Nicola Prezza

Scientific CV

▬▬▬▬ Education

Jan 1, 2014 – Dec 31, 2016	**Ph.D, Computer Science, Mathematics, and Physics**, *Department of Computer Science, University of Udine, Via delle Scienze, 206, 33100 Udine, Italy*, Thesis: *Compressed Computation for Text Indexing*, Final grade: passed cum laude.
Sep 26, 2011 – Oct 18, 2013	**Master's degree in Computer Science**, *Department of Computer Science, University of Udine, Via delle Scienze, 206, 33100 Udine, Italy*, Final grade: 110/110 cum laude. GPA: 3.90.
Sep 25, 2008 – Dec 14, 2011	**Bachelor's degree in Computer Science**, *Department of Computer Science, University of Udine, Via delle Scienze, 206, 33100 Udine, Italy*, Final grade: 110/110 cum laude. GPA: 3.73.

▬▬▬▬ Employments

Feb 1, 2020 –	**Assistant professor of Computer Science**, *Department of Business and Management, LUISS University, Rome.*
Mar 1, 2018 – Jan 31st, 2020	**PostDoc researcher**, *Department of Computer Science, University of Pisa, Italy*, Project: *application of compressed indexes to aligment-free methods in bioinformatics*.
jan 1, 2017 – Mar 28, 2018	**PostDoc researcher**, *AlgoLoG, DTU Compute, Technical University of Denmark, Denmark*, Project: *Algorithms for the compression of big repetitive data*.
Jan 1, 2014 – Dec 31, 2016	**Ph.D.-student in Computer Science**, *University of Udine, Italy*, Project: *Compressed Computation for Text Indexing*.

▬▬▬▬ Teaching

PhD courses

Lug 30, 2018 – Aug 18, 2018	**Compact data structures (Ph.D course)**, *University of Coruña.* `https://www.udc.es/gl/iss/courses/courses_2018/Compact_data_structures/index.html`, 6 ECTS.
Sep 15, 2017 – Dec 8, 2017	**Compact data structures (Ph.D course)**, *Technical University of Denmark, Kgs. Lyngby, Denmark.* `www2.compute.dtu.dk/courses/02951`, 5 ECTS.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ nprezza@luiss.it • 🌐 nicolaprezza.github.io
in it.linkedin.com/in/nicola-prezza-573898b2 • ⌨ github.com/nicolaprezza   1/8

University courses and lectures

Fall 2019-2020	**Informatica (Introduction to Computer Science)**, *Bachelor Degree in in Economics and Management. Luiss Guido Carli University, Roma.*, 6 ETCS.
Fall 2019-2020	**Fondamenti di Programmazione e Laboratorio (Computer Programming)**, *Bachelor degree in Mathematics, University of Pisa* `http://pages.di.unipi.it/bodei/CORSO_FP_19/FP/index.html`, 6 ECTS.
28-29 Jan, 25-26 Feb. 2019	**Introduction to Compressed Computation**, *University of Udine, dipartimento di Scienze Matematiche, Informatiche e Fisiche*, 8 hours.
Fall 2018-2019	**Fondamenti di Programmazione e Laboratorio (Computer Programming)**, *Bachelor degree in Mathematics, University of Pisa* `http://pages.di.unipi.it/bodei/CORSO_FP_18/FP/index.html`, 6 ECTS.
24 Jul. 2017 – 28 Jul. 2017	**Aligning DNA sequences on compressed collections of genomes**, The CODATA-RDA Research Data Science Applied workshops on Extreme sources of data, Bioinformatics and IoT/Big-Data Analytics, *ICTP, Trieste*, 6 hours.
April-May 2016	**Compressed self-indexing**, *Four lectures within the master's degree course "Algoritmica" (coordinated by prof. Alberto Policriti). University of Udine*, 8 hours.
April-May 2015	**Compressed self-indexing**, *Four lectures within the master's degree course "Algoritmica" (coordinated by prof. Alberto Policriti). University of Udine*, 8 hours.

Other

Fall 2019-2020	**Coding with Phyton - Edition 2**, *Academig Gym of the Luiss Guido Carli University. 10 hours.*
Fall 2019-2020	**Coding with Phyton - Edition 1**, *Academig Gym of the Luiss Guido Carli University. 10 hours.*
Fall 2018-2019	**Programming with Phyton**, *High school DINI, Pisa, Italy*, 15 hours.

Awards

Jun 29, 2018	**PhD award UniUD for best scientific PhD thesis 2017**, *University of Udine.*
Jun 21, 2018	**"Best italian young researcher in Theoretical Computer Science 2018"**, *Italian Chapter of the European Association for Theoretical Computer Science (EATCS)*, `https://www.eatcs.org/index.php/italian-chapter/565`.
Jun 20, 2016	**Best talk award**, *MatBio 2016, King's college London*, Talk title: *A Randomized In-Place LCE Data Structure*.
Jan 1, 2014	**PhD fellowship**, *Department of Computer Science, University of Udine, Italy*.

Principal investigator of research projects

1. Principal Investigator of a project funded by Huawei on *lossless data compression*.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • ○ *github.com/nicolaprezza* *2/8*

Open-source software systems

I'm author of several open source software systems, most of which are hosted at `https://github.com/nicolaprezza`. Follows a list of the most relevant such repositories, with a list of the relative publications.

1. Compressed dynamic data structures library `DYNAMIC`. URL: `https://github.com/xxsds/DYNAMIC`. Developer of the library. The library at the moment counts 77 *stars* e 10 *forks* on GitHub. Work published in [26]. The library contains implementations of algorithms published in [7][32][29][30].

2. Compressed data structure `r-index`. URL: `https://github.com/nicolaprezza/r-index`. Developer of the library. The library at the moment counts 24 *stars* e 5 *forks* on GitHub. Work published in [1][19].

3. Compressed data structures library `BWTIL`. URL: `https://github.com/nicolaprezza/BWTIL`. Developer of the library. The library at the moment counts 19 *stars* e 5 *forks* on GitHub. Work published in [33].

4. DNA alilgner `ERNE 2`. URL: `http://erne.sourceforge.net/`. Main developer of the version 2 of the software. Work published in [9][8][35].

5. Tool `bwt2lcp` to compute the LCP data structure from the Burrows-Wheeler transform of a text. Developer of the library. URL: `https://github.com/nicolaprezza/bwt2lcp`. Work published in [17].

6. Tool `ebwt2InDel` to find DNA variations from the Burrows-Wheeler transform of a DNA reads dataset. Developer of the tool. URL: `https://github.com/nicolaprezza/ebwt2InDel`. Work published in [22][3].

7. Compressed data structure `lz-rlbwt`. Developer of the tool. URL: `https://github.com/nicolaprezza/lz-rlbwt`. Work published in in [31][27].

8. Data compression tool `re-pair`. Developer of the repository. URL: `https://github.com/nicolaprezza/Re-Pair`. Work published in [25].

Talks at International Conferences

This is a list of the talks I have given at international conferences. Invited talks are excluded: for those please refer to the following section. In addition to these talks, I have given more than 20 other talks at international workshops (not reported here for brevity).

1. *Regular Languages meet Prefix Sorting*. 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2020). January 5-8 2020, Salt Lake City.

2. *Online Algorithms on Antipowers and Antiperiods*. 26th International Symposium on String Processing and Information Retrieval, (SPIRE 2019), 8/10/2019, Segovia (Spain).

3. *Optimal Rank and Select Queries on Dictionary-Compressed Text*. 30th Annual Symposium on Combinatorial Pattern Matching, (CPM 2019), 18/6/2019, Pisa.

4. *Space-Efficient Computation of the LCP Array from the Burrows-Wheeler Transform*. 30th Annual Symposium on Combinatorial Pattern Matching (CPM 2019), 18/6/2019, Pisa.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • �ن *github.com/nicolaprezza*   *3/8*

5. *At the Roots of Dictionary Compression: String Attractors*. 50th Annual ACM Symposium on the Theory of Computing (STOC 2018), June 27, 2018, Los Angeles.

6. *In-Place Sparse Suffix Sorting*. Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), January 7-10, 2018, New Orleans.

7. *Optimal-time text indexing in BWT-runs bounded space*. Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), January 7-10, 2018, New Orleans.

8. *Succinct Partial Sums and Fenwick Trees*. String Processing and Information Retrieval (SPIRE), September 26-29, 2017, Palermo.

9. *From LZ77 to the Run-Length Burrows-Wheeler Transform, and Back*. Symposium on Combinatorial Pattern Matching (CPM), July 4-6, 2017, Warsaw.

10. *A Framework of Dynamic Data Structures for String Processing*. Symposium on Experimental Algorithms (SEA), June 21-23, 2017, King's College London.

11. *Space-Efficient Re-Pair Compression*. Data Compression Conference (DCC), March 26-29, 2017, Snowbird (UT).

12. *A Randomized LCE Data Structure* (best student talk award). Mathematical foundations for bioinformatics (MatBio), July 20, 2016, King's College London.

13. *Computing LZ77 in Run-Compressed Space* . Data Compression Conference (DCC), March 29 - April 1, 2016, Snowbird (UT).

14. *Fast Online Lempel-Ziv Factorization in Compressed Space*. String Processing and Information Retrieval (SPIRE), September 3, 2015, King's college London.

15. *Average linear time and compressed space construction of the Burrows-Wheeler transform*. 9th International Conference on Language and Automata Theory and Applications (LATA), March 2-6, 2015, Nice, France.

16. *Hashing and Indexing: succinct data structures and smoothed analysis*. Symposium on Algorithms and Computation (ISAAC), December 15-17, 2014, Jeonju, Korea.

Invitations

Invited Speaker

○ Invited Speaker at *18th Symposium on Experimental Algorithms* (**SEA 2020**), June 16-18 2020, Catania. `http://www.sea2020.dmi.unict.it/`.

○ Invited Speaker at *5th International Workshop on Innovative Algorithms for Big Data* (**IABD 2019**), October 31 2019, Kyoto. `https://sites.google.com/view/iabd2019/iabd2019`.

○ Invited Highlight Speaker at *30th Annual Symposium on Combinatorial Pattern Matching* (**CPM 2019**), June 18-20 2019, Pisa. `http://cpm2019.di.unipi.it/highlights/`.

○ Invited Speaker at *19th Italian Conference on Theoretical Computer Science* (**ICTCS 2018**), September 19-20 2018, Urbino. `http://www.sti.uniurb.it/events/ictcs2018/program.pdf`

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • ⌂ *github.com/nicolaprezza* *4/8*

International invitation-based seminars

o Dagstuhl Seminar "Computation over Compressed Structured Data" (16431), Schloss Dagstuhl, Germany, October 23-28 2016.

o Shonan Seminar "Computation over Compressed Structured Data", Shonan Village Center (SVC), Japan, October 9-12 2017.

o Dagstuhl Seminar "Synergies between Adaptive Analysis of Algorithms, Parameterized Complexity, Compressed Data Structures and Compressed Indices" (18281), Schloss Dagstuhl, Germany, July 8-13 2018.

o Dagstuhl Seminar "25 Years of the Burrows-Wheeler Transform" (19241), Schloss Dagstuhl, Germany, June 10-14 2019.

Visiting Researcher

o University of Palermo, visiting Dr. Gabriele Fici, September 24-27 2019.

o University of Chile, Santiago, Chile, visiting prof. Gonzalo Navarro, October 3-4 2019.

o University Diego Portales, Santiago, Chile, visiting prof. Travis Gagie, October 5-6 2019.

o Universit Tor Vergata and LUISS of Roma, visiting prof. Giuseppe Italiano, January 21-25 2019.

o University of Udine, visiting prof. Alberto Policriti, January 28-31 and January 25 February 1 2019.

o University of Helsinki, Finland, visiting Dr. Simon Puglisi, February 14-19 2016.

Committees

Committees

o Organizing Committee of **ALGO 2020**, an event including the *28th European Symposium on Algorithms* (**ESA 2020**) and other five satellite workshops, Pisa, September 7-10, 2020.

o Program Committee of *27th International Symposium on String Processing and Information Retrieval* (**SPIRE 2020**), Orlando, Florida, 13-15 Ottobre 2020.

o Program Committee of *31th Annual Symposium on Combinatorial Pattern Matching* (**CPM 2020**), Copenhagen, Denmark, June 17-19 2020.

o Organizing Committee of *30th International Workshop on Combinatorial Algorithms* (**IWOCA 2019**), Pisa, July 23-25 2019.

o Organizing Committee of *30th Annual Symposium on Combinatorial Pattern Matching* (**CPM 2019**), Pisa, June 18-20 2018.

o Co-chair of *13th Workshop on Compression, Text and Algorithms* (**WCTA 2018**), Lima, Perù, October 12 2018.

o Program Committee of *29th Annual Symposium on Combinatorial Pattern Matching* (**CPM 2018**), Qingdao, China, July 2-4 2018.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • ⌂ *github.com/nicolaprezza* *5/8*

List of all Publications

International journals

[1] Travis Gagie, Gonzalo Navarro, and Nicola Prezza, 2020. *Fully-Functional Suffix Trees and Optimal Text Searching in BWT-runs Bounded Space*. Journal of the ACM 67, 1, Article 2.

[2] Antonio Fariña, Miguel A. Martinez-Prieto, Francisco Claude, Gonzalo Navarro, Juan J. Lastra-Díaz [r], Nicola Prezza [r], and Diego Seco [r], 2019 ([r] = reproducibility reviewers). *On the Reproducibility of Experiments of Indexing Repetitive Document Collections*. Information Systems, Volume 83, Pages 181-194.

[3] Nicola Prezza, Nadia Pisanti, Marinella Sciortino, and Giovanna Rosone, 2019. *SNPs detection by eBWT positional clustering*. Algorithms for Molecular Biology, 14(1), p.3.

[4] Gonzalo Navarro and Nicola Prezza, 2019. *Universal compressed text indexing*. Theoretical Computer Science, 762, pp.41-50.

[5] Philip Bille, Travis Gagie, Inge Li Gørtz, and Nicola Prezza, 2018. *A separation between RLSLPs and LZ77*. Journal of Discrete Algorithms, 50, pp.36-39.

[6] Karin Engström, Tomasz K. Wojdacz, Francesco Marabita, Philip Ewels, Max Käller, Francesco Vezzi, Nicola Prezza, Joel Gruselius, Marie Vahter, and Karin Broberg, 2017. *Transcriptomics and methylomics of CD4-positive T cells in arsenic-exposed women*. Archives of toxicology 91, no. 5 (2017): 2067-2078.

[7] Alberto Policriti and Nicola Prezza, 2018. *LZ77 computation based on the run-length encoded BWT*. Algorithmica. 80(7), pp.1986-2011.

[8] Nicola Prezza, Francesco Vezzi, Max Käller, and Alberto Policriti, 2016. *Fast, accurate, and lightweight analysis of BS-treated reads with ERNE 2*. BMC bioinformatics 17, no. 4: 69.

[9] Alberto Policriti and Nicola Prezza, 2015. *Fast randomized approximate string matching with succinct hash data structures*. BMC bioinformatics 16, no. 9: S4.

International conferences

[10] Nicola Prezza and Giovanna Rosone, 2020. *Faster Online Computation of the Succinct Longest Previous Factor Array*. Accepted at Computability in Europe (CiE 2020), University of Salerno - Fisciano campus, Italy.

[11] Tomasz Kociumaka, Gonzalo Navarro and Nicola Prezza, 2020. *Towards a Definitive Measure of Repetitiveness*. Accepted a 14th Latin American Theoretical Informatics Symposium (LATIN 2020), Università di of São Paulo, Brasile.

[12] Philip Bille, Mikko Berggren Ettienne, Travis Gagie, Inge Li Gørtz, and Nicola Prezza, 2020. *Decompressing Lempel-Ziv Compressed Text*. Accettato a Data Compression Conference (DCC 2020).

[13] Jarno Alanko, Giovanna D'Agostino, Alberto Policriti, and Nicola Prezza, 2020. *Regular Languages meet Prefix Sorting*. To appear in the Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), Salt Lake City, 2020.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • ⌂ *github.com/nicolaprezza*   *6/8*

[14] Davide Della Giustina, Nicola Prezza and Rossano Venturini, 2019. *A new Linear-time Algorithm for Centroid Decomposition*. Proceedings of the 26th International Symposium on String Processing and Information Retrieval (SPIRE).

[15] Mai Alzamel, Alessio Conte, Daniele Greco, Veronica Guerrini, Costas Iliopoulos, Nadia Pisanti, Nicola Prezza, Giulia Punzi and Giovanna Rosone, 2019. *Online Algorithms on Antipowers and Antiperiods*. Proceedings of the 26th International Symposium on String Processing and Information Retrieval, (SPIRE).

[16] Nicola Prezza, 2019. *Optimal Rank and Select Queries on Dictionary-Compressed Text*. Proceedings of the 30th Annual Symposium on Combinatorial Pattern Matching, (CPM).

[17] Nicola Prezza and Giovanna Rosone, 2019. *Space-Efficient Computation of the LCP Array from the Burrows-Wheeler Transform*. Proceedings of the 30th Annual Symposium on Combinatorial Pattern Matching (CPM).

[18] Dominik Kempa and Nicola Prezza, 2018. *At the Roots of Dictionary Compression: String Attractors*. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, June 25-29, 2018, Los Angeles, CA. pp. 827-840. ACM.

[19] Travis Gagie, Gonzalo Navarro, and Nicola Prezza, 2018. *Optimal-Time Text Indexing in BWT-runs Bounded Space*. Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA). Pages 1459-1477.

[20] Nicola Prezza, 2018. *In-Place Sparse Suffix Sorting*. Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA). Pages 1496-1508.

[21] Travis Gagie, Gonzalo Navarro, and Nicola Prezza, 2018. *On the Approximation Ratio of Lempel-Ziv Parsing*. In *Proceedings of the 13th Latin American Symposium on Theoretical Informatics (LATIN)*. pp. 490-503. Springer, Cham.

[22] Nicola Prezza, Nadia Pisanti, Marinella Sciortino, and Giovanna Rosone, 2018. *Detecting mutations by eBWT*. In *18th International Workshop on Algorithms in Bioinformatics, (WABI)*. Vol. 113, pp. 1-15. Schloss Dagstuhl-Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing.

[23] Dominik Kempa, Alberto Policriti, Nicola Prezza, and Eva Rotenberg, 2018. *String Attractors: Verification and Optimization*. In *26th Annual European Symposium on Algorithms (ESA)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[24] Nicola Prezza, 2018. *On String Attractors*. Proceedings of the 19th Italian Conference on Theoretical Computer Science (ICTCS), Urbino, Italy. September 18-20. Pages 12-16.

[25] Philip Bille, Inge Li Gørtz, and Nicola Prezza, 2017. *Space-Efficient Re-Pair Compression*. In *Proceedings of the Data Compression Conference (DCC)*. pp. 171-180. IEEE.

[26] Nicola Prezza, 2017. *A Framework of Dynamic Data Structures for String Processing*. Proceedings of the *16th International Symposium on Experimental Algorithms (SEA)*.

[27] Djamal Belazzougui, Fabio Cunial, Travis Gagie, Nicola Prezza, and Mathieu Raffinot, 2017. *Flexible indexing of repetitive collections*. In *Proceedings of the 13th Conference on Computability in Europe (CiE)*, pp. 162-174. Springer, Cham.

LUISS University, viale Romania 32 – 00197 Rome – Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • ⌕ *github.com/nicolaprezza* *7/8*

[28] Philip Bille, Anders Roy Christiansen, Nicola Prezza, and Frederik Rye Skjoldjensen, 2017. *Succinct Partial Sums and Fenwick Trees*. In *24th International Symposium on String Processing and Information Retrieval (SPIRE)*, pp. 91-96. Springer, Cham.

[29] Alberto Policriti and Nicola Prezza, 2017. *From LZ77 to the Run-Length Encoded Burrows-Wheeler Transform, and Back*. In *proceedings of the 28th Annual Symposium on Combinatorial Pattern Matching (CPM)*, 2017.

[30] Alberto Policriti and Nicola Prezza, 2016 *Computing LZ77 in run-compressed space*. In *Data Compression Conference (DCC)*, 2016, pp. 23-32. IEEE.

[31] Djamal Belazzougui, Fabio Cunial, Travis Gagie, Nicola Prezza, and Mathieu Raffinot, 2015. *Composite repetition-aware data structures*. In *Proceedings of the 26th Symposium on Combinatorial Pattern Matching (CPM)*, pp. 26-39. Springer, Cham.

[32] Alberto Policriti and Nicola Prezza, 2015. *Fast online lempel-ziv factorization in compressed space*. In *Proceedings of the 22nd International Symposium on String Processing and Information Retrieval (SPIRE)*, pp. 13-20. Springer, Cham.

[33] Alberto Policriti, Nicola Gigante, and Nicola Prezza, 2015. *Average linear time and compressed space construction of the Burrows-Wheeler transform*. In *9th International Conference on Language and Automata Theory and Applications (LATA)*, pp. 587-598. Springer, Cham.

[34] Alberto Policriti and Nicola Prezza, 2014. *Hashing and indexing: Succinct data structures and smoothed analysis*. In *25th International Symposium on Algorithms and Computation (ISAAC)*, pp. 157-168. Springer, Cham.

[35] Nicola Prezza, Cristian Del Fabbro, Francesco Vezzi, Emanuale De Paoli, and Alberto Policriti, 2012. *ERNE-BS5: aligning BS-treated sequences by multiple hits on a 5-letters alphabet*. In *Proceedings of the ACM conference on bioinformatics, computational biology and biomedicine (ACM-BCB)*, pp. 12-19. ACM.

PhD thesis

Nicola Prezza, 2017. *Compressed Computation for Text Indexing*. *Università di Udine*. Supervisor: Prof. Alberto Policriti. *Best scientific PhD thesis 2017* award at the university of Udine.

LUISS University, viale Romania 32 − 00197 Rome − Italy
✉ *nprezza@luiss.it* • 🌐 *nicolaprezza.github.io*
in *it.linkedin.com/in/nicola-prezza-573898b2* • *github.com/nicolaprezza* *8/8*