

David Novak

✉ davidnovak9000@gmail.com

in david-novak-04b65989

🐙 davnovak.github.io

Bioinformatician with proficiency in statistical data analysis, machine learning and software engineering

Employment History

- 2020–(Q1 2025) **Bioinformatics researcher.** FWO Strategic Basic research fellow at Saeys Lab, Center for Inflammation Research, VIB-UGent. Ghent, Belgium.
Research into applied ML and deep learning in cytometry and single-cell transcriptomics analysis pipelines with focus on interactive exploratory analyses, multivariate statistics, large differential expression analyses. Algorithm development for interpretable dimensionality reduction.
- 2021–(Q1 2025) **Assistant lecturer: machine learning.** Department of Applied Mathematics, Computer Science and Statistics, Ghent University. Ghent, Belgium.
Guiding practicals and student projects. Consulting role for experimental design and HPC use.
- 2018–2020 **Programmer and bioinformatics researcher.** Childhood Leukaemia Investigation Prague (CLIP), 2nd Faculty of Medicine, Charles University & Motol University Hospital. Prague, Czech Republic.
Development of automated pipelines for processing of flow and mass cytometry clinical and research data. R and C++ development of a novel semi-supervised trajectory inference solution.
- 2012–2020 **Translator and assistant.** MIDA Consulting. Prague, Czech Republic.
Czech-English translations and client-oriented research into European Union subsidy programmes and maintenance of supported projects.
- 2018 **Lecturer: ESL.** Channel Crossings. Prague, Czech Republic.
Designing and teaching an individual CAE prep course and a conversational English course.
- 2017–2018 **Lecturer: computer programming for children.** Logiscool. Prague, Czech Republic.
Teaching elementary-school students principles of algorithmisation.
- 2015 **Research intern: cellular neurophysiology.** Institute of Physiology, Czech Academy of Sciences. Prague, Czech Republic.
Primer design, bacterial transformation, plasmid preparation, IHC staining and fluorescent microscopy.




Education

- 2020 – (Q1 2025) **Ph.D. Bioinformatics.** Department of Applied Mathematics, Computer Science and Statistics, Faculty of Sciences, Ghent University, Ghent, Belgium. Supervisor: Prof Yvan Saeys.
Topic: Interpretable machine learning and statistics for exploration and modelling in single-cell biology.
- 2018 – 2020 **M.Sc. Bioinformatics** at Charles University, Prague, Czech Republic.
Thesis title: Studying lymphocyte development using mass cytometry.
- 2015 – 2018 **B.Sc. Biological Sciences** at Charles University, Prague, Czech Republic.
Thesis title: Human lymphopoiesis and its examination via single-cell analysis.


Research Publications

Journal Articles



- 1 M. Bakardjieva, O. Pelák, M. Wentink, *et al.*, “Tvbilindi algorithm identifies branching developmental trajectories of human b-cell development and describes abnormalities in rag-1 and was patients,” *European Journal of Immunology*, vol. n/a, no. n/a, p. 2 451 004, DOI: <https://doi.org/10.1002/eji.202451004>, eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/eji.202451004>.

- 2 T. Liechti, S. Van Gassen, M. Beddall, *et al.*, “A robust pipeline for high-content, high-throughput immunophenotyping reveals age- and genetics-dependent changes in blood leukocytes,” *Cell Reports Methods*, vol. 3, no. 10, p. 100 619, 2023, ISSN: 2667-2375.  DOI: <https://doi.org/10.1016/j.crmeth.2023.100619>.
- 3 A. Couckuyt, R. Seurinck, A. Emmaneel, *et al.*, “Challenges in translational machine learning,” *Human Genetics*, vol. 141, pp. 1–16, Sep. 2022.  DOI: 10.1007/s00439-022-02439-8.
- 4 T. A. Liechti, M. Beddall, S. Van Gassen, *et al.*, “Leveraging high-dimensional flow cytometry to reveal the human immune system at a population-wide scale,” *The Journal of Immunology*, vol. 206, no. 1 *supplement*, pp. 26.11–26.11, May 2021, ISSN: 0022-1767.  DOI: 10.4049/jimmunol.206.Supp.26.11.




Conference Papers

- 1 D. Novak, S. Van Gassen, and Y. Saeys, “GroupEnc: Encoder with group loss for global structure preservation,” in *BNAIC/BeNeLearn 2023*, Delft, Netherlands, 2023.  URL: https://bnaic2023.tudelft.nl/static/media/BNAICBENELEARN_2023_paper_46.9317ce00beb72bf31803.pdf.

Pre-prints

- 1 D. Novak, C. de Bodt, P. Lambert, J. A. Lee, S. Van Gassen, and Y. Saeys, “Interpretable models for scRNA-seq data embedding with multi-scale structure preservation,” 2024.  DOI: 10.1101/2023.11.23.568428. eprint: <https://www.biorxiv.org/content/early/2024/10/03/2023.11.23.568428.full.pdf>.
- 2 J. Stuchly, D. Novak, N. Brdickova, *et al.*, “Deconstructing Complexity: A Computational Topology Approach to Trajectory Inference in the Human Thymus with tviblinDi,” May 2024.  DOI: <https://doi.org/10.7554/eLife.95861.1>.

Skills and competencies

Languages	 Strong reading, writing and speaking competencies for English, Czech, Slovak. Conversational German. Basic Dutch.
IT	 Python data analysis, ML/AI and package development: <i>PyTorch</i> , <i>TensorFlow</i> , <i>numba</i> , <i>Optuna</i> , <i>pytest</i> . R statistical modelling, visualisation and frontend development: <i>ggplot2</i> , <i>Shiny</i> . Algorithm development in C++: <i>OpenMP</i> , <i>Boost</i> . Other languages & frameworks: <i>.NET</i> , <i>xUnit</i> , <i>Java</i> , <i>JavaScript</i> , <i>React</i> , <i>HTML</i> , <i>CSS</i> , <i>SQL</i> , <i>Bash</i> , <i>Slurm</i> , <i>Docker</i> , <i>AWS</i> , <i>Git</i> , <i>GitLab CI/CD</i> .
Misc.	 Statistical data analysis, machine learning, NGS data analysis, design and maintenance of large computational cytometry and single-cell ‘omics workflows, molecular biology theory, scientific writing, teaching.

In May 2024 I co-organised the *Computational Cytometry Summer School* at VIB-UGent, leading a session on statistical analyses and interpretation of experimental cytometry data.

In 2019 I co-taught R data analysis and statistics at the 2nd Faculty of Medicine, Charles University.

Throughout 2014–2015 I volunteered at the Thomayer hospital in Prague as part of the Lékořice foundation, holding weekly visits with elderly hospitalised patients.

References

References from employers, colleagues and collaborators will be made available upon request.