

# Joel P. Barmettler

Senior Software Engineer, AI/ML

*Published researcher. Startup founder. Enterprise AI architect.*

Zürich, Switzerland · joel.barmettler@gmail.com · +41 78 912 86 56

German (native), Swiss German (native), English (full professional proficiency)

github.com/joelbarmettlerUZH · linkedin.com/in/joel-barmettler-b9ab361b7 · joelbarmettler.xyz

Podcast · Google Scholar

## Experience

---

### Senior AI Consultant / Architect

September 2023 – present

*bbv Software Services AG, Switzerland*

- Architected and led a team of 5 engineers building the Swiss AI Hub, an enterprise AI infrastructure platform. One-command deployment of 35 integrated containers, runnable air-gapped on a single GPU or scaled on any Swiss Cloud provider. Deployed across 15 projects for 12 enterprise clients.
- Designed the Swiss AI Agent Protocol: typed event workflows over NATS/JetStream with dependency injection, human-in-the-loop escalation, and MCP tool integration. Built three SDKs (Agent, Pipeline, Process) so developers ship agents that inherit auth, tracing, streaming UI, and cost tracking out of the box.
- Built the multi-agent runtime: agent-to-agent delegation, dynamic agent discovery and registration, shared user- and org-level memory (Mem0), and agentic processes that orchestrate multi-step workflows across specialized agents. Every decision traceable end-to-end via Langfuse and OpenTelemetry.
- Engineered the data pipeline architecture: document ingestion from 70+ cloud storage providers via Rclone, PDF/OCR parsing (MinerU), chunking, embedding, and vector indexing (Milvus) with full lineage tracking in Dagster.
- Drove the technical strategy, pre-sales engineering, and product roadmap that established a new revenue-generating AI business unit within the company from the ground up. Led two large government tenders and 12 enterprise client engagements.

### Founder & CTO

April 2019 – August 2023

*PolygonSolutions GmbH, Switzerland*

- Co-founded a software consultancy at 22 while studying at UZH. Scaled the team from 3 to 18 engineers, delivering 62 projects for 25 clients across 5 industries, managing multiple projects in parallel.
- Led technical delivery of key client projects including a computer vision package recognition system for Swiss Post (TensorFlow, OpenCV) and AI-supported solutions for Klapp.
- Owned all business functions: technical roadmapping, financial planning, talent acquisition (40+ coding interviews, 26 hires), and client relationships (100+ sales meetings).
- Established the engineering culture: code review standards, automated testing, mentorship program coaching junior engineers into senior roles. Built the full-stack practice around Vue/Nuxt, Python, and AWS.

### Artificial Intelligence Engineer (Internship)

March 2019 – July 2019

*BE Semiconductor Industries (Besi), Zug, Switzerland*

- Developed a CNN-based defect detection system for semiconductor die bonding, replacing a traditional computer vision algorithm. The model achieved 99% accuracy at production speeds of several thousand chips per hour with micrometer-precision requirements.
- Managed the full data science lifecycle: data collection, labeling, model training (TensorFlow, TensorBoard), and integration into the existing production line. Trained the R&D team through workshops and documentation, establishing the foundation for subsequent ML projects at Besi.

### Lead Data Engineer

August 2018 – March 2019

*coinpaper.io, Switzerland*

- Architected ETL pipelines (Python, Spotify Luigi, AWS) processing gigabytes of daily data from GitHub, Reddit, Telegram, and Google Trends. Built sentiment analysis and scoring algorithms for automated fundamental analysis.

## Education

---

### Master of Science in Informatics

University of Zurich UZH, Switzerland

*summa cum laude*

Graduated July 2024

*Grade (max 6.0)*

Major 90: Artificial Intelligence

5.9

Minor 30: Informatics

5.3

Thesis: *ConceptFormer – A Novel Architecture for Integrating Knowledge Graphs into Large Language Models, achieving a 348% improvement in fact-retrieval accuracy.*

6.0

### Bachelor of Science in Informatics

University of Zurich UZH, Switzerland

*summa cum laude*

Graduated April 2021

*Grade (max 6.0)*

Major 150: Software Systems

5.6

Minor 30: Neuroinformatics

5.4

Thesis: *Implementation of a physically-based sun and sky model into the open-source rendering engine *apleseed*, navigating a 600,000+ line C++ codebase.*

6.0

## Publications

---

Joel Barmettler, Abraham Bernstein, and Luca Rossetto. *ConceptFormer: Towards Graph-Native Grounding of Large Language Models via Latent Concept Injection*. In Companion Proceedings of the ACM Web Conference 2026 (WWW '26), GLOW Workshop. 2026.

## Skills

---

**AI Platform Architecture** — Designed end-to-end enterprise AI infrastructure: LLM gateway and inference (LiteLLM, vLLM, Speeches), asset-based data pipelines (Dagster, Rclone, MinerU), vector search and RAG (Milvus, BGE-M3), agent memory with knowledge graphs (Mem0, Neo4j), security and data sovereignty (Keycloak, Presidio, Traefik). Agent evaluation and end-to-end observability (Langfuse, OpenTelemetry, SigNoz).

**Software Engineering** — Python, TypeScript. Extensive production code across dozens of projects in multiple industries. Backend (FastAPI, NestJS, Express), cloud (AWS, Azure, GCP, OpenStack), databases (PostgreSQL, Milvus, Redis, MongoDB), frontend (Vue, Nuxt). Protocols and integration: event-driven messaging (NATS/JetStream), REST, WebSockets, SSE, MCP. AI-assisted development with Claude Code (Skills, MCP servers, subagents).

**AI/ML Research** — Knowledge graph integration into LLMs, factual grounding, embedding space manipulation. Model training on NVIDIA GPU cluster infrastructure. Hands-on with PyTorch, Hugging Face Transformers, PyTorch-BigGraph, Weights & Biases, SPARQL, and Optuna for hyperparameter optimization.

**Technical Leadership** — Scaled a team from 3 to 18 engineers. Architecture decisions, code review, hiring, mentoring, and pre-sales. Co-founded and ran a software company for 4 years. Built a new AI business unit at bbv.

## Talks & Media

---

**Attention Heads** (Podcast)

2023–2026

German-language AI podcast, 27 episodes covering technical deep-dives, industry analysis, and social implications.

**CH Open Source AI Conference** (Workshop)

2026

Building a RAG system from scratch in 6 hours with open-source coding agents.

**Swiss AI Impact Forum** (Panel)

2024

AI's technical, ethical, and economic dimensions with SwissCognitive, ZHAW, and ETH AI Center.

**SWICO** (Keynote)

2024

AI fundamentals and practical adoption for Swiss SMEs. Zurich.

**Swiss AI Conference** (Workshop)

2024

Hands-on AI Agents workshop using the Swiss AI Hub. Rotkreuz.

**bbv Software Services** (Keynote & Webinars)

2023–2026

Opening keynote on generative AI. 6 public webinars on ChatGPT internals, AI agents, and AI trends.