Petr Pouč

Senior GenAI Developer

Address: Mladá Boleslav, Czech Republic
Email: petrpouc0@gmail.com
LinkedIn: linkedin.com/in/petr-pouč-274850237
Github: https://github.com/surViVeCZ

Professional Profile

AI Researcher and Engineer specializing in machine learning and cybersecurity, with a strong track record in developing data-driven, AI-based solutions. Proficient in Python, data science, and machine learning models, with a background in cloud services and scalable applications. Proven ability to manage complex projects from concept to execution through collaborative efforts.

Education

Jakub Škoda Gymnasium, Přerov (2011 – 2019)
Eight-year comprehensive study

Brno University of Technology, Faculty of Information Technology
Bachelor's Degree in Information Technology (2019 – 2022)

Brno University of Technology, Faculty of Information Technology
Master's Degree in Cybersecurity (2022 – 2024)
Thesis: *Optimization of Classification Models for Malicious Domain Detection*
https://www.vut.cz/studenti/zav-prace/detail/154617

Certificates

CyberOps Associate
Cisco, issued May 2023

Microsoft Certified: Azure Fundamentals
Microsoft, issued February 2023, Credential ID: I630-9753

Work Experience

Brno University of Technology — **March 2023 — Present**

Research Intern *March 2023 — April 2024*
Contributed to applied AI research focused on network security. Collected and labeled domain datasets, built automatic evaluation tooling, and applied machine learning methods to detect malicious activity in network traffic.

AI Researcher — FETA Project *April 2024 — July 2025*
Worked on the Flow-based Encrypted Traffic Analysis (FETA) project for the Ministry of the Interior of the Czech Republic, detecting threats in encrypted communications through domain-centric analysis. Improved classifiers for malicious, phishing, and DGA domains by enhancing feature vectors, building a robust dataset verification pipeline, and iterating on model quality.

AI Researcher — THOR Project *January 2026 — Present*
Working on the THOR project, a follow-up to FETA developed in collaboration with Cisco, focused on advancing threat detection in encrypted network traffic. Building on FETA's domain-centric detection methods and datasets to develop next-generation classification models for real-world deployment.

Lakmoos AI — June 2023 — June 2026

Data Science Team Lead *June 2023 — August 2024, full-time*
Led a team of data scientists developing predictive analytics models and data workflows. Owned the transformation of client data into formats suitable for the AI backend, including the design of data generators and quality standards for datasets feeding AI models.

AI Software Engineer / Data Engineer *July 2024 — March 2025, part-time*
Developed and optimized synthetic data generators, integrating real-world data to create enriched population clones. Applied classification, regression, and data augmentation strategies to improve model robustness. Designed ETL tools and automated pipelines for large-scale datasets, and built dashboards for monitoring data quality and model performance.

Data Supervision & Advisory *March 2025 — June 2026, part-time*
Advised on data quality, synthetic data generation methodology, and statistical validation of AI-generated populations. Supported research initiatives, including benchmarking of synthetic persona models against official statistical sources across multiple European markets.

Green:Code (Škoda Auto) — March 2025 — June 2026

AI Engineer
Designed and deployed internal AI applications that streamlined workflows and improved information accessibility across the organization. Built production RAG systems for retrieving and summarizing information from large collections of technical documentation, and a recommender system for employee development. Developed agentic AI features, tool-calling workflows and LLM observability practices (Langfuse) within an internal AI platform, working with vector databases at scale and containerized deployment on AWS.

Viable One — July 2026 — Present

Senior GenAI Developer *Project: Našeptávač / Hyperpersonalization (banking)*
Delivered end-to-end, in record time, an AI assistant for branch and contact-centre bankers, covering solution architecture, backend integration and prompt tuning. Currently building further components of the Orchestrator platform, including a production feedback loop for continuous improvement of model outputs.

Research Projects

Project FETA (Flow-based Encrypted Traffic Analysis)
A collaboration between CESNET, CTU Prague, and BUT Brno funded by the Ministry of the Interior of the Czech Republic under the IMPAKT 1 program. The project aims to develop new technologies for monitoring encrypted communication and identifying security threats in network infrastructure.

My main contribution was the implementation of an automated verification pipeline for creating ground-truth datasets, as well as participation in refining classification models and data preprocessing. The project was also related to my master's thesis, which is available on GitHub.

Publications and Research Outputs

A multi-dimensional DNS domain intelligence dataset for cybersecurity research
Co-author of a peer-reviewed journal article presenting a unique, multi-dimensional dataset for network security research, focusing on the classification and analysis of internet domains (phishing, malware, and benign traffic). The dataset includes comprehensive metainformation from DNS records, TLS handshakes, WHOIS/RDAP services, IP data, and geolocation details for **over a million internet domains**. Published open-access in *Data in Brief*.
DOI: 10.1016/j.dib.2025.112062

Phishing and Benign Domain Dataset (DNS, IP, WHOIS/RDAP, TLS, GeoIP)

Together with the team, I contributed to the creation of a large dataset containing 432,572 verified benign domains and 36,993 verified phishing domains, which served as the basis for many other research topics. The dataset is publicly available under the Creative Commons Attribution 4.0 International license, and details can be found on the BUT FIT portal.

Which AI? A comparative evaluation of generative AI and neuro-symbolic AI for silicon sampling
Co-author of a four-study benchmark of generative and neuro-symbolic AI as synthetic survey respondents, validated against official European statistics and a human panel. Contributed software and data curation. Under review (Revision 2), *Psychology & Marketing*.
Replication package: 10.5281/zenodo.22753761

Technical Skills

Most of my work today is building AI applications that people actually use. That means RAG systems over large document collections, AI agents, and the tooling around them. I spend a lot of time on the parts that make these systems work in production, like getting retrieval quality right, keeping vector databases fast and cheap at scale, and setting up observability so you know what your LLM is really doing. All of this sits on a solid data science background. I started in machine learning and data engineering, and that still shapes how I build AI systems today.

AI Engineering: RAG systems, AI agents and tool calling, prompt design, LLM observability and evaluation (Langfuse), working with Anthropic Claude, OpenAI and AWS Bedrock

Vector Search & Retrieval: Qdrant, FAISS, Chroma, embedding models, semantic search, recommender systems, scaling and cost optimization for collections with tens of millions of records

Machine Learning: Classification, regression, data augmentation, model evaluation, generative models

Languages & Data: Python (pandas, NumPy, scikit-learn, PyTorch), SQL, ETL and data pipelines, dashboards (Plotly Dash)

Cloud & DevOps: AWS (ECS, CloudWatch, Parameter Store, Bedrock), Azure, Docker, CI/CD, Git, Linux

About Me and Future Goals

I would describe myself as a hardworking, reliable, and non-conflicting person. In addition, I am thorough and a little bit of a perfectionist, with a strong attention to detail. I have excellent problem-solving skills and am able to think creatively to find effective solutions. I am also highly adaptable and eager to learn new things, even when they are unfamiliar to me. Outside of work, I enjoy spending time with friends, and, like many in tech, my productivity often leans on a good cup of coffee to keep ideas flowing.

Looking ahead, one of my primary goals is to enhance my communication skills to become more articulate and confident in conveying complex ideas. I recognize that while my technical expertise is strong, there is room to improve in translating those ideas clearly to others. Additionally, I'm keen to deepen my technical knowledge and broaden my experience, working with others to bring innovative solutions to life.

November 27th, 2026, Mladá Boleslav



Ing. Petr Pouč