

Bertalan Radostyán

Data Scientist / Engineer

rad.bertalan@gmail.com
[linkedin.com/in/bertalan-radostyan](https://www.linkedin.com/in/bertalan-radostyan)



Currently working as a Data Engineer, with more than 7 years of relevant experience, supported by a strong academic background. Experienced across data science/engineering, machine/deep learning, software development, and AI engineering.

Experience

Data Engineer / Szallas Group

May 2025 - PRESENT, Budapest

- Architected and delivered an end-to-end data platform from concept to production.
- Designed scalable batch and real-time data processing architecture, including CDC-based ingestion pipelines.
- Led cloud infrastructure architecture and operations, covering provisioning, reliability, monitoring, and deployment environments.
- Introduced and operationalized the use of AI (LLMs) across engineering workflows, design processes, and internal self-service tools to improve productivity and decision-making.
- Tech stack:
 - Programming languages: Python, SQL
 - Data/ETL tools: dbt, Airbyte
 - DevOps: Docker, Kubernetes, GitLab CI
 - Cloud: GCP, BigQuery, GKE, VertexAI

Data Scientist / Sanofi

November 2023 - May 2025, Budapest

- Delivered end-to-end AI solutions for commercial use-cases in the pharmaceutical domain.
- Developed and maintained a decision-support system for sales representatives to improve engagement with healthcare professionals.
- Designed an AI assistant for pharmaceutical documents using LLMs, RAG, and knowledge graphs.
- Tech stack:
 - Programming languages: Python, SQL
 - ML tools/packages: Scikit-Learn, Tensorflow, Pandas
 - MLOps: MLflow
 - AI/LLMs: LangChain, Neo4j

Data Scientist / AutSoft Zrt.

February 2019 - October 2023, Budapest

Data Platforms & Integration:

- Designed analytics schemas and optimized complex queries in transactional environments.

- Built ETL pipelines and a centralized data warehouse and analytics platform.
- Automated banking system integrations and data processing workflows.
- Designed data lake architecture with real-time ingestion from multiple source systems.
- Provisioned and operated Apache Kafka clusters on Kubernetes (replication, monitoring, performance tuning).

Machine Learning & Decision Support

- Developed forecasting models for product and energy consumption.
- Built decision-support systems including customer segmentation and trading signal prediction.
- Delivered Industry 4.0 ML solutions for production environments.
- Created self-service BI reporting and dashboards for business users.

Advanced Analytics & Computer Vision

- Implemented OCR pipelines for ID and document recognition.
- Developed algorithms for processing medical 3D scans.
- Conducted predictive maintenance research and internal data science training.

Tech stack

- Programming languages: Python, SQL, C++
 - ML tools/packages: Scikit-Learn, PyTorch, Tensorflow
 - Other tools: Apache Kafka, Apache Spark
 - DevOps: Kubernetes, Grafana, Prometheus
 - MLOps: MLflow
 - Cloud: AWS
-

Data Science Researcher/ AutSoft Kft.

June 2018 - December 2018, Budapest

- Predictive maintenance in production line environments.
 - Tech stack: Apache Hadoop, Spark, Kafka, PyTorch (RNN, LSTM)
-

Software Developer / AutSoft Kft.

March 2017 - August 2017, Budapest

- Full-stack and mobile development.
 - Tech stack: ASP.NET, Android, Node.js, Angular
-

Software Researcher and Developer / Budapest University of Technology and Economy

April 2014 - June 2016, Budapest

- Educational software research and development.
- Tech stack: ASP.NET, Android

Education

MSc in Computer Engineering / Budapest University of Technology and Economy

September 2017 - June 2019, Budapest

Diploma thesis: Fast Data-based Production Support System

BSc in Computer Engineering / Budapest University of Technology and Economy

September 2013 - January 2017, Budapest

Thesis: Adaptive Figural Abstraction Test with Generated Exercises

Core Competencies

Data Engineering & Platforms

- Data modeling, ETL/ELT design, CDC ingestion
- Data warehousing & data lake architecture
- Batch and real-time processing pipelines

Machine Learning & AI

- Forecasting, segmentation, and decision-support models
- LLM applications, RAG systems, knowledge graphs
- Computer vision and advanced analytics pipelines

Cloud & DevOps

- Kubernetes, Docker, CI/CD pipelines
- Observability, monitoring, and reliability engineering
- GCP and AWS data platforms

Programming

- Python, SQL, C#

Collaboration

- Stakeholder communication and requirement translation
- Estimation and solution design in cross-functional teams

Languages

- English – C1
- German – A2