

## Curriculum vitae

### PERSONAL INFORMATION **Maciej Bartoszek, Ph.D., Eng.**

ul. ██████████ Warsaw (Poland)

+48 ██████████ [✉ maciej.bartoszek@mini.pw.edu.pl](mailto:maciej.bartoszek@mini.pw.edu.pl)

<http://www.bartoszek.rexamine.com>

Gender Male | Date of birth ██████████ 1989 | Nationality Polish

### WORK EXPERIENCE

#### March 2016 – Present **Research and Teaching Assistant**

Faculty of Mathematics and Information Science,  
Warsaw University of Technology  
ul. Koszykowa 75, 00-662 Warsaw (Poland)

- Taught subjects/languages: *R*, *Python*, *Algorithms and Data Structures*, *C++*, *SAS*,
- research: source code similarity, graphs similarity, clustering, hierarchical clustering
- co-author of a book *Data Processing and Analysis in Python (Przetwarzanie i analiza danych w języku Python)*, PWN, <https://ksiegarnia.pwn.pl/Przetwarzanie-i-analiza-danych-w-jezyku-Python,634359876,p.html>.

#### October 2016 – Present **Trainer**

Sages, <http://www.sages.com.pl/>,  
ul. Nowogrodzka 62c,  
02-002 Warsaw (Poland)

- Python Trainer at data science bootcamp: <https://kodolamacz.pl/bootcamp-datascience/>
- R trainer at data science postgraduate studies in Faculty of Electronics and Information Technology, Warsaw University of Technology, <http://datascience.ii.pw.edu.pl/>
- Author of popular science articles

#### July 2015 – December 2015 **Internship**

Systems Research Institute,  
Polish Academy of Sciences  
ul. Newelska 6, 01-447 Warsaw (Poland)

- part of Interdisciplinary Ph.D. Studies in Information Technology
- developing a new hierarchical clustering algorithm
- developing of the algorithm ended up writing an article in the Information Sciences journal (IF 4.832)

### EDUCATION

#### November 2018 **Ph.D. Thesis: “A source code similarity assessment system for functional programming languages based on machine learning and data aggregation methods”** graduated with distinction

Faculty of Mathematics and Information Science,  
Warsaw University of Technology, Warsaw (Poland)

[http://bartoszek.rexamine.com/wp-content/uploads/2018/10/Maciej\\_Bartoszek\\_doktorat.pdf](http://bartoszek.rexamine.com/wp-content/uploads/2018/10/Maciej_Bartoszek_doktorat.pdf)

2013–2015 **Ph.D. studies - Thesis Title: 'Anti-plagiarism system for R language'**  
 Institute of Computer Science , Polish Academy of Sciences, Warsaw (Poland)

2012–2013 **Master of Science in Computer Science** graduated with distinction  
 Faculty of Mathematics and Information Science,  
 Warsaw University of Technology, Warsaw (Poland)  
 – speciality “design of CAD/CAM systems”  
 – computer graphics  
 – physics simulations  
 – shaders  
 – geometric modeling

2008–2012 **Bachelor of Science in Computer Science** graduated with distinction  
 Faculty of Mathematics and Information Science,  
 Warsaw University of Technology, Warsaw Poland  
 – parallel programming  
 – algorithms and data structures  
 – CUDA  
 – business applications  
 – databases

## PERSONAL SKILLS

Mother tongue Polish

Other languages	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C1	C2	C1	C1	C2

Levels: A1/A2: Basic user - B1/B2: Independent user - C1/C2: Proficient user  
[Common European Framework of Reference \(CEF\) level](#)

Driving licence A, B

## HONORS AND GRANTS

October 2017 3rd degree Warsaw University of Technology rector's award for didactic achievements

2015–2017 Co-investigator in Research project 2014/13/D/HS4/01700 (NCN) Construction and analysis of methods of information resources producers' quality management.

October 2015 Distinction for high valuation by students for laboratory classes

2014/15 Scholarship for academic achievements at Institute of Computer Science, Polish Academy of Sciences

2013/14 Scholarship for academic achievements at Institute of Computer Science, Polish Academy of Sciences

2013 top 5% of students in the academic year 2012/2013 at the Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland

2012 top 5% of students in the academic year 2012/2013 at the Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland

## TRAININGS

- R training in Narodowy Bank Polski
- R programming
  - code profiling
  - parallel programming
  - Rcpp – C++ chunks
  - knitr – report creating
- Sages, kodolamacz.pl, Bootcamp Data Science
- vectorized calculations: Numpy
  - data frames, data engineering, Pandas
  - data visualizations, Matplotlib
  - data cleaning, reading files in different formats in Python: .xls, .xml, .json
  - web scraping in Python
- Data science postgraduate studies in Faculty of Electronics and Information Technology, Warsaw University of Technology, <http://datascience.ii.pw.edu.pl/>
- R programming
  - data frames, data engineering, dplyr
  - data visualizations, ggplot2
  - regex, operation on strings

## PROJECTS

- 2015 Distributed Calculation of R code using computer cluster
- 2014 Web interface for antiplagiarism system, <http://www.similar.rexamine.com>
- 2013 Physics implementation and project management in Mars Effect, mars lander simulator, <http://www.mini.pw.edu.pl/~porter/ppvr/MarsEffect/index.html>

## COURSES

- Institute of Computer Science, Polish Academy of Sciences
- Advanced Software Development
  - Data Analysis with R
  - Introduction to Machine Learning
  - Advanced Statistical Methods
  - Game and Decision Theory
  - Monte Carlo methods: Algorithms & applications
  - Advanced Topics in Machine Learning
  - Biological Inspirations in Computational Intelligence Algorithms
  - Mining massive datasets
  - Natural language processing
- SAS Institute
- Data processing with SAS (part I)

## SKILLS

- Programming languages and technologies
- C++, Boost, Qt
  - C#, WPF, WinForms, ASP.NET MVC, Visual Studio
  - R, Python
  - PHP
  - JavaScript
  - PostgreSQL, SQLite
  - Linux, Bash

- Scientific interests
- algorithms and data structures
  - machine learning
  - information retrieval
  - similarity measures
  - graphs, similarity of graphs
  - hierarchical clustering

## SOCIAL

---

January 2015 – December 2015 Chairman of Ph.D student council

## PUBLICATIONS

---

- Books**
- Gagolewski M., Bartoszuik M., Cena A., *Genie: Przetwarzanie i analiza danych w języku Python (Data Processing and Analysis in Python)*, Wydawnictwo Naukowe PWN, 2016 (accepted for publication)
- Journals**
- Gagolewski M., Bartoszuik M., Cena A., Genie: A new, fast, and outlier-resistant hierarchical clustering algorithm, *Information Sciences* **363**, 2016, pp. 8-23. doi:10.1016/j.ins.2016.05.003
- Proceedings**
- Bartoszuik M., Gagolewski M., *Binary aggregation functions in software plagiarism detection*, In: Proc. FUZZ-IEEE'17, IEEE, 2017, no. 8015582. doi:10.1109/FUZZ-IEEE.2017.8015582
  - Gagolewski M., Cena A., Bartoszuik M., *Hierarchical clustering via penalty-based aggregation and the Genie approach*, In: Torra V. et al. (Eds.), *Modeling Decisions for Artificial Intelligence (Lecture Notes in Artificial Intelligence 9880)*, Springer, 2016, pp. 191-202. doi:10.1007/978-3-319-45656-0\_16
  - Bartoszuik M., Beliakov G., Gagolewski M., James S., *Fitting aggregation functions to data: Part I – Linearization and regularization*, In: Carvalho J.P. et al. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II (Communications in Computer and Information Science 611)*, Springer, 2016, pp. 767-779. doi:10.1007/978-3-319-40581-0\_62
  - Bartoszuik M., Beliakov G., Gagolewski M., James S., *Fitting aggregation functions to data: Part II – Idempotentization*, In: Carvalho J.P. et al. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II (Communications in Computer and Information Science 611)*, Springer, 2016, pp. 767-779. doi:10.1007/978-3-319-40581-0\_63
  - Bartoszuik M., Gagolewski M., *Detecting similarity of R functions via a fusion of multiple heuristic methods*, In: Alonso J.M., Bustince H., Reformat M. (Eds.), *Proc. IFSA/EUSFLAT 2015*, Atlantis Press, 2015, pp. 419-426,
  - Bartoszuik M., *Solving systems of polynomial equations: a novel end condition and root computation method*, In: *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems*, M. Ganzha, L. Maciaszek, M. Paprzycki (eds). ACSIS, Vol. 2, pages 543–552 (2014), doi:10.15439/2014F183,
  - Bartoszuik M., Gagolewski M., *A fuzzy R code similarity detection algorithm*, In: Laurent A. et al. (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part III (CCIS 444)*, Springer-Verlag, Heidelberg, 2014, pp. 21-30.

Popular science articles  
and tutorials

- Bartoszek M., Trzeci sezon Black Mirror, 24 stycznia 2019, <https://bezprawnik.pl/trzeci-sezon-black-mirror/>
- Bartoszek M., Web scraping, 17 stycznia 2019, <https://kodolamacz.pl/blog/wyzwanie-python-7-web-scraping/>
- Bartoszek M., Wyjątki oraz operacje na plikach, 10 stycznia 2019, <https://kodolamacz.pl/blog/wyzwanie-python-6-wyj%C4%85tki-oraz-operacje-na-plikach/>
- Bartoszek M., Zaawansowane aspekty programowania obiektowego, 3 stycznia 2019, <https://kodolamacz.pl/blog/wyzwanie-python-5-zaawansowane-aspekty-programowania-obiektowego/>
- Bartoszek M., Programowanie obiektowe, 27 grudnia 2018, <https://kodolamacz.pl/blog/wyzwanie-python-4-programowanie-obiektowe/>
- Bartoszek M., Algorytmy i struktury danych, 20 grudnia 2018, <https://kodolamacz.pl/blog/wyzwanie-python-3-algorytmy-i-struktury-danych/>
- Bartoszek M., Podstawowe instrukcje, 13 grudnia 2018, <https://kodolamacz.pl/blog/wyzwanie-python-2-podstawowe-instrukcje/>
- Bartoszek M., Hello World, 6 grudnia 2018, <https://kodolamacz.pl/blog/wyzwanie-python-1-hello-world/>
- Bartoszek M., Jakie kompetencje powinien posiadać początkujący Data Scientist, 31 stycznia 2018, <https://nofluffjobs.com/blog/jakie-kompetencje-powinien-posiadac-poczatkujacy-data-scientist/>
- Bartoszek M., Data scientiści – kim są i ile zarabiają w Polsce i na świecie?, 8 grudnia 2017, <http://www.sages.com.pl/blog/data-scientisci-kim-sa-i-ile-zarabiaja-w-polsce-i-na-swiecie/>
- Bartoszek M., 3 najlepsze Python IDE dla Data Scientistów, magazyn Programista (8), 2017, s. 4-7
- Bartoszek M., Dziesięć najczęściej popełnianych błędów przez początkujących „Data Scientistów”, 17 października, 2017, źródło: <http://blog.pclab.pl/cotuzgrzyta/Dziesi%C4%99%C4%87.najcz%C4%99%C5%9Bciej.pope%C5%82nianych.b%C5%82%C4%99d%C3%B3w.przez.pocz%C4%85tkuj%C4%85cych.Data.Scientist%C3%B3w,863>