# David Novak, PhD

davidnovak9000@gmail.com david-novak-04b65989 PORTFOLIO: davnovak.github.io

Bioinformatician proficient in statistical data analysis, machine learning, and software engineering

## Employment history

2025– **Bioinformatics consultant.** *Subcontractor with Burns Life Sciences Consulting GmbH.*
Providing consulting services, both advisory and hands-on, to a diverse array of clients working with biological data.

2020–2025 **Bioinformatics researcher: exploratory modelling of single-cell data.**
*Saeys Lab, VIB-UGent, Belgium.*
Working with NIH immunologists, I designed *iidx*: an end-to-end interactive pipeline for large-scale differential analysis of complex age- and sex-associated immunophenotype changes, enabling the largest study of its kind on a 2196-donor flow cytometry dataset.
Having started a collaboration with UCLouvain, I developed *ViVAE* and *ViScore*: a novel VAE-based scRNA-seq dimension-reduction model with QC measures grounded in differential geometry, and a framework for robust evaluation of embeddings.

2021–2025 **Assistant lecturer: machine learning.** *Ghent University, Belgium.*
I have tutored over 200 students over 4 years, and have designed and taught graduate-level practical sessions. Emphasis on modern supervised ML use for predictive models in biomedicine.

2018–2020 **Programmer and researcher: computational cytometry.**
*Childhood Leukaemia Investigation Prague (CLIP), Czech Republic.*
I co-developed *tviblindi*: a semi-supervised trajectory inference solution using persistent homology, accessible to wet lab scientists within the group. This allowed us to build and publish multi-organ models of human B-cell and T-cell development.

2012–2020 **Translator and assistant: EU subsidy programmes.** *MIDA Consulting, Czech Republic.*

2018 **Lecturer: ESL.** *Channel Crossings, Czech Republic.*

2017–2018 **Lecturer: computer programming for children.** *Logiscool, Czech Republic.*

2015 **Research intern: cellular neurophysiology.** *Czech Academy of Sciences, Czech Republic.*

## Education

2020 – 2025 **PhD in Bioinformatics.** *Ghent University, Belgium. Supervisor: Prof Yvan Saeys.*

2018 – 2020 **MSc in Bioinformatics.** *Charles University, Czechia. Supervisor: Dr Jan Stuchly.*

2015 – 2018 **BSc in Biological Sciences.** *Charles University, Czechia. Supervisor: Prof Tomas Kalina.*

## Skills

IT Advanced data analysis, statistics, ML in *R* and *Python*. Application of modern deep learning frameworks. HPC use for large analyses. Frontends in *R Shiny*, *React*.
Familiar with *.NET*, *xUnit*, *C++*, *Java*, *HTML/CSS*, *SQL*, *Bash*, *Slurm*, *Docker*, *AWS*, *Git*, *GitLab CI/CD*, *OpenMP*, *numba*, *Optuna*, *Nextflow*.

Comp-bio Bulk & single-cell NGS and high-dim cytometry data analysis. Large and maintainable single-cell omics workflows. Interpretable machine learning. Consulting and tutoring. Scientific writing. Advanced data viz.

Languages Fluent English, Czech, Slovak. Conversational German. Basic Dutch.