

Alexey Berezka

AI & DIGITAL TRANSFORMATION
CONSULTANT · AI / LLM ENGINEER

CONTACT

@ almiberezka@gmail.com
☎ +49 176 4295 0906
✚ Germersheim, Germany · open to relocation
in linkedin.com/in/alexberezka
🔗 berezka.me

AVAILABILITY

From October 2026 · Full-time · Germany-wide or remote

EU work permit · no sponsorship required

LANGUAGES

Russian	Native
English	Fluent
Czech	Fluent
German	B2

CORE SKILLS

AI & Generative AI
LLMs (Claude, GPT-4o, Gemini) · RAG · Agentic / Multi-Agent Systems · LLMOps · Prompt Engineering · MCP · Embeddings / Semantic Search · Computer Vision

AI Strategy & Transformation
AI Adoption · Use-Case Identification · BPMN · ROI / Business Case · Stakeholder Management · MCDM / AHP

Programming
Python · TypeScript · C/C++ · SQL · Bash

Infrastructure & MLOps
Docker · Linux · nginx · CI/CD (GitHub Actions) · FastAPI · Git · n8n · Power BI

Governance & Domain
EU AI Act · GDPR · ISO/IEC 42001 · ISO 9001 / IATF 16949 · TISAX · Automotive · Manufacturing SMEs

PROFILE

I **build and run production AI** — and make it pay for the business. 5+ years of engineering across automotive, aviation and analytics; today I operate **LLM systems daily**: RAG pipelines, a **200+ agent orchestrator**, and LLMOps with an eval gate that **cut model cost ~95%**. Currently driving the **AI & digital-transformation** programme at **Kyocera Automotive**; M.Sc. at KIT (thesis: decision models for AI adoption in manufacturing). At home in both boardroom and codebase — ROI, stakeholders, **EU AI Act / GDPR** — in DE / EN / CZ / RU.

EXPERIENCE

AI & Digital Transformation (Werkstudent) May 2026 – Sep 2026
Kyocera Automotive and Industrial Solutions GmbH · Dietzenbach, Germany · KI und Digitale Transformation

- ▶ Greenfield **AI & digitalisation programme** for the European B2B sales hub — coordinating bilingually (DE/EN) between business, IT and the parent group.
- ▶ **Process discovery**: stakeholder interviews and workshops, **BPMN** maps of quality, customer-service and sales workflows, **use-case identification**.
- ▶ **Tool evaluation & pilot**: **RAG**, **Copilot Studio**, **Azure OpenAI** (EU Data Zone), **Claude API** — vendor evaluation and honest ROI case under **EU AI Act / GDPR**.

Embedded Software Developer Dec 2019 – Mar 2023
ZF Engineering · Pilsen, Czech Republic · automotive ECU development

- ▶ Built automated HiL/SiL validation pipelines for safety-critical ECUs across four OEM programmes — **cut validation cycles by 25%**.
- ▶ Led release-readiness reviews under ASIL-D; defined quantitative go/no-go criteria — the gating logic my KIT thesis now formalises.
- ▶ Established structured 8D root-cause analysis; defect recurrence fell quarter over quarter.

Engineer — Aviation Measurement Systems Jul 2017 – Jun 2018
Research & Production Association "Typhoon" · Russia

- ▶ Designed measurement-system architecture for aviation weather monitoring under strict regulatory constraints; translated certification requirements into technical specs — the same feasibility-assessment discipline now applied to AI use-cases.

Data Analytics Engineer May 2016 – Jun 2017
Stream · Moscow, Russia · marketing analytics

- ▶ Built customer-segmentation models (SAS + Python) and ETL pipelines — **cut manual reporting workload by 40%**; converted weekly reports into self-service dashboards.

SELECTED AI SYSTEMS — BUILT & OPERATED IN PRODUCTION

JARVIS — Multi-Agent AI Assistant. Self-hosted 24/7 system: a meta-orchestrator coordinating **200+ specialist agents** over custom MCP servers, with voice (Whisper) and email/calendar tools. *Python · Claude API · MCP · Docker*

DeepSignal — Autonomous LLM Pipeline. Ingests market & news data and publishes analysis on schedule, zero manual steps; vendor-neutral model adapter (Anthropic / OpenAI / DeepInfra / Ollama) plus a shadow-eval gate **cut model cost ~95%**. *LLMOps · Eval gating · Backtested*

Hybrid RAG Knowledge System. Sub-second grounded search over a 4,500-note base — dense embeddings (bge-m3) + sparse FTS5 + reranker, ~2,000-document index, nightly reindex. *RAG · Hybrid retrieval*

Self-Hosted AI Infrastructure. Production stack on a dedicated server — Docker, nginx, systemd, CI/CD, FastAPI, custom MCP servers; vendor-portable by design. *Docker · CI/CD · MCP*

EDUCATION & RESEARCH

M.Sc. Information Systems Engineering & Management — Karlsruhe Institute of Technology (KIT) · HECTOR School · **Sep 2024 – Oct 2026**

Master's Thesis: Multi-Criteria Decision Model for Evaluating AI-Based Image Processing Feasibility in Manufacturing SMEs (submission Sep 2026).

M.Sc. Electronics & Applied Informatics — University of West Bohemia, Pilsen · **Sep 2019 – Aug 2021**

Thesis: Development of Electronic Systems for a Stratospheric Balloon Mission. Focus: embedded systems, hardware-software co-design, C/C++.

B.Sc. Radio Engineering (Telecommunication) — MIREA, Moscow · **Sep 2014 – Aug 2018**

Thesis: Space Multi-Scale Phase Ranging System. Focus: telecommunication systems and signal processing.

Applied research: adversarial backtesting & walk-forward validation of ML trading strategies; multi-strategy portfolio (Sharpe 3.78, walk-forward-validated 2013–2025).