

Patrik P. Süli

Software Architect · AI Engineer · PhD Candidate · Founder of KenguTech

EMAIL	patrik.suli@mailbox.org
UNIVERSITY	suli.patrik@uni-obuda.hu
WEB	PatrikPSuli.com
LINKEDIN	patrik-süli
ORCID	0009-0001-9481-3664
RESEARCHGATE	Patrik-Sueli
GOOGLE SCHOLAR	h4ASwVUAAAAJ
LOCATION	Martonvásár, Hungary

PROFILE

Software architect with a decade of professional JVM engineering, currently leading enterprise microservice systems sized for millions of requests at KenguTech, the studio I founded in 2022. PhD candidate at Obuda University and Assistant Lecturer at the John von Neumann Faculty of Informatics, with peer-reviewed research on applied large language models, agentic systems for scientific workflow, and LLM evaluation methodology.

EXPERIENCE

HUN-REN Centre for Agricultural Research (ATK)

Sep 2026 – Present

Software Architect · AI Engineer

Architecting research informatics for plant breeding and crop science at the HUN-REN Agricultural Research Centre in Martonvásár: modernising the scientific software and data infrastructure and embedding LLM and machine learning capabilities into the research workflow.

KenguTech

Dec 2022 – Present

Software Architect · Founder

Founded the studio and lead time-and-materials engagements end-to-end: architecture, hands-on delivery, and operational ownership. Specialised in enterprise-grade Java microservice systems sized for millions of requests, alongside applied LLM solutions.

Customer Portal Re-Platforming

2025 – Present

Software Architect · Lead Developer · Hungarian Electricity & Natural Gas Provider

Architect and lead developer on the modernisation of a public portal serving millions of utility customers. Decomposing a legacy monolith into a high-availability microservice topology aligned with modern DevOps and cloud-native practice, with reliability and operational maintainability as primary constraints.

STACK: Java · Spring Boot · Maven · REST · Microservices · PostgreSQL · Redis · RabbitMQ · Debezium · Keycloak

High-Throughput Billing Platform

2023 – 2025

Software Architect

Co-architect on a 25+ microservice distributed billing platform sized for millions of customers. Led the planning and specification of the high-throughput data and integration pathways underpinning the platform's correctness and performance guarantees.

STACK: Java · Spring Boot · Maven · REST · Microservices · PostgreSQL · Redis · Elasticsearch · Kafka · Debezium · Keycloak · GitLab CI · Kubernetes

AW Server DevSecOps Pipeline

2022 – 2023

DevOps Engineer · GE HealthCare

Owned the GitLab CI/CD pipeline for AW Server, GE HealthCare's enterprise medical-imaging visualization platform deployed across hospital networks. Built and maintained the surrounding Bash and Python tooling, plus the vulnerability-scanning workflow that gated each release.

STACK: Python · Bash · GitLab CI · Nessus · SCAP · RHEL · SLES

Obuda University · John von Neumann Faculty of Informatics

Aug 2024 – Present

Assistant Lecturer & Researcher

Teach and supervise at the Biomatics and Applied AI Institute alongside an active research agenda. Focus areas: applied large language models, agentic systems for scientific workflow, and LLM evaluation methodology.

Freelance

Jan 2018 – Nov 2022

Software Developer · Consultant

Nearly five years of independent and contract delivery across Java (Spring Boot, Quarkus), TypeScript (NestJS, Angular), with occasional C# work. Across the engagements I shipped microservice backends, full-stack products, and the CI/CD pipelines and environments around them, often serving as the de-facto DevOps engineer too.

Hungarian e-Health Integration Gateway

2022 – 2022

Software Architect · Lead Developer

Software architect and lead developer on the gateway that translated private-hospital messaging into EESZT, the Hungarian national e-Health infrastructure every healthcare provider must report to. Also led the team's day-to-day task management and delivery rhythm.

STACK: Java · Spring · Spring Boot · Maven · REST · MongoDB · Redis · Docker · GitLab CI · Proxmox

Enterprise Conversational AI Platform

2021 – 2022

Software Architect · Lead Developer · Talk-A-Bot Kft.

Software architect and lead developer on Talk-A-Bot's enterprise conversational-AI platform, which serves brands operating under regulated-communications constraints. Built the natural-language-processing pipelines around external NLP services and the API surface that exposed them.

STACK: Java · Spring · Spring Boot · Maven · REST · Vue.js · JPA · Hibernate · PostgreSQL · Flyway · Tomcat · Docker

Real-Time E-Learning Platform

2020 – 2020

Senior Backend Developer

Built the Spring Boot microservices behind the WebSocket-based real-time chat layer of a synchronous e-learning platform sized for classroom-scale concurrency. Stood up the Docker-driven CI/CD pipeline that delivered them to production.

STACK: Java · JHipster · Spring Boot · Maven · JPA · Hibernate · Liquibase · Oracle · PostgreSQL · Elasticsearch · WebSocket · Docker · Kubernetes

MÁV Rail Document Interoperability Service

2020 – 2020

Senior Backend Developer

Spring Boot microservices that integrated MÁV's document workflows with EU rail-interoperability services, with .NET Core components handling the cross-runtime interop boundary.

STACK: Java · Spring Boot · Maven · JPA · Hibernate · Liquibase · Oracle · REST · Swagger · C# · .NET Core

Unified Multi-Tenant HR Platform

2020 – 2020

Full-Stack Developer

Full-stack delivery of a multi-tenant HR product: NestJS backend, Angular frontend, plus the GitLab-driven dev and test environments and pipelines around them.

STACK: NestJS · Node.js · TypeScript · Angular · MongoDB · TypeORM · Mongoose · Docker · GitLab CI

Unified Transit API Gateway

2019 – 2019

Backend Developer

Quarkus-based aggregation service unifying multiple public-transport operators behind a single REST API, normalising the disparate provider responses into one consistent schema.

STACK: Java · Quarkus · Maven · JPA · Hibernate · Liquibase · Oracle · PostgreSQL · REST · Swagger

ELBA Architectural CAD Platform

2018 – 2019

Full-Stack Developer

Spring Boot microservices powering an architectural-CAD SaaS, with integrated OTP Simple payment for subscription billing. Also delivered the dev, test, and prod environments and the CI/CD pipelines that connected them.

STACK: Java · Spring Boot · Maven · Gradle · JPA · Hibernate · Liquibase · Oracle · PostgreSQL · REST · Angular · Docker

NETI IT Consulting Ltd.

Jul 2016 – Jun 2019

> *Software Developer*
>
> Mar 2017 – Jun 2019
>
> Delivered cross-platform C++ desktop applications on the Qt framework, then moved into full-stack development on the company's flagship product line with JavaFX and Spring Boot.

> *Trainee Developer*
>
> Jul 2016 – Mar 2017
>
> First professional role. Contributed production features to a graph-designer application in C++ with the Qt framework as part of the platform team.

EDUCATION

Ph.D. · Applied Informatics

2024 – 2028

Obuda University · Doctoral School of Applied Informatics & Applied Mathematics

M.Sc. · Computer Science

2022 – 2024

Eötvös Loránd University

B.Sc. · Computer Science Engineering 2013 – 2017

John von Neumann University

TECHNICAL SKILLS

LANGUAGES	Java · TypeScript · Python · C++ · C# · Bash
FRAMEWORKS	Spring Boot · Quarkus · Helidon · NestJS · FastAPI · JHipster · Angular · Vue.js
DATA & MESSAGING	PostgreSQL · MongoDB · Redis · Oracle · Elasticsearch · pgvector · Qdrant · Kafka · RabbitMQ · Debezium
INFRA & DEVOPS	Docker · Kubernetes · GitLab CI · GitHub Actions · ArgoCD · Keycloak · Nessus · SCAP · RHEL
AI / LLM	LangChain · LangGraph · OpenAI · Anthropic · Llama · Qwen · DeepSeek · Ollama
TESTING & QUALITY	JUnit · Testcontainers · Pact · k6 · Gatling · SonarQube · SpotBugs

OPEN SOURCE

benchgap — *Co-author* hollorol/benchgap

Fills the empty cells of LLM leaderboards. Calibrates benchmarks against each other on the models measured on both, predicts each missing score from that model's own results, and attaches a cross-validated error and confidence level to every estimate. Data refreshed daily from BenchLM, with a public OpenAPI 3.1 JSON API. Built with Roland Hollós.

Distribird — *Author* HUN-REN-AI1Science/Distribird

Multi-agent LangGraph pipeline that turns scientific literature into informative Bayesian priors. Searches Semantic Scholar and OpenAlex in parallel, extracts reported numerical values, and fits a probability distribution via AIC model selection. Built with HUN-REN AI1Science.

SpotBugs — *Contributor* spotbugs/spotbugs

6 merged pull requests to the widely-used Java static-analysis tool, including two new CERT Java detectors with full test cases: SER11-J (prevent overwriting of externalizable objects) and VNA04-J (ensure calls to chained methods are atomic). The atomicity detector connects to the published Acta Cybernetica paper on multithreading atomicity.

SELECTED PUBLICATIONS

Süli, P. P., Eigner, G., Zrubka, Z. (2026).

Emergent Health Preferences in Large Language Models: A Time-Trade-Off Analysis Using EQ-5D-3L.

IEEE 30th Jubilee International Conference on Intelligent Engineering Systems (INES 2026), Budapest, Hungary · Conference paper · DOI: 10.1109/INES69513.2026.11661097

Yang, J., Dotoli, M., Wang, Y., Wang, X., Liu, Y., Tian, Y., Ge, J., Ni, Q., Carli, R., **Süli, P. P.**, Horti, D., Allgöwer, F., Werbos, P. J., Shen, Z. (2026).

Automation 5.0: The Step to Systems Intelligence for a Sustainable Future.

IEEE Transactions on Automation Science and Engineering · Journal article · DOI: 10.1109/TASE.2026.3668032

Zrubka, Z., Marczisák, B., **Süli, P. P.**, Eigner, G., Gáspár, J., Benedek, M., Sziklai, B. (2025).

Piloting ModiBot: A Large Language Model-Based Moderator in Normal and Emotionally Challenging Focus Group Interactions.

IEEE 25th International Symposium on Computational Intelligence and Informatics (CINTI 2025), Budapest, Hungary · Conference paper · DOI: 10.1109/CINTI67731.2025.11311748

Süli, P. P., Tafferner-Gulyás, V., Kovács, L., Eigner, G. (2025).

A Survey of Large Language Models as Research Assistants.

IEEE 25th International Symposium on Computational Intelligence and Informatics (CINTI 2025), Budapest, Hungary · Conference paper · DOI: 10.1109/CINTI67731.2025.11311841

Ding, W., Wang, F.-Y., Wang, X., Ni, Q., Ge, J., Fan, L., **Süli, P. P.**, Hartveg, Á., Szatmáry, R., Simon, B., Kovács, L. (2025).

DeSci for Control Education: From CDIO to iCDIOS.

IFAC PapersOnLine · Elsevier · Journal article · DOI: 10.1016/j.ifacol.2025.08.055

Süli, P. P., Marczisák, B., Csibrák, T., Eigner, G., Zrubka, Z. (2025).

Augmenting virtual brainstorming workshops via agentic AI.

International Conference and Summer Institute on GenAI Literacy (ICGAL 2025), Budapest, Hungary · Abstract

Zhang, T., Tian, Y., Lin, F., Huang, J., **Süli, P. P.**, Ni, Q., Qin, R., Wang, X., Wang, F.-Y. (2025).

CoordField: Coordination Field for Agentic UAV Task Allocation in Low-altitude Urban Scenarios.

arXiv e-print · Preprint · DOI: 10.48550/arXiv.2505.00091

Süli, P. P., Knoll, J., Porkoláb, Z. (2025).

Multithreading Atomicity: Static Analysis Checkers.

Acta Cybernetica · University of Szeged · Journal article · DOI: 10.14232/actacyb.311949

Nagy, M., Simon, B., Szász, L., Siket, M., Dénes-Fazakas, L., Eigner, G., **Süli, P. P.**, Kovács, L., Szilágyi, L. (2024).

Web application development for diabetes patients.

IEEE 18th International Symposium on Applied Computational Intelligence and Informatics (SACI 2024), Timişoara, Romania · Conference paper · DOI: 10.1109/SACI60582.2024.10619716

WORKING PAPERS & IN PROGRESS

Süli, P. P., Eigner, G., Hollós, R. (2026).

benchgap: Mind the Gap: Explicit Cross-Benchmark Calibrations for Estimating Missing LLM Evaluation Scores.

In preparation · status: in progress

Most LLMs are evaluated on only a few benchmarks: across 401 models and 146 benchmark versions, coverage is just 8.3%. benchgap estimates the missing scores without running any evaluations, calibrating each benchmark against same-capability benchmarks with monotone curves selected by leave-one-out cross-validation, and attaching an error and confidence level to every estimate. In a 15% holdout it fills 71% of masked cells at 4.6 pp MAE, 23% lower than a matrix-completion baseline. It also quantifies the harness tax, the disagreement between evaluation setups on the same models. Estimates are updated daily at benchgap.net.

Süli, P. P., Eigner, G., Hollós, R. (2026).

SchemaEvolve: Data Model Refinement for LLM-Based SQL Code Generation.

Submitted to ICLR 2027 · status: submitted

Keeps the Text-to-SQL model fixed and optimises the database schema it has to read instead. An evolutionary loop proposes restructured schemas as code, each candidate must preserve the original data, and fitness is the execution accuracy of a small local translator. On 13 real databases from Spider 2.0-Lite, 20 generations raise the mean accuracy of three small local models from 27.1% to 50.6% at about $3.50 per database, and the two stronger evolved translators overtake Claude Opus 4.6 reading the original schemas. A local 27B mutator recovers most of the gains at zero API cost.

Süli, P. P., Eigner, G., Hollós, R. (2026).

Distribird: Literature-Informed Prior Distribution Design for Bayesian Model Calibration.

Submitted to Scientific Reports (Nature Portfolio) · status: preprint · arXiv:2608.11210 · Research Square

A multi-agent LangGraph pipeline that turns scientific literature into informative Bayesian priors. Searches Semantic Scholar and OpenAlex in parallel, extracts numerical values, and fits a probability distribution via AIC model selection, improving MCMC sampling efficiency in 80% of cases across 20 parameters in ten scientific domains. Preprints are live on arXiv and Research Square; the manuscript is under review at Scientific Reports.

Süli, P. P., Balázs, E., Eigner, G. (2026).

When Graders Drift and Models Hallucinate: Auditing Human Grading with Local LLMs.

35th EAEEIE Conference, Ruse, Bulgaria · status: presented

A GDPR-compliant grading framework that runs four open-weight LLMs against teacher grades on six C# programming exams (159 submissions, 73 students). The best model reaches 75.3% subtask-level exact-match and ICC 0.974 with the human grader; the ensemble flags 122 likely teacher errors, 85% of which are stricter (not more lenient) than the model consensus. Presented at the 35th Annual Conference of the European Association for Education in Electrical and Information Engineering; proceedings publication forthcoming.

Süli, P. P., Hollós, R., Szabó, G., Eigner, G. (2026).

Syxplain: A Neurosymbolic Framework for LLM-Guided Symbolic Regression and Automated Scientific Discovery.

Poster at HAICON26 · In preparation · status: in progress

A neurosymbolic framework that pairs LLM-guided symbolic regression with automated scientific discovery, closing the loop from data to candidate equation to written explanation. Rather than only steering the search, it interprets the entire Pareto front and grounds every inference in real literature via deep formula research. P. P. Süli and R. Hollós contributed equally and share first authorship.

Buday, V., **Süli, P. P.**, Klofsten, M., Eigner, G., Ge, J., Ding, W., Ni, Q., Wang, F.-Y. (2026).

University Students in the Knowledge Transfer Value Chain: LLM-Mediated Knowledge Sourcing in Higher Education.

Triple Helix · revised after peer review · status: in revision

Examines how LLMs may shape knowledge sourcing, the first stage of the knowledge transfer value chain in higher education. Combines a structured literature review, persona simulation (25 student personas prompted across GPT-4o, DeepSeek, Llama 3.3 and Qwen3), and a small qualitative post-test with real students at Obuda University. The simulated students use LLMs mainly for brainstorming, information acquisition and knowledge structuring, while still preferring face-to-face interaction for collaborative learning and unfamiliar material.