

SZYMON GÓRKA

Software/ML engineer

PROFILE

I'm a passionate generalist who thrives on exploring all aspects of technology, from software development to hardware tinkering and machine learning. No challenge is too big or too small – I love diving into different tech fields and constantly expanding my skill set. Outside the tech world, I'm an avid fencer, always sharpening my skills in this dynamic and strategic sport.

EDUCATION

Łódź University of Technology • 2023

Engineer Degree in Computer Science at the EEIA faculty

EXPERIENCE

Digica, 11.2022 - present

Generalist - worked on broad range of projects - mostly related to AI on embedded, classic ML, mobile apps, python backend developer

Septic Solutions, 06.2024 - present

Backend/infrastructure python developer and shareholder. Large scale application for managing septic tanks at country scale. My work include both building the side and cloud infrastructure.

LANGUAGES

Polish native
English B2

PROJECTS

LLM EU ACT search engine

Project developed at Digica. Partnered with a university to develop embeddings and a local chat model based on the EU AI Act, aimed at advising the legal status of LLM-related projects. My contributions included:

- Researching and selecting the best local models
- Testing and implementing various text-splitting and embedding methods

Markov Chain Model in predictive maintenance

Collaborated with a leading smart home solutions company in North America to develop predictive models for home occupancy based on alarm signal data. My key contributions included:

- Cleaning and processing large datasets (~0.5TB per batch)
- Building and training machine learning models
- Communicating insights and data findings to the client
- Utilizing the Azure Databricks environment for data processing and model development

Thesis: "Effective Algorithms for Genome Similarity Approximation"

In my thesis, I optimized several well-known algorithms for genome comparison, addressing a key problem in bioinformatics. The result was one of the fastest algorithms globally for genome similarity estimation. My work involved:

- Writing highly optimized C++ code
- Analyzing disassembled code to identify performance bottlenecks
- Handling large datasets (5-10k files, hundreds of gigabytes)