Prof. **Marek Gagolewski** 13 December 2024

Associate Professor in Data Science
Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland
ul. Koszykowa 75, 00-662 Warsaw, Poland, Math Building (MiNI), Room 550

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

Web: https://www.gagolewski.com
Email: marek⁎gagolewski.com
ORCID: 0000-0003-0637-6028
Nationality: Polish, Australian

---

1 Highlights

Researcher in Data Science (mathematical modelling of complex phenomena and developing *usable*, general-purpose algorithms and data analysis software)

- Research interests: machine learning, data clustering, data fusion, aggregation, *metrics, prototype learning, computational statistics, mathematical modelling for sports analytics, informetrics, economics, science of science, etc.
- Area editor (aggregation functions and data science) in *Fuzzy Sets and Systems*
- Author/editor of 99 publications, including journal papers in outlets such as *Proceedings of the National Academy of Sciences (PNAS)*, *Journal of Statistical Software*, *R Journal*, *Journal of Classification*, *Information Fusion*, *International Journal of Forecasting*, *Statistical Modelling*, *Physica A: Statistical Mechanics and Its Applications*, *Information Sciences*, *Knowledge-Based Systems*, *IEEE Transactions on Fuzzy Systems*, and *Journal of Informetrics*
- Current h-index = 20 (G**gle Scholar)

Free (Libre) and Open Source Data Analysis Software Developer

- Author and maintainer of the fast and robust *Genie* hierarchical clustering algorithm (see the Python and R package *genieclust*)
- Author and maintainer of *stringi* – one of the most often downloaded R packages (text/natural language processing; over 72,000,000 downloads)

Data Science, Machine Learning, and Statistical Computing Tutor & Trainer

- Current: Warsaw University of Technology (Warsaw, Poland)
- Past: Deakin University (Melbourne, Australia), Data Science Retreat (Berlin, Germany)
- Author of the open-access textbooks *Deep R Programming* and *Minimalist Data Wrangling with Python*

2 Qualifications

11.2020 — Deakin University, Melbourne, VIC, Australia
Graduate Certificate of Higher Education: Learning And Teaching (E575)

10.2017 — Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
DSc (*Habilitation* – Higher Doctorate) in Computer Science

12.2011 — Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
PhD in Computer Science

06.2008 — Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland
BEng and MSc (by Research) in Computer Science (cum laude) [GPA 4.964/5.000]

3 Employment History

04.2024 – … — Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland
Associate Professor in Data Science (05.2024 – …)
Vice-chair of the Teaching and Learning Committee for BSc and MSc in Data Science (04.2024 –)
First-Year Academic Liaison for BSc in Data Science (04.2024 –)

04.2024 – … — Department of Stochastic Methods, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Associate Professor (04.2024 – …)

09.2019 – 03.2024 — School of Information Technology, Deakin University, Melbourne, VIC, Australia
Senior Lecturer in Applied Artificial Intelligence/Data Science [50% research/40% teaching/10% service load; corresponds to an Associate Professor position in North American universities (tenured)] (09.2019 – 03.2024)
Deputy Course Director – BSc Degree in Data Science (03.2022 – 03.2024)

10.2008 – 09.2019 — Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland
Associate Professor in Data Science (01.2018 – 09.2019)
(Founding) Vice-chair of the Teaching and Learning Committee for BSc and MSc in Data Science (10.2016 – 09.2019)
Supervisor of the Data Science studies (01.2018 – 09.2019)
First-Year Academic Liaison for BSc in Data Science (10.2017 – 09.2019)
Assistant Professor (04.2012 – 12.2017)
Teaching and Research Assistant (10.2008 – 02.2012)

07.2008 – 08.2019 — Department of Stochastic Methods, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Associate Professor (04.2018 – 08.2019)
Assistant Professor (02.2012 – 03.2018)
Research Assistant (07.2008 – 01.2012)

4 Short-term Research Visits and Casual Employment

07.2014 – 07.2019 Data Science Retreat, Berlin, Germany
Python, R, and Data Science Tutor (19 batches)

07.2017 – 08.2017 Deakin University, Melbourne, VIC, Australia
School of Information Technology
Visiting Professor

04.2015 – 06.2015 IRAFM, University of Ostrava, Czechia
Visiting Professor

03.2013 – 06.2013 Slovak University of Technology in Bratislava, Slovakia
Visiting Professor

5 Scholarships and Awards

1. Ministry of Education and Science, Poland, award for significant achievements in teaching: design and implementation of a new innovative course of study – Master of Data Science – at the Faculty of Mathematics and Information Science, Warsaw University of Technology, 2022 (team)
2. Deakin University's School of IT best paper award, 2022
3. Warsaw University of Technology award for excellence in research, 2024 (with A. Cena, 1st degree)
4. Warsaw University of Technology award for excellence in research, 2022 (with M. Bartoszuk and A. Cena, 1st degree)
5. Warsaw University of Technology IDUB PW best paper award (3 papers), 2020
6. Warsaw University of Technology award for excellence in research, 2020 (individual, 1st degree)
7. Golden Chalk for teaching excellence, Faculty of Mathematics and Information Science, Warsaw University of Technology, 2019 (individual, 2nd degree, i.e., silver)
8. Warsaw University of Technology award for excellence in teaching, 2017 (with M. Bartoszuk and A. Cena, 3rd degree)
9. Ministry of Science and Higher Education, Poland, scholarship for young researchers, 2015 (36 months)
10. Warsaw University of Technology award for excellence in teaching, 2015 (with K. Bobecka-Wesołowska and P. Grzegorzewski, 3rd degree)
11. Foundation for Polish Science (FNP), scholarship for young researchers – START Program, 2013 (12 months)
12. Warsaw University of Technology award for excellence in research, 2012 (with P. Grzegorzewski, 1st degree)
13. Warsaw University of Technology award for excellence in research, 2010 (with P. Grzegorzewski, 1st degree)
14. Ministry of Science and Higher Education, Poland, scholarship for research achievements for students, 2007

6 Research Activities

6.1 Research Projects

1. Chief Investigator (02.2021–03.2024), Partner Investigator (04.2024–02.2025): Australian Research Council, 2021 ARC Discovery Project DP210100227, *Beyond black-box models: Interaction in eXplainable Artificial Intelligence*, Deakin University, Australia; Other CIs: Gleb Beliakov–lead, Simon James (Deakin), and Enrique Herrera-Viedma (University of Granada), 2021 (48 months)
2. Partner Investigator: The Czech Science Foundation (GAČR), research project 18-06915S, *New approaches to aggregation operators in analysis and processing of data*, University of Olomouc, Czechia; Lead CI: Radomír Halaš, 2018 (36 months)
3. Lead Chief Investigator: National Science Centre, Poland, research project NCN Sonata 2014/13/D/HS4/01700, *Construction and analysis of methods of information resources producers' quality management*, Systems Research Institute, Polish Academy of Sciences; PIs: Maciej Bartoszuk and Anna Cena, 2015 (30 months)

6.2 Research Students

Current PhD students:

1. Łukasz Brzozowski (PhD cand., since 2022; Warsaw University of Technology)

I was the principal supervisor of the following PhD students:

1. Jan Lasek (PhD – completed 2019; Warsaw University of Technology) – *New Data-Driven Rating Systems for Association Football*
2. Maciej Bartoszuk (PhD cum laude – completed 2018; Warsaw University of Technology) – *A Source Code Similarity Assessment System for Functional Programming Languages Based on Machine Learning and Data Aggregation Methods* (in Polish)
3. Anna Cena (PhD – completed 2018; Systems Research Institute, Polish Academy of Sciences) – *Adaptive Hierarchical Clustering Algorithms Based on Data Aggregation Methods* (in Polish)

I was the supervisor of the following MSc (by Research) students (Warsaw University of Technology):

1. Dawid Stelmach (MSc – completed 2020) – *Video Anomaly Detection as a One-Class-Classification Problem* (in Polish)
2. Maciej Kurek (MSc – completed 2020) – *Survival Analysis of the Time-to-Score in Sports* (in Polish)
3. Piotr Wawrzyniak (MSc – completed 2019) – *Boxing Board Results Prediction Based on Neural Networks* (in Polish)
4. Michał Hadryś (MSc – completed 2019) – *Comparison of models for match outcome prediction* (in Polish)
5. Piotr Smuda (MSc – completed 2018) – *A Music Recommendation System* (in Polish)
6. Mateusz Jabłoński (MSc – completed 2016) – *Dynamic Report Generation based on Jupyter Kernels* (in Polish)
7. Natalia Potocka (MSc – completed 2016) – *Text Clustering Based on String Metrics* (in Polish)
8. Piotr Frukacz (MSc – completed 2015) – *Mobile Salesman Assistant on Salesforce platform and Google API* (in Polish)
9. Norbert Ryciak (MSc – completed 2015) – *A Text Topic Modelling-Based Recommender System Utilising the Latent Dirichlet Allocation Method* (in Polish)
10. Emma Sanderson (MSc – completed 2015) – *New Methods for Calculating Optimal Safety Stocks at Procter&Gamble*
11. Dawid Janocha (MSc – completed 2014) – *Continuous Integration in Software Engineering* (in Polish)

6.3 Publications

6.3.1 Research Monographs and Textbooks (7)

1. Gagolewski, M., *Deep R Programming*, Melbourne, v1.0.0 edition, 2023, 456 pp., URL: https://deepr.gagolewski.com/
2. Gagolewski, M., *Minimalist Data Wrangling with Python*, Melbourne, v1.0.3 edition, 2023, 442 pp., URL: https://datawranglingpy.gagolewski.com/
3. Gagolewski, M., *Algorytmy i postawy programowania w języku C++ (Introduction to Algorithms and Programming in C++)*, Melbourne, v1.2.0 edition, 2022, 209 pp., URL: https://github.com/gagolews/aipp
4. Gagolewski, M., Bartoszuk, M., Cena, A., *Przetwarzanie i analiza danych w języku Python (Data Processing and Analysis in Python)*, Wydawnictwo Naukowe PWN, Warsaw, 2016, 369 pp., URL: https://github.com/gagolews/Analiza_danych_w_jezyku_Python
5. Gagolewski, M., *Programowanie w języku R. Analiza danych, obliczenia, symulacje (R Programming. Data Analysis, Computing, Simulations)*, Wydawnictwo Naukowe PWN, Warsaw, 2nd edition, 2016, 550 pp., URL: https://github.com/gagolews/Programowanie_w_jezyku_R
6. Gagolewski, M., *Data Fusion: Theory, Methods, and Applications*, Institute of Computer Science, Polish Academy of Sciences, Warsaw, 2015, 290 pp., URL: https://github.com/gagolews/datafusion
7. Grzegorzewski, P., Gagolewski, M., Bobecka-Wesołowska, K., *Wnioskowanie statystyczne z wykorzystaniem środowiska R (Statistical Inference with R)*, Politechnika Warszawska, Warsaw, 2014, 183 pp.

6.3.2 Edited Volumes (3)

8. Halaš, R., Gagolewski, M., Mesiar, R. (eds.), *New Trends in Aggregation Theory*, Springer, 2019, 348 pp.
9. Ferraro, M., Giordani, P., Vantaggi, B., Gagolewski, M., Gil, M., Grzegorzewski, P., Hryniewicz, O. (eds.), *Soft Methods for Data Science*, Springer, 2017, 535 pp.
10. Grzegorzewski, P., Gagolewski, M., Hryniewicz, O., Gil, M. (eds.), *Strengthening Links Between Data Analysis and Soft Computing*, Springer, 2015, 294 pp.

6.3.3 Journal Articles (54)

11. Gagolewski, M., Normalised clustering accuracy: An asymmetric external cluster validity measure, *Journal of Classification*, 2024
12. Gagolewski, M., Cena, A., Bartoszuk, M., Brzozowski, L., Clustering with minimum spanning trees: How good can it be?, *Journal of Classification*, 2024
13. Bertoli-Barsotti, L., Gagolewski, M., Siudem, G., Żogała-Siudem, B., Equivalence of inequality indices in the three-dimensional model of informetric impact, *Journal of Informetrics* **18**(4), 101566, 2024
14. Bertoli-Barsotti, L., Gagolewski, M., Siudem, G., Żogała-Siudem, B., Gini-stable Lorenz curves and their relation to the generalised Pareto distribution, *Journal of Informetrics* **18**(2), 101499, 2024
15. Wu, J-Z., Beliakov, G., James, S., Gagolewski, M., Random generation of linearly constrained fuzzy measures and domain coverage performance evaluation, *Information Sciences* **659**, 120080, 2024
16. Gagolewski, M., Cena, A., James, S., Beliakov, G., Hierarchical clustering with OWA-based linkages, the Lance–Williams formula, and dendrogram inversions, *Fuzzy Sets and Systems* **473**, 108740, 2023
17. Boczek, M., Gagolewski, M., Kaluszka, M., Okolewski, A., A benchmark-type generalization of the Sugeno integral with applications in bibliometrics, *Fuzzy Sets and Systems* **466**, 108479, 2023
18. Żogała-Siudem, B., Cena, A., Siudem, G., Gagolewski, M., Interpretable reparameterisations of citation models, *Journal of Informetrics* **17**(1), 101355, 2023
19. Gagolewski, M., A framework for benchmarking clustering algorithms, *SoftwareX* **20**, 101270, 2022, URL: https://clustering-benchmarks.gagolewski.com/
20. Siudem, G., Nowak, P., Gagolewski, M., Power laws, the Price Model, and the Pareto type-2 distribution, *Physica A: Statistical Mechanics and its Applications* **606**, 128059, 2022
21. Gagolewski, M., stringi: Fast and portable character string processing in R, *Journal of Statistical Software* **103**(2), 1–59, 2022, URL: https://stringi.gagolewski.com/
22. Beliakov, G., Gagolewski, M., James, S., Reduction of variables and constraints in fitting antibuoyant fuzzy measures to data using linear programming, *Fuzzy Sets and Systems* **451**, 266–284, 2022
23. Geras, A., Siudem, G., Gagolewski, M., Time to vote: Temporal clustering of user activity on Stack Overflow, *Journal of the Association for Information Science and Technology* **73**(12), 1681–1691, 2022
24. Gagolewski, M., Żogała-Siudem, B., Siudem, G., Cena, A., Ockham's index of citation impact, *Scientometrics* **127**, 2829–2845, 2022
25. Mrowiński, M., Gagolewski, M., Siudem, G., Accidentality in journal citation patterns, *Journal of Informetrics* **16**(4), 101341, 2022
26. Cena, A., Gagolewski, M., Siudem, G., Żogała-Siudem, B., Validating citation models by proxy indices, *Journal of Informetrics* **16**(2), 101267, 2022
27. Gagolewski, M., Bartoszuk, M., Cena, A., Are cluster validity measures (in)valid?, *Information Sciences* **581**, 620–636, 2021, URL: https://github.com/gagolews/optim_cvi
28. Beliakov, G., Gagolewski, M., James, S., Hierarchical data fusion processes involving the Möbius representation of capacities, *Fuzzy Sets and Systems* **433**, 1–21, 2022
29. Bartoszuk, M., Gagolewski, M., T-norms or t-conorms? How to aggregate similarity degrees for plagiarism detection, *Knowledge-Based Systems* **231**, 107427, 2021
30. Lasek, J., Gagolewski, M., Interpretable sports team rating models based on the gradient descent algorithm, *International Journal of Forecasting* **37**(3), 1061–1071, 2021

31. Gagolewski, M., genieclust: Fast and robust hierarchical clustering, *SoftwareX* **15**, 100722, 2021, URL: https://genieclust.gagolewski.com/
32. Pérez-Fernández, R., Gagolewski, M., De Baets, B., On the aggregation of compositional data, *Information Fusion* **73**, 103–110, 2021
33. Beliakov, G., Gagolewski, M., James, S., DC optimization for constructing discrete Sugeno integrals and learning nonadditive measures, *Optimization* **69**(12), 2515–2534, 2020
34. Bartoszuk, M., Gagolewski, M., SimilaR: R Code Clone and Plagiarism Detection, *R Journal* **12**(1), 367–385, 2020, URL: https://CRAN.R-project.org/package=SimilaR
35. Siudem, G., Żogała-Siudem, B., Cena, A., Gagolewski, M., Three dimensions of scientific impact, *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **117**, 13896–13900, 2020
36. Coroianu, L., Fullér, R., Gagolewski, M., James, S., Constrained ordered weighted averaging aggregation with multiple comonotone constraints, *Fuzzy Sets and Systems* **395**, 21–39, 2020
37. Cena, A., Gagolewski, M., Genie+OWA: Robustifying hierarchical clustering with OWA-based linkages, *Information Sciences* **520**, 324–336, 2020
38. Gagolewski, M., Pérez-Fernández, R., De Baets, B., An inherent difficulty in the aggregation of multidimensional data, *IEEE Transactions on Fuzzy Systems* **28**, 602–606, 2020
39. Beliakov, G., Gagolewski, M., James, S., Robust fitting for the Sugeno integral with respect to general fuzzy measures, *Information Sciences* **514**, 449–461, 2020
40. Geras, A., Siudem, G., Gagolewski, M., Should we introduce a dislike button for academic papers?, *Journal of the Association for Information Science and Technology* **71**(2), 221–229, 2020
41. Coroianu, L., Gagolewski, M., Grzegorzewski, P., Piecewise linear approximation of fuzzy numbers: Algorithms, arithmetic operations and stability of characteristics, *Soft Computing* **23**(19), 9491–9505, 2019, URL: https://CRAN.R-project.org/package=FuzzyNumbers
42. Beliakov, G., Gagolewski, M., James, S., Aggregation on ordinal scales with the Sugeno integral for biomedical applications, *Information Sciences* **501**, 377–387, 2019
43. Pérez-Fernández, R., De Baets, B., Gagolewski, M., A taxonomy of monotonicity properties for the aggregation of multidimensional data, *Information Fusion* **52**, 322–334, 2019
44. Gagolewski, M., James, S., Beliakov, G., Supervised learning to aggregate data with the Sugeno integral, *IEEE Transactions on Fuzzy Systems* **27**(4), 810–815, 2019
45. Beliakov, G., Gagolewski, M., James, S., Pace, S., Pastorello, N., Thilliez, E., Vasa, R., Measuring traffic congestion: An approach based on learning weighted inequality, spread and aggregation indices from comparison data, *Applied Soft Computing* **67**, 910–919, 2019
46. Lasek, J., Gagolewski, M., The efficacy of league formats in ranking teams, *Statistical Modelling* **18**(5–6), 411–435, 2018
47. Gagolewski, M., Penalty-based aggregation of multidimensional data, *Fuzzy Sets and Systems* **325**, 4–20, 2017
48. Mesiar, R., Gagolewski, M., H-index and other Sugeno integrals: Some defects and their compensation, *IEEE Transactions on Fuzzy Systems* **24**(6), 1668–1672, 2016
49. Beliakov, G., Gagolewski, M., James, S., Penalty-based and other representations of economic inequality, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **24(Suppl.1)**, 1–23, 2016
50. Gagolewski, M., Bartoszuk, M., Cena, A., Genie: A new, fast, and outlier-resistant hierarchical clustering algorithm, *Information Sciences* **363**, 8–23, 2016, URL: https://genieclust.gagolewski.com/
51. Lasek, J., Szlavik, Z., Gagolewski, M., Bhulai, S., How to improve a team's position in the FIFA ranking – A simulation study, *Journal of Applied Statistics* **43**(7), 1349–1368, 2016
52. Żogała-Siudem, B., Siudem, G., Cena, A., Gagolewski, M., Agent-based model for the bibliometric h-index – Exact solution, *European Physical Journal B* **89**(21), 2016
53. Cena, A., Gagolewski, M., Mesiar, R., Problems and challenges of information resources producers' clustering, *Journal of Informetrics* **9**(2), 2015
54. Cena, A., Gagolewski, M., OM3: Ordered maxitive, minitive, and modular aggregation operators – Axiomatic and probabilistic properties in an arity-monotonic setting, *Fuzzy Sets and Systems* **264**, 138–159, 2015

55. Gagolewski, M., Spread measures and their relation to aggregation functions, *European Journal of Operational Research* **241**(2), 469–477, 2015
56. Gagolewski, M., Mesiar, R., Monotone measures and universal integrals in a uniform framework for the scientific impact assessment problem, *Information Sciences* **263**, 166–174, 2014
57. Gagolewski, M., On the relationship between symmetric maxitive, minitive, and modular aggregation operators, *Information Sciences* **221**, 170–180, 2013
58. Coroianu, L., Gagolewski, M., Grzegorzewski, P., Nearest piecewise linear approximation of fuzzy numbers, *Fuzzy Sets and Systems* **233**, 26–51, 2013, URL: https://CRAN.R-project.org/package=FuzzyNumbers
59. Gagolewski, M., Scientific impact assessment cannot be fair, *Journal of Informetrics* **7**(4), 792–802, 2013
60. Gagolewski, M., Mesiar, R., Aggregating different paper quality measures with a generalized h-index, *Journal of Informetrics* **6**(4), 566–579, 2012
61. Gagolewski, M., Bibliometric impact assessment with R and the CITAN package, *Journal of Informetrics* **5**(4), 678–692, 2011, URL: https://CRAN.R-project.org/package=CITAN
62. Gagolewski, M., Grzegorzewski, P., Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals, *International Journal of Approximate Reasoning* **52**(9), 1312–1324, 2011
63. Gagolewski, M., Grzegorzewski, P., A geometric approach to the construction of scientific impact indices, *Scientometrics* **81**(3), 617–634, 2009
64. Rowiński, T., Gagolewski, M., Preferencje i postawy wobec pomocy online (Attitudes towards online counselling and psychotherapy), *Studia Psychologica UKSW* **7**, 195–210, 2007

6.3.4 Papers in Edited Volumes and Proceedings (35)

65. Beliakov, G., James, S., Wu, J-Z., Gagolewski, M., *A Tractable Approach to Fitting the Choquet Integral for Explainable Prediction and Analysis*, in: *Proc. 2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2024, pp. 10611753
66. Coroianu, L., Gagolewski, M., *Penalty-based data aggregation in real normed vector spaces*, in: Halaš, R. et al. (eds.), *New Trends in Aggregation Theory*, Springer, 2019, pp. 160–171
67. Beliakov, G., Gagolewski, M., James, S., *Least median of squares (LMS) and least trimmed squares (LTS) fitting for the weighted arithmetic mean*, in: Medina, J. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*, Springer, 2018, pp. 367–378
68. Gagolewski, M., James, S., *Fitting symmetric fuzzy measures for discrete Sugeno integration*, in: Kacprzyk, J. et al. (eds.), *Advances in Fuzzy Logic and Technology 2017*, Springer, 2018, pp. 104–116
69. Bartoszuk, M., Gagolewski, M., *Binary aggregation functions in software plagiarism detection*, in: *Proc. 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2017, pp. 8015582
70. Cena, A., Gagolewski, M., *OWA-based linkage and the Genie correction for hierarchical clustering*, in: *Proc. 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, 2017, pp. 8015652
71. Gagolewski, M., Cena, A., Bartoszuk, M., *Hierarchical clustering via penalty-based aggregation and the Genie approach*, in: Torra, V. et al. (eds.), *Modeling Decisions for Artificial Intelligence*, Springer, 2016, pp. 191–202
72. Cena, A., Gagolewski, M., *Fuzzy k-minpen clustering and k-nearest-minpen classification procedures incorporating generic distance-based penalty minimizers*, in: Carvalho, J. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II*, Springer, 2016, pp. 445–456
73. Bartoszuk, M., Beliakov, G., Gagolewski, M., James, S., *Fitting aggregation functions to data: Part I – Linearization and regularization*, in: Carvalho, J. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II*, Springer, 2016, pp. 767–779
74. Bartoszuk, M., Beliakov, G., Gagolewski, M., James, S., *Fitting aggregation functions to data: Part II – Idempotization*, in: Carvalho, J. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II*, Springer, 2016, pp. 780–789
75. Bartoszuk, M., Gagolewski, M., *Detecting similarity of R functions via a fusion of multiple heuristic methods*, in: Alonso, J., Bustince, H., Reformat, M. (eds.), *Proc. IFSA/EUSFLAT'15*, Atlantis Press, 2015, pp. 419–426

76. Cena, A., Gagolewski, M., *A K-means-like algorithm for informetric data clustering*, in: Alonso, J., Bustince, H., Reformat, M. (eds.), *Proc. IFSA/EUSFLAT'15*, Atlantis Press, 2015, pp. 536–543
77. Gagolewski, M., *Sugeno integral-based confidence intervals for the theoretical h-index*, in: Grzegorzewski, P. et al. (eds.), *Strengthening Links Between Data Analysis and Soft Computing*, Springer, 2015, pp. 233–240
78. Gagolewski, M., *Normalized WD_pWAM and WD_pOWA spread measures*, in: Alonso, J., Bustince, H., Reformat, M. (eds.), *Proc. IFSA/EUSFLAT'15*, Atlantis Press, 2015, pp. 210–216
79. Gagolewski, M., Lasek, J., *The use of fuzzy relations in the assessment of information resources producers' performance*, in: *Proc. 7th IEEE International Conference Intelligent Systems IS'2014, Vol. 2: Tools, Architectures, Systems, Applications*, Springer, 2015, pp. 289–300
80. Gagolewski, M., Lasek, J., *Learning experts' preferences from informetric data*, in: Alonso, J., Bustince, H., Reformat, M. (eds.), *Proc. IFSA/EUSFLAT'15*, Atlantis Press, 2015, pp. 484–491
81. Lasek, J., Gagolewski, M., *Estimation of tournament metrics for association football league formats*, in: *Selected problems in information technologies (Proc. ITRIA'15 vol. 2)*, Institute of Computer Science, Polish Academy of Sciences, 2015, pp. 67–78
82. Lasek, J., Gagolewski, M., *The winning solution to the AAIA'15 Data Mining Competition: Tagging firefighter activities at a fire scene*, in: Ganzha, M., Maciaszek, L., Paprzycki, M. (eds.), *Proc. FedCSIS'15*, IEEE, 2015, pp. 375–380
83. Cena, A., Gagolewski, M., *Clustering and aggregation of informetric data sets*, in: *Computational methods in data analysis (Proc. ITRIA'15 vol. 1)*, Institute of Computer Science, Polish Academy of Sciences, 2015, pp. 5–26
84. Gagolewski, M., *Some issues in aggregation of multidimensional data*, in: Baczyński, M., De Baets, B., Mesiar, R. (eds.), *Proc. 8th International Summer School on Aggregation Operators (AGOP 2015)*, University of Silesia, 2015, pp. 127–132
85. Cena, A., Gagolewski, M., *Aggregation and soft clustering of informetric data*, in: Baczyński, M., De Baets, B., Mesiar, R. (eds.), *Proc. 8th International Summer School on Aggregation Operators (AGOP 2015)*, University of Silesia, 2015, pp. 79–84
86. Coroianu, L., Gagolewski, M., Grzegorzewski, P., Adabitabar Firozja, M., Houlari, T., *Piecewise linear approximation of fuzzy numbers preserving the support and core*, in: Laurent, A. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part II*, Springer, 2014, pp. 244–254
87. Bartoszuk, 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*, Springer, 2014, pp. 21–30
88. Gagolewski, M., Dębski, M., Nowakiewicz, M., *Efficient algorithm for computing certain graph-based monotone integrals: The l_p-indices*, in: Mesiar, R., Bacigal, T. (eds.), *Proc. Uncertainty Modeling*, STU Bratislava, 2013, pp. 17–23
89. Cena, A., Gagolewski, M., *OM3: Ordered maxitive, minitive, and modular aggregation operators – Part I: Axiomatic analysis under arity-dependence*, in: Bustince, H. et al. (eds.), *Aggregation Functions in Theory and in Practise*, Springer, 2013, pp. 93–103
90. Cena, A., Gagolewski, M., *OM3: Ordered maxitive, minitive, and modular aggregation operators – Part II: A simulation study*, in: Bustince, H. et al. (eds.), *Aggregation Functions in Theory and in Practise*, Springer, 2013, pp. 105–115
91. Gagolewski, M., *Statistical hypothesis test for the difference between Hirsch indices of two Pareto-distributed random samples*, in: Kruse, R. et al. (eds.), *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, Springer, 2013, pp. 359–367
92. Gagolewski, M., *On the relation between effort-dominating and symmetric minitive aggregation operators*, in: Greco, S. et al. (eds.), *Advances in Computational Intelligence, Part III*, Springer, 2012, pp. 276–285
93. Gagolewski, M., Grzegorzewski, P., *Axiomatic characterizations of (quasi-) L-statistics and S-statistics and the Producer Assessment Problem*, in: Galichet, S. et al. (eds.), *Proc. EUSFLAT/LFA'11*, Atlantis Press, 2011, pp. 53–58
94. Rowiński, T., Gagolewski, M., *Internet a kryzys*, in: Jankowska, M., Starzomska, M. (eds.), *Kryzys: Pułapka czy szansa?*, WN Akapit, Warsaw, 2011, pp. 211–224
95. Gagolewski, M., Grzegorzewski, P., *Arity-monotonic extended aggregation operators*, in: Hüllermeier, E. et al. (eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Springer, 2010, pp. 693–702
96. Gagolewski, M., Grzegorzewski, P., *S-statistics and their basic properties*, in: Borgelt, C. et al. (eds.), *Combining Soft Computing and Statistical Methods in Data Analysis*, Springer, 2010, pp. 281–288

97. Gagolewski, M., Grzegorzewski, P., *Metody i problemy naukometrii (Methods and problems of scientometrics)*, in: Rowiński, T., Tadeusiewicz, R. (eds.), *Psychologia i informatyka. Synergia i kontradykcje*, Wyd. UKSW, Warsaw, 2010, pp. 103–125
98. Gagolewski, M., Grzegorzewski, P., *O pewnym uogólnieniu indeksu Hirscha*, in: Kawalec, P., Lipski, P. (eds.), *Kadry i infrastruktura nowoczesnej nauki: Teoria i praktyka, Proc. 1st Intl. Conf. Zarządzanie Nauką*, Wydawnictwo Lubelskiej Szkoły Biznesu, Lublin, 2009, pp. 15–29
99. Gagolewski, M., Grzegorzewski, P., *Possible and necessary h-indices*, in: Carvalho, J. et al. (eds.), *Proc. IFSA/EUSFLAT'09*, IFSA, 2009, pp. 1691–1695

6.4 Talks (Conferences, Seminars, etc.)

6.4.1 Invited Plenary Lectures and Tutorials

1. *Jak badać metody analizy skupień? (Methods for analysis of cluster analysis methods)*, XLIX Konferencja „Statystyka Matematyczna – Chęciny 2024", Chęciny, Poland, 09–12.12.2024 (In Polish)
2. *Clustering and aggregation*, 16th International Conference on Fuzzy Set Theory and Applications – FSTA 2022, Liptovský Ján, Slovakia, 04.02.2022 (online)
3. *Clustering on MSTs*, International Student Conference on Applied Mathematics and Informatics ISCAMI'18, Malenovice, Czechia, 10–13.05.2018
4. *Stochastic properties of and agent-based models for the Hirsch index and other discrete Sugeno integrals*, 14th International Conference on Fuzzy Set Theory and Applications – FSTA 2018, Liptovský Ján, Slovakia, 02.02.2018
5. *Aggregation of multidimensional data: A review*, 9th International Summer School on Aggregation Operators – AGOP 2017, Skövde, Sweden, 21.06.2017
6. *Penalty-based fusion of complex data, computational aspects, and applications*, International Symposium on Aggregation and Structures – ISAS 2016, University of Luxembourg, 06.07.2016

6.4.2 Other Invited

7. *R package* stringi, Text Analysis Developers' Workshop 2018, New York University, New York, NY, US, 20–21.04.2018
8. *Algorytmy analizy skupień oparte na MST*, Studencka konferencja zastosowań matematyki DwuMIan'18, Warsaw, Poland, 24.03.2018
9. *R package* stringi, Text Analysis R Developers' Workshop 2017, London School of Economics, London, England, 21–22.04.2017
10. Genie*: A new, fast, and outlier-resistant hierarchical clustering algorithm and its R interface*, European R Users Meeting, Poznań, Poland, 14.10.2016
11. *Can the scientific assessment process be fair?*, Workshop on Research Evaluation, Free University of Bozen-Bolzano, Italy, 10.05.2013

6.4.3 Seminars

12. *Aggregation of multidimensional data: A review*, School of Information Technology, Deakin University, Melbourne-Burwood, VIC, Australia, 21.07.2017
13. *Genie: Nowy, szybki i odporny algorytm analizy skupień*, Seminarium IBS PAN, Warszawa, Poland, 23.05.2017
14. *Agregacja danych: Teoria, metody i zastosowania*, Wykład dla słuchaczy Studiów Doktoranckich IBS PAN, Warszawa, Poland, 05.03.2016
15. `^(R|ICU|i18n|regex)+$`, Seminarium Matematyczne Metody Informatyki, Instytut Matematyki, University of Silesia, Katowice, Poland, 20.04.2015
16. *Data aggregation from an algorithmic perspective*, IRAFM Seminar, University of Ostrava, Czechia, 04.06.2015
17. *Indeks Hirscha i okolice*, Seminarium CeON, ICM UW, Warsaw, Poland, 12.03.2014
18. *Scientific impact assessment – State of the art: Agregačné funkcie: teória a aplikácie* (*Aggregation functions: theory and applications*), Seminár z modelovania neurčitosti, Katedra matematiky a deskriptívnej geometrie, SvF STU, Bratislava, Slovakia, 17.04.2013

19. *Penalty-based data aggregation in real normed vector spaces*, 10th International Summer School on Aggregation Operators (AGOP), Olomouc, Czechia, 1–4.07.2019
20. *Fitting symmetric fuzzy measures for discrete Sugeno integration*, 10th International Conference of EUSFLAT, Warsaw, Poland, 11–15.09.2017
21. *Binary aggregation functions in software plagiarism detection*, IEEE International Conference on Fuzzy Systems (IEEE FUZZ'17), Naples, Italy, 9–12.07.2017
22. *Binary aggregation functions in software plagiarism detection*, 3rd International Symposium on Fuzzy Sets and Uncertainty Modeling (ISFS 2017), Rzeszów, Poland, 19–20.05.2017
23. *Hierarchical clustering via penalty-based aggregation and the Genie approach*, 13th International Conference on Modeling Decisions for Artificial Intelligence (MDAI), Sant Julià de Lòria, Andorra, 20.09.2016
24. *Fitting aggregation functions to data: Part I – Linearization and regularization*, 16th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Eindhoven, The Netherlands, 23.06.2016
25. *Some issues in aggregation of multidimensional data*, 8th International Summer School on Aggregation Operators (AGOP), Katowice, Poland, 07.07.2015
26. *Normalized WD_pWAM and WD_pOWA spread measures*, International Conference of IFSA/EUSFLAT 2015, Gijon, Spain, 02.07.2015
27. *Sugeno integral-based confidence intervals for the theoretical h-index*, 7th International Conference on Soft Methods in Probability and Statistics (SMPS), Warsaw, Poland, 24.09.2014
28. *OM3: Ordered maxitive, minitive, and modular aggregation operators – Part I: Axiomatic analysis under arity-dependence*, 7th International Summer School on Aggregation Operators (AGOP), Pamplona, Spain, 16–19.07.2013
29. *Statistical hypothesis test for the difference between Hirsch indices of two Pareto-distributed random samples*, 6th International Conference on Soft Methods in Probability and Statistics (SMPS), Konstanz, Germany, 04–06.10.2012
30. *On the relation between effort-dominating and symmetric minitive aggregation operators*, 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Catania, Italy, 09–13.07.2012
31. *Porównanie wybranych estymatorów teoretycznego indeksu Hirscha*, XXXVII Konferencja Statystyka Matematyczna, Wisła, Poland, 05–09.12.2011
32. *Axiomatic characterizations of (quasi-) L-statistics and S-statistics and the Producer Assessment Problem*, 7th International Conference of EUSFLAT/LFA, Aix-Les-Bains, France, 18–22.07.2011
33. *Podstawowe właściwości S-statystyk*, XXXVI Konferencja Statystyka Matematyczna, Wisła, Poland, 06–10.12.2010
34. *S-Statistics and their basic properties*, 5th International Conference on Soft Methods in Probability and Statistics (SMPS), Oviedo, Spain, 28.09–01.10.2010
35. *Arity-monotonic extended aggregation operators*, 13th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Dortmund, Germany, 28.06–02.07.2010
36. *Uogólniony indeks Hirscha a dwupróbkowe testy dla rodziny rozkładów Pareto II rodzaju*, XXXV Konferencja Statystyka Matematyczna, Wisła, Poland, 07–11.12.2009
37. *O pewnym uogólnieniu indeksu Hirscha*, 1st International Conference on "Scientific Management", Lublin, Poland, 20–22.11.2009
38. *Possible and necessary h-indices*, 6th International Conference of IFSA/EUSFLAT, Lisbon, Portugal, 20–24.07.2009

7 Reviewing and Other Academic Activities

- Area editor (aggregation functions and data science) in *Fuzzy Sets and Systems* (2021–)
- Member of the Research Council (Computer Science and Telecommunication); Warsaw University of Technology (2019–)
- Member of the Scientific Council; Systems Research Institute, Polish Academy of Sciences (2011–)

- Member of the Faculty Council; Faculty of Mathematics and Information Science, Warsaw University of Technology (2017–)
- Research Seminar "MADAM: Methods for Analysis of Data – Algorithms and Modelling" – Chair (https://madam.mini.pw.edu.pl/) (2017–)
- Scientific program committee member/chair for:

    1. 20th Conference on Complex Systems (CCS 2024), London, England
    2. 13th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT/AGOP/FQAS 2023), Palma, Mallorca, Spain
    3. 11th International Summer School on Aggregation Operators (AGOP 2021), Bratislava, Slovakia – Program Chair
    4. 19th World Congress of the International Fuzzy Systems Association and 12th Conference of the European Society for Fuzzy Logic and Technology (IFSA/EUSFLAT 2021), Bratislava, Slovakia
    5. 10th International Summer School on Aggregation Operators (AGOP 2019), Olomouc, Czechia
    6. 11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019), Prague, Czechia
    7. 2nd International Symposium on Aggregation and Structures (ISAS 2018), Valladolid, Spain
    8. 3rd Conference on Information Technology, Systems Research and Computational Physics (ITSRCP'18), Cracow, Poland
    9. 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA/SCIS 2017), Otsu, Japan
    10. 1st International Symposium on Aggregation and Structures (ISAS 2016), Luxembourg
    11. 16th World Congress of the International Fuzzy Systems Association and 9th Conference of the European Society for Fuzzy Logic and Technology (IFSA/EUSFLAT 2015), Gijon, Spain

- Special session organiser at:

    1. IEEE World Congress on Computational Intelligence (WCCI 2020), Glasgow (UK) – FUZZ-IEEE-6 Special Session *Aggregation Structures: New Trends and Applications*
    2. 10th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2017), Warsaw, Poland – Special Session *Algorithms for Data Aggregation and Fusion*
    3. 16th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2016), Eindhoven, The Netherlands – Special Session *Computational Aspects of Data Aggregation and Complex Data Fusion*

- Organising committee member/chair for:

    1. 10th International Summer School on Aggregation Operators (AGOP 2019), Olomouc, Czechia – Conference Chair
    2. 10th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2017), Warsaw, Poland – Stream on Data Analysis Coordinator
    3. 8th International Conference Soft Methods in Probability and Statistics (SMPS 2016), Rome, Italy
    4. 8th International Summer School on Aggregation Operators (AGOP 2015), Katowice, Poland
    5. 7th International Conference Soft Methods in Probability and Statistics – SMPS 2014, Warsaw, Poland
    6. 37th Conference *Statystyka Matematyczna – Wisła 2011*, Poland

- Reviewer of research project proposals for:

    1. Australian Research Council (ARC) (2)
    2. The National Fund for Scientific and Technological Development of Chile (Fondo Nacional de Desarrollo Científico y Tecnológico – FONDECYT) (1)
    3. Slovak Research and Development Agency (Agentúra na podporu výskumu a vývoja – APVV) (4)

- Reviewer of PhD theses of:

    1. dr Jana Borzová, Faculty of Science, P. J. Šafárik University in Košice, Slovakia, 2018
    2. Hossein Yazdani, Faculty of Electronics, Wrocław University of Science and Technology, Poland; 2018 and 2020 (re-review)

- PhD/DSc committee member of:

    1. dr Tomasz Rybotycki, Systems Research Institute, Polish Academy of Sciences, 2023
    2. dr Weronika Gutfeter, Warsaw University of Technology, 2023
    3. dr Grzegorz Gołaszewski, Systems Research Institute, Polish Academy of Sciences, 2022
    4. dr hab. Barbara Pękala, Systems Research Institute, Polish Academy of Sciences, 2019

```
## Warning in scan(file = file, what = what, sep = sep, quote = quote, dec = dec,
: EOF within quoted string
```

- Peer-reviewer for the following international journals (301 reviews written):

    1. *Accountability in Research: Ethics, Integrity and Policy* (1)
    2. *ACM Transactions on Knowledge Discovery from Data* (1)
    3. *ACM Transactions on Mathematical Software* (5)
    4. *Advances in Statistical Analysis (German Statistical Society)* (3)
    5. *Afrika Mathematica* (1)
    6. *Annals of Operations Research* (2)
    7. *Computational and Applied Mathematics* (1)
    8. *Computers and Operations Research* (2)
    9. *Control and Cybernetics* (1)
    10. *Data Mining and Knowledge Discovery* (4)
    11. *Demonstratio Mathematica* (1)
    12. *Engineering Applications of Artificial Intelligence* (1)
    13. *European Journal of Operational Research* (18)
    14. *Expert Systems with Applications* (1)
    15. *Foundations of Computing and Decision Sciences* (1)
    16. *Fundamenta Informaticae* (1)
    17. *Fuzzy Optimization and Decision Making* (3)
    18. *Fuzzy Sets and Systems* (39)
    19. *Group Decision and Negotiation* (1)
    20. *IEEE Access* (1)
    21. *IEEE Transactions on Big Data* (1)
    22. *IEEE Transactions on Emerging Topics in Computational Intelligence* (2)
    23. *IEEE Transactions on Fuzzy Systems* (59)
    24. *Information Fusion* (9)
    25. *Information Sciences* (45)
    26. *Intelligent Systems with Applications* (2)
    27. *International Journal of Applied Mathematics and Computer Science* (5)
    28. *International Journal of Approximate Reasoning* (4)
    29. *International Journal of Computational Intelligence Systems* (4)
    30. *International Journal of Forecasting* (1)
    31. *International Journal of Sports Science and Coaching* (5)
    32. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* (5)
    33. *International Transactions in Operational Research* (1)
    34. *Journal of Applied Analysis* (1)
    35. *Journal of Engineering Education* (1)
    36. *Journal of Informetrics* (12)
    37. *Journal of Intelligent and Fuzzy Systems* (3)
    38. *Journal of Open Source Software* (3)
    39. *Journal of the Association for Information Science and Technology* (7)
    40. *Knowledge and Information Systems* (1)
    41. *Knowledge-Based Systems* (3)
    42. *Mathematical Problems in Engineering* (1)

43. *Pervasive and Mobile Computing* (1)
44. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* (1)
45. *Quantitative Science Studies* (1)
46. *R Journal* (3)
47. *RUDN Journal of Mathematics, Information Sciences and Physics* (1)
48. *Scientific Reports* (1)
49. *Scientometrics* (23)
50. *Social Sciences and Humanities Open* (1)
51. *Soft Computing* (3)
52. *Statistical Modelling* (2)
53. *Wiley Interdisciplinary Reviews (WIREs) Data Mining and Knowledge Discovery* (1)

and international conferences (72 reviews written; IFSA/EUSFLAT 2009, IPMU 2010, IPMU 2012, SMPS 2014, EUSFLAT 2015, IPMU 2016, ISAS 2016, SMPS 2016, EUSFLAT 2017, IFSA/SCIS 2017, EUSFLAT 2019, FUZZ-IEEE 2020, IPMU 2020, AGOP 2021, FUZZ-IEEE 2021, IFSA/EUSFLAT 2021, FUZZ-IEEE 2022, IEEE-SAIC 2022, EUSFLAT 2023, FUZZ-IEEE 2023, Complex Systems 2024, IPMU 2024)

8 Teaching-Related Activities

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

- (Founding) Vice-chair of the Teaching and Learning Committee for the BSc and MSc degrees in Data Science (2016–2019, 2024–); key responsibilities: moulding, developing, and implementing a new degree in Data Science
- First-Year Academic Liaison for BSc in Data Science (2018–2019, 2024–); key responsibilities:
  - counselling students with regards to their academic goals and how to meet them
- Supervisor of the Data Science studies (a.k.a. Dean's Proxy for Data Science Studies) (2018–2019); key responsibilities:
  - managing, assessing, and synchronising all data science subjects
  - handling undergraduate and graduate students' admissions
  - coordinating students' transfers, providing advice regarding degree requirements and exchange programs (such as within the Erasmus framework)
  - lecturer/teacher allocation
- Member of the Teaching and Learning Committee for the BSc degree in Mathematics and Data Analysis (2019–)
- Initiator, supervisor, and mentor of the *Data Science* Student Club (2014–2019)
- Courses (units, subjects) taught:
  - Structured Data Processing
    New Unit Developer, Chair, and Lecturer (3 editions), 2018–2019, 2024–
  - Introduction to Programming and Data Processing
    New Unit Developer, Chair, and Lecturer (2 editions), 2017–2019
  - Data Processing in R and Python
    New Unit Developer, Chair, and Lecturer (3 editions), 2016–2019
  - Data Processing and Analysis in Python
    New Unit Developer, Chair, and Lecturer (4 editions), 2015–2019
  - Programming and Data Analysis in R
    New Unit Developer, Chair, and Lecturer (7 editions), 2012–2019

  - Algorithms and Introduction to Programming
    New Unit Developer, Chair, and Lecturer (6 editions), 2010–2016
  - Advanced R Programming
    New Unit Developer, Chair, and Lecturer (1 edition), 2013–2014
  - Mathematical Statistics I
    Tutor (4 editions), 2009–2012
  - Computer statistics
    Tutor (5 editions), 2008–2013
  - Programming in x86 Assembler
    Tutor (2 editions), 2010–2011
  - Algorithms and Data Structures II
    Tutor (4 editions), 2008–2011
  - Object-oriented Programming in C++
    Tutor (5 editions), 2007–2011

- Principal supervisor of 16 BSc students in Mathematics and Computer Science
- Principal supervisor of 11 MSc (by research) students in Mathematics, Computer, and Data Science

8.2 Deakin University, School of IT (Australia)

- Deputy Course Director for BSc in Data Science, 2022–2024
- Course Leadership Team Member for BSc in Artificial Intelligence, 2020–2023
- Course Leadership Team Member for MSc of Data Science, 2023–2024
- Academic mentor of 10 industry capstone projects at Deakin University – providing mentoring and assessment of students, 2020–2021
- Courses (units, subjects) taught:

  - SIT220/731 – Data Wrangling
    New Unit Developer, Chair, and Lecturer in 2022.T1, 2022.T3, 2023.T1, and 2023.T3
  - SIG731 – Data Wrangling (Master of Data Science (Global, online); via "Great Learning" in India)
    Unit Chair and Guest Lecturer in 2022.T3 and 2023.T3
  - SIT114 – Introduction to Artificial Intelligence
    New Unit Developer, Chair, and Lecturer in 2020.T1, 2021.T1, and 2022.T1
  - SIT752 – Introduction to IT Professional Practice
    Unit Chair in 2019.T3 and 2020.T1
  - SIT172 – Programming for Engineers
    Unit Campus Coordinator and Co-lecturer in 2020.T2 and 2021.T2

8.3 Other

Courses/units taught:

- Data Science Retreat, Berlin

  - NumPy, Pandas, TensorFlow, Advanced Python, Data Structures and Algorithms for Data Science, Introduction to R, Advanced R, Rcpp, Speeding up R and Python, String Processing, Good Development Practices in R
    Guest Lecturer (19 student batches), 2014–2019

- University of Silesia in Katowice

  - Introduction to Data Science in Python (online)
    Guest Lecturer, 2022

- Centre for Advanced Studies, Warsaw University of Technology

  - Python for Data Processing and Analysis
    Unit Developer and Chair in 2018

- Institute of Computer Science, Polish Academy of Sciences

  - Advanced Data Analysis Software Development in R (online)
    Unit Developer and Chair; three editions, 2014–2015

- Warsaw School of Information Technology

  - Statistical Decision Support Methods
    Tutor; three semesters, 2009–2011
  - Probability and Statistics
    Tutor; three semesters, 2008–2010

9 Open-Source Software Development and Industry Engagement

Author and maintainer of free (libre) and open-source source software (see my GitHub profile at https://github.com/gagolews/):

1. *genieclust* (https://genieclust.gagolewski.com/) – Python and R implementation of my fast and robust *Genie* hierarchical clustering algorithm
2. *clustering-benchmarks* (https://clustering-benchmarks.gagolewski.com/) – A framework for benchmarking clustering algorithms (including a package for Python)
3. *stringi* (https://stringi.gagolewski.com/) – text/natural language processing; one of the most often downloaded R packages (over 72,000,000 downloads)
4. *stringx* (https://stringx.gagolewski.com/) – drop-in replacements for base R string functions powered by *stringi*
5. *realtest* (https://realtest.gagolewski.com/) – a framework for unit testing for realistic minimalists, where we distinguish between expected, acceptable, current, fallback, ideal, or regressive behaviour; it can also be used for monitoring other software projects for changes
6. *genie* (http://cran.r-project.org/package=genie) – the reference R implementation of the *Genie* algorithm, now superseded by *genieclust*
7. *SimilaR* (http://cran.r-project.org/package=SimilaR) – code clones and plagiarism detection within R code chunks
8. *FuzzyNumbers* (http://cran.r-project.org/package=FuzzyNumbers) – R package implementing interval and fuzzy numbers arithmetic, and various piecewise linear approximation algorithms
9. *agop* (http://cran.r-project.org/package=agop) – aggregation operators in R
10. *CITAN* (http://cran.r-project.org/package=CITAN) – citation analysis toolpack for R
11. *TurtleGraphics* (http://cran.r-project.org/package=TurtleGraphics) – learn R programming while having a jolly time!

Other:

- Amongst top 3% StackOverflow users (https://stackoverflow.com/users/3309529/gagolews)
- G**gle *Summer of Code 2016* – Mentor of the *RE2 Regular Expressions in R* project (Student: Qin Wenfeng), 2016

- StackOverflow *Academic Research Partnership Program* – Supervisor of a research task related to quantitative determinants of the popularity of online content, 2019

Marek Gagolewski
13 December 2024