

Jan Malý

Interdisciplinary Master's student in Computational Linguistics combining a deep understanding of natural language processing with hands-on experience in LLM backend engineering and dataset development.

Brno | malyjan3581@gmail.com | +420 608 832 327

GitHub Page <https://janmalyy.github.io/> | LinkedIn <https://www.linkedin.com/in/jan-malyy/>

Education

2024-2026 | Computational Linguistics, Faculty of Informatics & Faculty of Arts, Masaryk University (Mgr.)

- Thesis: A web application for interpreting genetic variants mentioned in the literature

2021-2024 | Computational Linguistics, Faculty of Informatics & Faculty of Arts, Masaryk University (Bc.)

- Thesis: A web application for generating Czech dictation exercises

Working experience

05/2025-05/2026 | Data annotator, Masaryk University

- Developed and used a custom tool to generate and annotate Czech language datasets, accelerating the training pipeline for the Opravidlo 2.0 grammar checker

01-03/2025 | NLP Engineer Intern, AI Check, s. r. o.

- Engineered backend for an LLM-driven chatbot, integrating Retrieval Augmented Generation (RAG) to query a database (MongoDB)
 - Organised workflow and communication in our small team
-

Languages

English: B2 (FCE, 2021)

French: B2 (DELF, 2021)

Technical Skills

Python, JavaScript

PostgreSQL, MongoDB

Git, Docker

Hobbies

organising LARPs

volunteering with Hnutí Brontosaurus

hiking

literature