

# Eduard Caizer

## AI Systems Architect | Independent Consultant

Applied AI · Distributed Data · Architecture & Delivery

Based in Italy

### Profile

Hands-on architect and independent consultant with more than ten years in software engineering. I connect applied AI, backend services, distributed data and the infrastructure needed to operate them. My work spans conversational products, industrial monitoring and long-lived enterprise applications. I work from requirements through implementation, technical review and operational follow-through, with practical experience in mentoring and applied research.

### Where I Can Help

**From idea to a working product:** clarify the first useful journey, choose an appropriate architecture and deliver a testable increment.

**AI in an existing system:** connect inference, relevant knowledge and product workflows; measure quality, latency and operational cost rather than assume a larger model fixes the problem.

**Systems that remain workable:** improve data models, observability, release paths and failure recovery without replacing every existing component.

### Experience & Selected Work

#### Syncode Società Cooperativa a r.l.

2018–Present

*Co-founder and board member · Hands-on technical architecture and delivery*

- Co-founded an R&D cooperative working across industrial software, AI, connected devices and enterprise applications.
- Contribute across architecture, implementation, technical review and operations; work directly with client stakeholders, engineers and researchers.
- Mentor contributors and review designs and code while retaining hands-on responsibility for difficult cross-layer integrations.
- Developed a specialised multi-tenant application with university researchers, alongside industrial platforms and enterprise integrations.

#### Production conversational AI

2025–Present

*Self-hosted speech and language-model service · Syncode client delivery*

- Architected and delivered the service path from device sessions through speech recognition, language-model inference and speech synthesis to the application backend.
- Operated the service through the Christmas demand peak; built traffic-forecasting and capacity-calibration tooling to turn observed demand into load-test scenarios.
- Integrated vLLM, Whisper-based recognition and Piper speech synthesis with Kubernetes services and application protocols.
- Diagnosed model-serving memory and telemetry failures; combined recovery procedures with tracing, logs and metrics through OpenTelemetry, Grafana, Loki, Tempo and Prometheus.

#### Industrial data, monitoring & applied ML

2017–Present

*Industrial research followed by cooperative product delivery*

- Designed Cassandra-backed time-series services with tenant-scoped access, prepared reads and writes, and realtime delivery to applications.
- Shaped data layouts around explicit product queries and historical telemetry, balancing useful read paths with partition and retention constraints.
- Worked on industrial monitoring, production-planning, MES and SCADA-style applications connecting machine data with operator-facing workflows.
- Applied machine learning to predictive maintenance, fault detection and computer-vision tasks; co-authored research in these areas.

### Technical Range

**Software & data:** TypeScript, Node.js, Effect-TS, React, Hono, Java/Spring, Python, SQL and Bash; PostgreSQL/pgvector, Cassandra, MongoDB and Redis; REST and WebSocket integration.

**AI & infrastructure:** retrieval and context engineering, local and hosted model integration, vLLM, Ollama, speech pipelines; Linux, Kubernetes/K3s, Docker, Terraform, GitOps and observability.

**Working style:** incremental delivery, architecture and code review, reproducible checks, incident diagnosis and measured trade-offs.

## Independent Platform Development

---

### Rezyak

2025–Present

*Personal platform design and daily use · Building on experiments begun in 2024*

- Build an agent workspace connecting persistent context, typed actions, model execution, local tools and reviewable artifacts.
- Use it daily for software work, research and coordination across coding agents; preserve source context and results so work can continue beyond an individual conversation.
- Develop desktop and web surfaces for multiple workstreams, with explicit session ownership and delivery tracking.
- Explore portable interactive presentations and project kits that remain editable with ordinary development tools. Current development focuses on connecting that review workflow with cloud execution.

## Enterprise & Earlier Engineering

---

### TeamSystem · enterprise commerce platform

2018–2026

*Senior software engineering and technical-lead contributions · Project engagements*

- Contributed to a long-lived commercial platform across React applications, Java services, integration services and data-processing pipelines.
- Worked across requirements, implementation, code review and operational troubleshooting in a multi-team environment.
- Evolved existing components incrementally, preserving product continuity and coordinating interfaces with the wider engineering team.

### Vanillainnovations

2015–2018

*Lead / Head Developer*

## Research, Mentoring & Teaching

---

- Supervised a university internship and mentor engineers through architecture discussions, code review and hands-on technical coaching.
- Co-authored conference research on predictive maintenance, online fault detection, cyber-physical manufacturing, ML-as-a-Service for fault diagnosis and industrial sensing.
- Teach advanced AI and cloud infrastructure; develop practical explanations and interactive learning material connecting concepts with experiments and observable results.

## Education

---

### University of Trento

2019

*Master's degree*

### University of Bologna

2016

*B.Sc. Computer Science*

## Languages

---

Italian · English · Romanian · Russian

## Engagements

---

Independent architecture, technical assessment and scoped implementation work. Availability, responsibilities and contracting arrangements are agreed for each engagement.