# Shahriar Yazdipour [Call me Sha]

www.yazdipour.com

Software Engineer (Cloud & AI) yazdipour@outlook.com
Phone: +49 163 1932054 linkedin.com/in/yazdipour/
Open to relocate anywhere in Germany & Switzerland

## Summary

Shahriar is a Software Engineer focused on **distributed systems**, **cloud infrastructure**, and **backend platforms**, designing and operating high-availability services on **AWS** and **Azure**. Drives reliability engineering, observability and performance tuning for event-driven and real-time workloads, owning systems end-to-end from architecture and build pipelines to production operations. Brings a strong **ML/LLM** background as a force multiplier for developer productivity and automation.

## Work Experience

- **Ubisoft — Full-Stack Software Engineer, Dusseldorf, Germany**: Aug 2022 — Present
  - Designed and operated event-driven microservices and serverless workloads on **AWS (CDK)** and **Kubernetes** for cloud gaming services to handle volatile traffic, improving high availability and scaling reliably.
  - Optimized databases and caching (query tuning, indexes, connection pooling) for APIs to address latency and throughput bottlenecks, keeping p95 latency within SLOs and increasing request throughput by **40%** during global events.
  - Enhanced production observability (service health checks, metrics, structured logs) for critical services to streamline incident triage, reducing investigation time for high-severity incidents.
  - Standardized **CI/CD** and build pipelines for Java/C#/Python services using **Maven**, multi-stage **Docker** builds, and Git-based workflows (GitLab CI) to enable reproducible builds, artifact/version management.
- **TomTom — Software Engineer, Berlin, Germany**: Aug 2020 — Jul 2022
  - Developed low-latency, real-time traffic processing services in **Java Quarkus** with Protobuf over REST/MQTT to handle data streams, doubling processed message throughput.
  - Provisioned and operated **Azure Kubernetes Service** infrastructure with **Terraform** and **Helm** to support high-availability deployments, introducing blue/green and canary strategies that removed planned downtime for releases.
- **GFPishro — Mobile Developer**: Oct 2017 — Jan 2018
  - Developed offline-first GIS mobile applications for **Android (Java)** and **iOS (Xamarin/C#)** to support field data collection in low-connectivity environments, improving data capture reliability and reducing manual reconciliation effort.

## Education

- **M.Sc. Computer Science & Systems Engineering**: TU Ilmenau, Germany
  - **Master's Thesis:** Analysis and Implementation of **Transformer Based Text-to-SQL Models** using NLP Techniques.
  - Building an end-to-end ML pipeline for data preparation, model training, and evaluation of transformer-based solutions for Text-to-SQL and also evaluating **LLMs (GPT)** and Transformers (SEOSS) with PICARD for this problem.
  - Conducted a deep evaluation of the effectiveness and efficiency of various encoder-decoder transformer architectures (Google T5, BERT, GPT) and using incrementally parsers for constrained auto-regressive decoding of language models, assessing performance across diverse datasets like WikiSQL and Spider.

## Skills

- **Languages**: Java, Python, C#, TypeScript/JavaScript, C++
- **Cloud & Infra**: AWS, Azure, Kubernetes, Docker, Terraform, AWS CDK
- **Backend and Web**: Spring, Quarkus, ASP.NET, React, Flask, FastAPI
- **Build & CI/CD**: Maven, Docker build optimization, GitHub Actions, GitLab CI, Jetbrains TeamCity
- **AI and ML**: PyTorch, LangChain, DSpy, N8N, Transformers and Data Processing
- **Databases**: PostgreSQL, MySQL, SQLite, DynamoDB
- **Spoken Languages**: English (Fluent), German (B1), Persian (Native)

## Selected Side Projects

- **Harpoon AI Agent — RAG LLM for Docs/Code/APIs**: 2024
  - Built an internal **LLM agent** using **LangChain** and **Azure OpenAI** to automate documentation and code generation over internal APIs.
- **TRMNL E-ink Dashboard Plugins**: 2025
  - Developed two production-grade plugins extending **TRMNL**'s privacy-respecting ambient e-ink dashboard: **Beszel Server Metrics** (system/container/network metrics, scheduled pushes) and **TelegramToTRMNL Bot** (images/PDF/EPUB ingestion).
  - Delivered Dockerized Node.js & Python services with secure access controls and cron-based update pipelines, improving reliability and ease of deployment for self-hosted users.
- **GraphQL Library and TUI for Omnivore Service**: 2024
  - Created **OmnivoreQL** (Python GraphQL client) and **OmnivoreX** (TUI) to simplify automation of Omnivore reading workflows, reducing multi-step tasks to single commands for power users.
  - Set up **GitHub Actions** CI/CD (linting, tests, packaging, PyPI publishing) and semantic versioning to ensure reproducible builds and fast iteration for contributors.
- **COVID-19 Opinion Leaders on Twitter — Data Mining & Graph Analysis**: 2020
  - Processed **300GB** of Twitter interaction data with **Dask** to construct large-scale social graphs and identify opinion leaders around COVID-19.
  - Leveraged **NetworkX**-based graph algorithms and clustering to surface key communities and accounts for further analysis.
- **Text-to-SQL Transformer — Evaluation and PICARD-based Optimizations**: 2022
  - Implemented and evaluated Transformer-based Text-to-SQL models with constrained decoding via **PICARD** to improve execution-safe SQL generation, increasing executable query rates on benchmarks like **Spider** and **WikiSQL** by **15%**.
  - Automated ML experiments with reusable pipelines (data preparation, training, evaluation, error analysis) to accelerate iteration across architectures and hyperparameters.
- **Text-to-SPARQL Question Answering Transformer**: 2021
  - Fine-tuned **T5/BERT** models in **PyTorch** and **Hugging Face Transformers** to translate natural-language questions into SPARQL queries over **DBpedia** and **Wikidata**, improving end-to-end QA accuracy on curated benchmarks.
  - Built data processing, training, and evaluation tooling to support reproducible research and deployment of the QA system.
- **MIRA Robot Navigation System**: 2019
  - Implemented **optical flow** and **color segmentation**-based navigation using the onboard camera in C++ with CUDA-accelerated **OpenCV**, achieving stable real-time performance on embedded hardware.
- **Open Source**: github.com/yazdipour

## Publications

- **GitHub Data Exposure using GraphQL Security Design Flaw (1k$ Bug Bounty)** 2019
  *S. Yazdipour, CCSE 2020 — 10th Int. Conf. on Computer Engineering, Ilmenau, Germany* *Published*
- **Decision Support for International Students Who Want to Fund Their Graduate Studies** 2017
  *S. Yazdipour, N. Taherian, IEEE Int. Conf. on Machine Learning and Data Science, India* *Published*