Virtual event.
17/09/2019	*Compressed Indexes for Fast Search of Semantic Data*. IIR conference presentation. Department of Information Engineering, Padova, Italy.
07/06/2019	*Ordered Set Problems*. Seminar. ISTI-CNR, Pisa, Italy.
08/03/2019	*Space- and Time-Efficient Data Structures*. PhD thesis defense. The University of Pisa, Pisa, Italy.
12/02/2019	*Fast Dictionary-based Compression for Inverted Indexes*. WSDM conference presentation. Melbourne Exhibition Center, Melbourne, Australia.
01/02/2019	*Indexing Compressed Data for Fast Retrieval*. Talk. The University of Pisa, Pisa, Italy.
15/11/2018	*Space- and Time-Efficient Data Structures*. PhD research results. The University of Pisa, Pisa, Italy.
29/10/2018	*Effective Web Graph Representations*. Seminar. The University of Pisa, Pisa, Italy.
17/05/2018	*On Optimally Partitioning Variable-Byte Index Data*. Seminar. RMIT University, Melbourne, Australia.
10/04/2018	*Elias-Fano Encoding: a powerful tool for data structure design*. Seminar. RIKEN AIP, Tokyo, Japan.
10/10/2017	*Space- and Time-Efficient Data Structures*. PhD research results. The University of Pisa, Pisa, Italy.
10/08/2017	*Efficient Data Structures for Massive N-Gram Datasets*. SIGIR conference presentation. Keio Plaza Hotel, Tokyo, Japan.
06/07/2017	*Dynamic Elias-Fano Representation*. CPM conference presentation. University Library of Warsaw, Warsaw, Poland.
06/06/2017	*Efficient Data Structures for Massive N-Gram Datasets*. IIR conference presentation. Università della Szizzera Italiana, Lugano, Switzerland.

17/10/2016	*Space- and Time-Efficient Data Structures* PhD thesis proposal. The University of Pisa, Pisa, Italy.
21/06/2016	*Elias-Fano Encoding: succinct representation of monotone integer sequences with search operations*. Seminar. The University of Pisa, Pisa, Italy.

Teaching Experience

02/2020 – 06/2020	Teacher for *Algorithmics and Laboratory - Corso B, code 008AA*, Bachelor Degree in Computer Science, University of Pisa, Italy
02/2019 – 06/2019	Assistant for *Algorithmics and Laboratory - Corso A, code 008AA*, Bachelor Degree in Computer Science, University of Pisa, Italy
10/2018 – 12/2018	Assistant for *Competitive Programming and Contests, code 645AA*, Master Degree in Computer Science, University of Pisa, Italy
09/2017 – 12/2017	Assistant for *Competitive Programming and Contests, code 645AA*, Master Degree in Computer Science, University of Pisa, Italy
10/2016 – 12/2016	Teacher for *Algorithmics and Laboratory - Corso di recupero, code 008AA*, Bachelor Degree in Computer Science, University of Pisa, Italy
02/2016 – 06/2016	Assistant for *Algorithmics and Laboratory - Corso A, code 008AA*, Bachelor Degree in Computer Science, University of Pisa, Italy

Professional Activities

2021	Member of the Program Committee of the 44-th edition of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR 2021).
2021	Member of the Program Committee of the 43-rd European Conference on Information Retrieval (ECIR 2021).
2020	Member of the Program Committee of the 14-th International ACM Conference on Web Search and Data Mining (WSDM 2021).
2020	Member of the Program Committee of the 29-th edition of the International ACM SIGIR Conference on Information and Knowledge Management, (CIKM 2020).
2020	Member of the Organizing Committee of the 28-th edition of the Annual European Symposium on Algorithms (ESA 2020).
2020	Member of the Program Committee of the 43-rd edition of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR 2020).
2019	Member of the Program Committee of the 42-nd edition of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR 2019).
2019	Member of the Organizing Committee of the 30-th edition of the International Symposium on Combinatorial Pattern Matching (CPM 2019).

2018	Member of the Program Committee of the 2-nd edition of the Workshop on Knowledge Graphs and Semantics for Text Retrieval and Analysis (KG4IR), in conjuction with ACM SIGIR 2018.
2017	Member of the Organizing Committee of the 24-th edition of the International Symposium on String Processing and Information Retrieval (SPIRE 2017).
2016	Student volunteer for the organization of the 39-th edition of the International ACM SIGIR Conference on Research and Development in Information Retrieval, (SIGIR 2016).

Languages

Italian	**Native**	CEFR level: C2
English	**Fluent**	CEFR level: C1
2018	**TOEFL iBT in English**. 100 (HIGH level)	
2008	**First Certificate in English (Level B2)**. University of Cambridge, Cambridge, United Kingdom	

Driving Licences

01/10/2010	Driving License of type B
23/01/2010	Driving License of type A

Web

Personal	http://pages.di.unipi.it/pibiri
GitHub	https://github.com/jermp
LinkedIn	https://www.linkedin.com/in/giulio-ermanno-pibiri-496ab3ab/