

CONTACT

Berlin, Germany
jakubfronczyk@gmail.com
+49 175 2297118
jakubfronczyk.dev
github.com/jakubfronczyk
linkedin.com/in/jakub-fronczyk

SKILLS

Programming
Go (Golang), TypeScript/JavaScript, Python, SQL, Bash

Infrastructure
Kubernetes (k3s, bare-metal), Docker, Terraform, Linux, CI/CD (GitHub Actions)

Cloud
AWS (ECS, Fargate, Lambda, EC2, API Gateway, EventBridge, Batch, S3, Cognito), Cloudflare

Observability
Prometheus, Grafana, Graylog, GlitchTip, structured logging, load testing

Data
PostgreSQL, Neo4j, Redis

APIs & Messaging
REST, SSE/streaming, BullMQ (Redis-backed job queues)

AI Serving & Integration
LLM inference serving with vLLM (self-hosted open models, continuous batching, OpenAI-compatible API), GPU serving on-prem and in-VPC, GraphRAG and retrieval, OpenAI, Anthropic, Gemini, tool calling, structured outputs, token accounting

Backend
FastAPI, Pydantic, SQLAlchemy, asyncio, Node.js, NestJS, tRPC, Bun, React/Next.js

Languages
English (C1), German (A2)

EDUCATION

B.Sc. Computer Science
University WSB Merito, Poznań
Oct 2018 – Mar 2022

Jakub Fronczyk

Senior Platform & Infrastructure Engineer · ML & Inference Serving

Platform and infrastructure engineer, 4+ years, working across cloud infrastructure, service architecture, and the product on top of them. TypeScript and Python day to day, with Go for performance-critical paths. I build the AI application layer that connects models and agents to real products, and the automated infrastructure underneath it, with AWS managed as code in Terraform, containerized deployments, CI/CD, and observability. I work closely with ML teams and build the layer that takes their models from research into production, on managed cloud or fully on-prem where data cannot leave the network.

EXPERIENCE

- Full-Stack Engineer (Platform & Infrastructure)

Backwell Tech

Berlin

Nov 2025 – Present

Full-stack and platform engineering across service architecture, data modelling, end-to-end application development, and the layers that serve the ML team's models in production.

Tech stack TypeScript, Python, Go, PostgreSQL, Neo4j, Docker, Terraform, AWS

KEY ACHIEVEMENTS
 - Own the inference and model-serving layer for the ML team's prediction and optimization models, and the deployment topology under it. The same containerized stack deploys to managed AWS or entirely inside a customer VPC, self-hosting open models through vLLM on GPU instances where data cannot leave the network.
 - Led Production Intelligence, an ontology-driven knowledge platform built on a Neo4j graph layer, with GraphRAG retrieval over the client's technical documents and ingestion adapters that let their existing systems write into the graph. Built as independent microservices so the deployed scope moves per client.
 - Built and installed the client's production-tracking app, now in daily use on their shop floor and feeding the platform as a data source. Did the on-site discovery and installation, and reworked the domain logic on site once the real workflow turned out to differ from the agreed specification.
 - Built a multi-tenant document service on PostgreSQL. Any internal app connects as a tenant, generates batches of 10,000+ PDFs with workers autoscaling to queue depth, and sends them as physical mail.
 - Set the team's engineering standards, and own the AWS account and CI/CD behind them. Infrastructure as code in Terraform, Docker images built and deployed from GitHub Actions, service and database ownership rules, and a templated repo every new service starts from.

○

Full-Stack Engineer

MAGIG Design + Technologies

Berlin

May 2023 – Nov 2025 · [getcassi.ai](#)

Core-team full-stack engineering on an AI-powered social media assistant, across backend services, frontend architecture, schema design, and the cloud infrastructure and release pipeline behind them.

Tech stack TypeScript, React, Next.js, Bun, tRPC, NestJS, PostgreSQL, AWS, Terraform, Docker

KEY ACHIEVEMENTS
 - Owned the deployment model for CASSI 1.0, an enterprise platform running entirely inside customer VPCs. Every generative model self-hosted on GPU instances, containerized with Docker, and served in-network with no calls to external providers, so customer data never left the perimeter and the product met hard data-residency requirements.
 - Rebuilt CASSI 2.0 as one of two engineers, migrating the product from that VPC-isolated enterprise deployment to a public B2C SaaS and re-architecting for model flexibility so newly released generative models could be adopted as they landed. Worked across the whole monorepo: backend, frontend, and the AWS infrastructure in Terraform, shipping brand-PDF analysis, persona-driven generation, multi-platform publishing, scheduling, OAuth, and Stripe billing.
 - Sole developer of a generative-AI experience used at internal Siemens events, returning a finished AI-generated image composited from an attendee selfie and event copy.
 - Owned the frontend of Vodafone's Next Generation Connectivity Platform, an AR/VR experience streamed from Unity to the browser, and built the AWS infrastructure behind it, from S3 build distribution and CloudFront CDN to Cognito authentication and per-user load-balanced instances autoscaled to session demand.

○

Frontend Developer

Nerdbord

Remote

Dec 2022 – May 2023

Built the company's site end to end on Next.js, with OAuth login, role-based access control, and subscription-gated routes.

Tech stack TypeScript, Next.js, React, NextAuth