



Filipp Trigub

AI Systems Engineer

Postal address

Friedrich-Wöhler-Str 1

53117, Bonn

Phone: +49 178 4042916

E-mail: filipp@trigub.tech

[LinkedIn](#) | [Web](#) | [GitHub](#)

Everyone wants to build a cool product, and I got the know-how to do so.

Last year I've **made 4,419 contributions on GitHub** to projects that I've pushed to production ready state. My work focusses on AI systems and integrations via clients, API and MCP, but even more importantly includes backend, frontend, cloud, testing, CI/CD and mobile. I build reliable solutions and am happy to do so with you.

Theoretical physicist with experience in **tech and strategy consulting**. Proven record of converting client needs into tech architecture proposals, improving these iteratively and implementing the final solution.

Previous clients include startups, consultancies, major banks, an **international central bank** and multinational companies in the pharmaceutical and cloud hosting industry.

Specialized in **Agentic AI application development**, bringing complex multi-agent systems from PoC to production. Beyond that, I am a seasoned **Python (9 years)**, **TS (3 years)** & **SQL (5 years)** developer with broad T-shaped skills including data processing, training and model implementation.

Comfortable working independently or in teams, remote or on-site.

Education

M.Sc. in theoretical Physics
Humboldt university of Berlin

GPA 3.6 (GER scale 1.4)
10.2013 – 03.2019

Thesis:

- “Makovian and Non-Markovian Dissipation Mechanisms in Nonequilibrium Dispersion Forces.”
- Modelled stochastic processes with memory effects mathematically. Validated the model numerically using an implementation in C.

Relevant knowledge

- Statistical Data Analysis: Mathematical foundation and practical application of data processing and machine learning on real world data.
- Statistical Mechanics: Theoretical Statistics at an advanced level with coding experience in Python, C and fortran.
- Fluctuation Induced Phenomena: Deep understanding of complex processes driven by randomness fostered by advanced mathematical and numerical exercises.

Specialization

End-to-end solution architecture for agentic AI systems applications

Coding Languages

Proficient: Python, TS, SQL

Familiar: R, C, C++, Flutter

Tools

Agentic: Claude Code, Cline, BMAD

LLM: LangGraph, Haystack, MCP, Vertex, Anthropic

ML: Tensorflow, Keras, SkLearn, PyTorch

Infra: Docker, Terraform, Azure CLI, GCP CLI, GitHub Actions, uv, poetry, GitLab

Cloud: Azure, AWS, TerraForm

Database: Oracle, PostgreSQL, MySQL, MongoDB, Weaviate, Qdrant

Languages

Bilingual: English, German

Fluent: French, Ukrainian

Professional project experience



— Title: Collaborative multi-agent community AI platform

Client: huuh.me

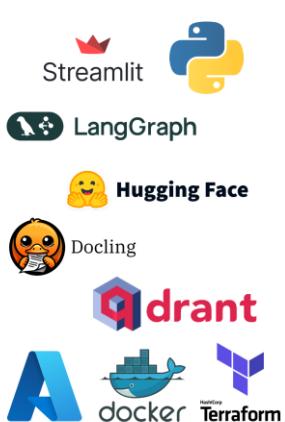
Duration: 6 months (01.2025-07.2025)

Goal:

- Enable content creators and community leads to monetize content AI-natively.

Solution:

- Built a customizable knowledge base to power a RAG system with semantic chunking, multi-format input, page attribution and metadata filtering.
- Built responsive react frontend, and backend systems handling auth, data management and auxiliary services as a RESTful API.
- Enabled interaction with the base powered by a multi-agent system.
- Enabled role-based access to base and agent to foster community collaboration through contributions and sharing of knowledge and AI tools.
- Enabled marketplace and monetization via stripe integration.



— Title: Tax consulting bot

Client: colmo Ltd.

Duration: 3 months (09.2024-12.2024)

Goal:

- Enable clients of the consulting company to get AI-generated answers they can trust.

Solution:

- Understood the proprietary data.
- Developed custom parsers for different data structures.
- Developed custom chunking approaches to preserve coherent text chunks whenever possible and reasonable.
- Implemented alignment optimization based on identified topics and base search as a fallback retrieval strategy.
- Deployed with UI and backend to the cloud.



— Title: Rapid MVP development for geo data startup

Client: HB Partners

Duration: 2 months (06.2025-08.2025)

Goal:

- Develop a presentation ready web app interacting dynamically with local KMZ files and supabase in less than 50 hours.

Solution:

- Established clear coordination with project lead.
- Refactored the existing codebase.
- Formulated clear requirements and implementation steps.
- Developed rapidly with Claude Code and agentic coding tools with a frequently shifting db schema and codebase layout.
- Assured the required functionality before deadline and within the budget of 50 hours.

Professional project experience



– **Title:** Multi-chain LLM copilot for academic teaching and studying
Employer: infolab.ai
Client: Startup
Duration: 14 months (10.2023-12.2024)

Goal:

- Build a sophisticated AI copilot to augment the students' learning experience and provide AI-derived insights to professors.

Solution:

- Build a multi-chain LLM system adapting to user needs at its own accord with a Weaviate vector DB based RAG system and evaluated it with Ragas.
- Build responsive react frontend, and backend systems handling auth, data management and auxiliary services as a RESTful API.
- Deployed and managed the app to the cloud in a production environment including the CI/CD via multi-stage deployment.



– **Title:** Full-stack application design and development
Employer: d-fine GmbH
Client: International central bank
Duration: 24 months (01.2021-12.2022)

Goal:

- Design and develop an application for statistical data analysis holistically and support users on-demand.

Solution:

- Designed and developed architectural extensions based on client demands and limitations.
- Successfully implemented over 500 stories in a 300k+ LoC codebase, which included a dockerized Python backend hosted in AWS with PostgreSQL and Oracle DB, as well as a JS frontend, both delivered via CI/CD.
- Ensured continuous stability through extensive unit and end-to-end testing.
- Served as Scrum Master for a team of 7 developers for over a year.



– **Title:** GPT-powered service staff assistant
Employer: d-fine GmbH
Client: International hosting services provider
Duration: 2 months (04.2023-06.2023)

Goal:

- Provide a GPT-powered chatbot using internal documents to support service staff and interact with customers directly.

Solution:

- Led the design and development of a lightweight GPT-powered chatbot for service-staff support for an international client.
- Contributed significantly to the initial design, technology selection, and architecture.
- Implemented a haystack pipeline with OpenAI embeddings to optimize GPT usage and integrated CI/CD and DevOps for rapid, collaborative development.
- Spearheaded technical development and introduced agile development practices, setting guidelines and conventions for the team.

PoC project experience



Title: Telegram Search&Chat

Client: SME

Duration: 1 month (07.2023)

Goal:

- Make Telegram searchable. Allow the automated exploration of a multitude of telegram channels via a GPT-bot fed by vectorized documents found by a search mechanism.

Solution:

- Retrieved chat contents from Telegram with Telethon and loaded these to a FAISS DB with OpenAI embeddings hosted in a FastAPI backend.
- Implemented question-answer pipeline with Haystack, using Dense Passage Retrieval and OpenAI generation.
- Provided a light-weight UI with user authentication to Telegram and subsequent chat window.
- Deployed the bot to Azure with Docker.



Title: AI Sommelier

Client: Online Retailer

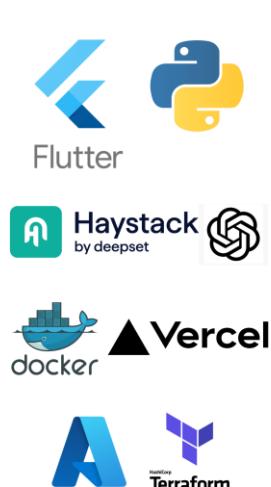
Duration: 1 month (08.2023)

Goal:

- Build a sales bot for a wine store with consistent identification and memorization of customer preferences to provide the customer with a sommelier-like experience.

Solution:

- Built an LLM agent with a recommendation tool based on a general haystack agent pattern design with a Streamlit UI.
- Implemented a specified meta-prompt to correctly and consistently identify customer preferences and remember them.
- Extracted these preferences programmatically to force usage of tool after a set number of preferences has been identified.
- Dockized and deployed to Azure with Terraform.



Title: Audio-Summarizer

Client: Startup

Duration: 1 month (09.2023)

Goal:

- Summarize audio recordings, specifically books, into concise bullet-point style summaries.

Solution:

- Converted audio recordings into text using STT, in this case Whisper and Huggingface models.
- Engineered a meta-prompt and summarization loop subdividing the recording in user designated parts (chapters) and then token-limit enforced chunks to extract a sensible and concise summary for each part as well as the entire recording.
- Build a Flutter based UI allowing users to summarize their recordings and receive the summary by mail.
- Dockerized and deployed to Vercel and Azure with Terraform and CI/CD.