



# Mykhailo Pavlov

📍 Bratislava, Slovakia

## Skills

Python



Java



C++



JavaScript



SQL



Git



GitHub Actions



Kubernetes



Docker



Linux



LLM integration



Prompt-based pipelines



## Languages

Slovak



English



Ukrainian



Russian



Computer Science student at Comenius University (graduating 2026) with hands-on experience in AI/ML, data pipelines, and LLM-powered applications. Strong interest in MLOps/DevOps, foundation models, and scalable AI solutions. Skilled in Python, TensorFlow, PyTorch, and data analysis libraries, with experience in Docker, CI/CD, and collaborative development via GitHub.

## Profiles

GitHub  
[Pavloffm](https://github.com/Pavloffm)

LinkedIn  
[mykhailo-pavlov](https://www.linkedin.com/in/mykhailo-pavlov)

## Education

Comenius University Bachelor's degree  
Computer Science September 2023 - June 2026

<https://fmph.uniba.sk/>

## Projects

GitHub h-index (Bachelor Thesis)

- Designed and implemented a scalable data collection pipeline to compute a GitHub-adapted h-index for public users based on repository stars
- Collected and processed data for ~3.5 million GitHub users using the GitHub GraphQL API, carefully managing API rate limits
- Built parallelized data ingestion using multiple authentication tokens, improving throughput by 4–5× and reducing runtime from months to days
- Designed a lightweight relational data storage architecture (SQLite) with indexed schemas for efficient large-scale analytics
- Engineered user-level features and trained a Random Forest regression model to predict GitHub h-index, achieving  $R^2 = 0.95$
- Implemented robust data cleaning, feature engineering, and reproducible train/test pipelines
- Thesis focused on scalable data pipelines, API-driven data ingestion, and production-oriented data processing

Asynchronous Telegram Data Parser with Dockerized PostgreSQL Storage

- Designed and deployed a production-ready async data ingestion pipeline processing 5,000+ messages/min
- Implemented robust error handling, rate-limit recovery, structured logging, and configuration management
- Containerized services using Docker and managed database connectivity in isolated environments
- Processed and stored 50,000+ records with full data integrity

Multilingual Anki Vocabulary Parser

- Built a Python tool using BeautifulSoup and LLMs to generate multilingual Anki flashcards from online dictionaries
- Developed a Flask web app supporting 100+ target languages and automated creation of 1,000+ cards
- Added logging, tests, and CI with GitHub Actions for reliability and maintainability

## Certifications

IBM Consulting Slovakia — Fellowship of Spring 04.11.2025  
IBM

Covered Spring Architecture, REST API Development, Persistence (Data Management with Spring), and Testing in Spring-Based Applications