

Sebastian Schkudlara

Software Architect | AI Solutions Architect

[linkedin.com/in/schkudlara](https://www.linkedin.com/in/schkudlara) · Germany / Spain

Languages: Spanish, German, English, Polish.



Profile

Software architect with 20+ years of experience in distributed systems, enterprise integration and technical leadership. Focused on AI architectures combining agents, retrieval-augmented generation (RAG) and model integration.

I translate business requirements into AI system architectures and phased delivery plans, defining how agents, models and enterprise systems work together. I combine architecture and technical leadership with hands-on implementation and production experience.

Core Expertise

- **Architecture:** system design, enterprise integration, APIs, microservices and event-driven systems.
- **AI systems:** agent and multi-agent architecture, workflow orchestration, MCP, RAG, model integration and human approval controls.
- **Deployment & operations:** AI models on private clusters, automation, CI/CD, evaluation and observability.
- **Technical leadership:** client-facing solution design, architecture reviews, delivery planning and team coordination.

Recent Experience

ITZBund · IT Services for the German Federal Administration

AI Platform / DevOps Consultant | Aug – Oct 2026

- Prepared and validated AI model deployments on private clusters for GPU-accelerated inference.
- Refactored Helm and Kustomize configurations for model deployment.
- Automated model configuration and regression testing across environments.
- Investigated integration issues and documented technical changes.

AI.IMPACT / Hoffmann und Campe · Hamburg

AI & Backend Architecture Consultant | Apr 2026 – Present

- Translate publishing processes into AI solution architectures and phased delivery plans, integrating enterprise data and systems.
- Implement model-output validation, decision traceability and human approval gates before actions are executed.
- Design and develop solutions using FastAPI, PostgreSQL, React and enterprise APIs.

TRAYLIX

Product Architect & Senior AI Engineer | 2024 – Present

- Design and build agent-based AI platforms, defining agent coordination, tool integration, information retrieval and access controls.
- Design and implement APIs and distributed services, with automated tests and deployments.
- Develop real-time voice infrastructure using WebRTC, speech recognition, streamed model responses and text-to-speech.

Professional Experience & AI Platforms

Monodon / Navantia · Madrid

Deeptech AI Specialist | Jul 2023 – Apr 2025

- Architected and implemented RAG systems for an industrial and maritime environment.
- Fine-tuned models and integrated AI solutions with enterprise systems.

Technologies: Python, PyTorch, Hugging Face, LangChain, vector databases and Kubernetes.

Babel Group · Madrid

AI Engineer | Jun – Dec 2024

- Developed enterprise RAG systems and NLP pipelines for summarization, entity extraction and sentiment analysis.
- Optimized model latency and operating costs, and integrated AI services with enterprise data infrastructure.

Technologies: Python, Hugging Face, LangChain, TensorFlow, Docker, Kubernetes and PostgreSQL.

ChainGO · Madrid

Head of Development | Feb 2019 – Jul 2024

- Led engineering for logistics platforms, automation and document processing.
- Designed microservices and event-driven integrations; developed and operated the services.

Agentero · San Francisco

Head of Development | May 2017 – Dec 2018

- Grew the engineering team from one to 15 people and introduced Kubernetes and CI/CD.
- Built real-time analytics systems using event-driven architectures.

Earlier Experience

Vizzuality · Backend Engineer | May 2014 – May 2017: APIs and data platforms for environmental and research projects.

Palabea · CTO & Founder | Jan 2007 – Feb 2010: led product, architecture and development of an online learning platform.

Zanox · Product Manager & Technical Lead | May 2003 – Sep 2007.

AI Platforms · TRAYLINX

- **SwitchAI:** multimodel gateway with an OpenAI-compatible API, routing, fallback and provider normalization.
- **Makakoo OS:** AI tool integration platform with shared memory, plugins, MCP, permissions and coordinated workflows.
- **Lope:** validation system using independent agents to review code, plans and technical decisions.

Technology Stack

AI: LangChain, LangGraph, PydanticAI, Hugging Face, PyTorch, TensorFlow, RAG, MCP and A2A.

Development & integration: Python, FastAPI, Ruby on Rails, Go, Node.js, TypeScript, React, REST, GraphQL, gRPC and OAuth 2.0.

Data: PostgreSQL, pgvector, Redis, MongoDB, Neo4j and Elasticsearch.

Cloud & operations: AWS, GCP, Azure, Kubernetes, Helm, Kustomize, Docker, Terraform, GitHub Actions, Prometheus, Grafana, Kafka and NATS.

Education & Links

Business Administration studies · Technische Universität Berlin, 1999–2007.

github.com/rschumann · github.com/makakoo · jevvellabs.com/cv.html