

Szymon Zawadzki

Młodszy Analityk Danych Naukowych

Kraków, Polska | szymon.zawadzki.bio@gmail.com | +48 513 284 917 | [linkedin.com/in/szymon-zawadzki-bioinfo](https://www.linkedin.com/in/szymon-zawadzki-bioinfo) | github.com/szawadzki-bio



Absolwent bioinformatyki na Uniwersytecie Jagiellońskim z doświadczeniem w analizie danych genomicznych w instytucie badawczym. Analizowałem zbiór 42.000 próbek z wykorzystaniem R (Bioconductor) i Pythona (pandas, scikit-learn). Współautor publikacji naukowej w recenzowanym czasopiśmie. Biegłe posługuję się metodami statystycznymi i narzędziami bioinformatycznymi.

■ DOŚWIADCZENIE

Stażysta Analityk Danych, Instytut Farmakologii im. Jerzego Maja PAN, Kraków
Staż w zespole bioinformatycznym instytutu badawczego.

Lip 2025 – Sty 2026

- Analiza danych transkryptomicznych (RNA-seq) obejmujących **42.000 próbek** z publicznych baz danych
- Opracowanie pipeline bioinformatycznego w **R (Bioconductor)** do analizy ekspresji genów
- Współautorstwo **publikacji naukowej** w recenzowanym czasopiśmie (IF: 3,8)
- Wizualizacja wyników w **ggplot2** i przygotowanie 18 rycin do publikacji i prezentacji konferencyjnych

Praktykant Bioinformatyk, Centrum Badań Genomowych (Uniwersytet Jagielloński), Kraków
Praktyki w laboratorium genomiki obliczeniowej.

Sty 2025 – Cze 2025

- Przetwarzanie danych sekwencjonowania **NGS** (Illumina) dla 3 projektów badawczych
- Mapowanie odczytów i analiza wariantów z wykorzystaniem **BWA, GATK i SAMtools**

■ WYKSZTAŁCENIE

Stopień i kierunek, Uniwersytet Jagielloński, Kraków

Paź 2021 – Cze 2026

Praca magisterska: *Analiza danych transkryptomicznych w identyfikacji biomarkerów chorób neurodegeneracyjnych z wykorzystaniem metod uczenia maszynowego*. Średnia ocen: 4,61.

■ UMIEJĘTNOŚCI

R (Bioconductor, tidyverse, ggplot2) • Python (pandas, scikit-learn, NumPy) • Analiza danych RNA-seq / NGS • Statystyka (testy parametryczne, regresja, PCA) • Bazy danych biologicznych (NCBI, Ensembl, UniProt) • Linux / Bash (skrypty) • Git / GitHub • BWA / GATK / SAMtools • LaTeX (publikacje naukowe)

■ JĘZYKI

Polski (ojczysty) • Angielski (C1)

■ PROJEKTY

Pipeline do analizy ekspresji genów

Wrz 2025 – Sty 2026

Autorski pipeline bioinformatyczny w R do analizy danych RNA-seq.

- Przetwarzanie danych od surowych odczytów do **analizy różnicowej ekspresji** (DESeq2)
- Udostępniony na GitHubie z **dokumentacją i testami** — 85 gwiazdek