Gianluca Mazza

Roles and responsibilities are my account of my professional history. Where public work exists, I link the artifact so it can be inspected.
Canonical: https://gianlucamazza.it/en/cv
I take LLM systems into production: state, recovery, eval, cost.
I have worked in software and infrastructure since 2012. As co-founder and CTO of Scaling Parrots, I led technical decisions on product and infrastructure. Today I help technical teams manage state, failures, output quality and costs in LLM systems.

Contact

- Email: info@gianlucamazza.it
- GitHub: https://github.com/gianlucamazza
- LinkedIn: https://linkedin.com/in/gianlucamazza
- ORCID: https://orcid.org/0009-0005-1462-3019
- Hugging Face: https://huggingface.co/gianlucamazza

AI Lead · confidential project

July 2026 – Present · Italy
I lead the development of an AI platform for a confidential project with a small team: architecture, development assisted by AI agents, CI/CD and releases.

Independent AI consultant · gianlucamazza.it

October 2026 – Present · Italy
I design LLM systems for teams taking them from prototype to daily use.
- I design agent systems that stay inspectable: typed tool contracts, durable state, and recovery paths that survive a restart.
- I build reproducible evals, so a change to a prompt or a model shows up as a measured difference rather than an impression.
- I research AI memory for agents and publish reproducible benchmarks under the emotional-memory project.

Independent research & AI consulting — Venere Labs

September 2024 – September 2026 · Italy
Until September 2026, Venere Labs was the name I used for my independent research and AI consulting work on inspectable agent systems and reproducible evaluation.
- Under that name I also worked on the emotional-memory project.

Related public work:
- emotional-memory — DOI 10.5281/zenodo.19972258: https://doi.org/10.5281/zenodo.19972258

Stack: LangChain, LangGraph, MCP, OpenAI/Anthropic

AI course instructor @ training provider for school staff

September 2026
I taught this course from September 7 to 10, 2026.
- Course for teachers and administrative staff.
- 12 hours / 4 modules.

Co-founder and CTO @ Scaling Parrots

April 2020 – November 2025 · Brescia, Italy
As co-founder and CTO, I led technical decisions on product and infrastructure.
- I coordinated development, product and infrastructure throughout delivery.
- I oversaw code review, continuous integration and release preparation.

Stack: Python, Rust, AWS

Product and systems consultant @ industrial client (external consultancy)

2025
As an external consultant I designed the architecture of a software system and defined the product, from requirements through technical delivery.
- I set the product boundaries around that architecture.
- I translated that design into an implementable technical plan and supported delivery with the client team.

Stack: Product design, Product architecture

Senior software developer @ Bizonweb

November 2019 – April 2020 · Brescia, Italy
I built bespoke web applications for client projects, each to the client's requirements.
- I wrote the backend services and REST APIs behind them, in Node.js.

Stack: Node.js, JavaScript, REST APIs

Earlier experience, 2017–2019: software development and infrastructure, including systems for financial-sector clients

Junior software developer @ MYTI
January 2015 – December 2016 · Brescia, Italy
I built responsive single-page web applications in AngularJS.
▪ I implemented responsive interface components with Angular Material.
Stack: AngularJS, Angular Material, REST APIs

Freelance software consultant
April 2013 – December 2014 · Italy
I took short software consulting engagements. I do not name clients.

DevOps and system administrator @ Atic
October 2012 – March 2013 · Brescia, Italy
I installed, supported and secured fuel-station systems, including redundancy, failover and disaster recovery procedures.
Stack: Linux, Redundancy, Disaster recovery

Microsoft developer, internship @ Lutech
May 2012 – September 2012 · Brugherio, Italy
I worked on Microsoft enterprise and cloud technologies during an internship.
▪ I developed applications on the Microsoft .NET stack.
▪ I deployed services on Azure and integrated enterprise workflows through BizTalk.
Stack: .NET, Microsoft Azure, BizTalk, SQL Server

Selected public work
Public code, documentation and research results.

emotional-memory
Memory layer for LLM systems with affective state encoding, a PyPI package, Zenodo DOI, pre-registered recall tests, and published negative results on LoCoMo, MADial-Bench, and ES-MemEval.
▪ https://github.com/gianlucamazza/emotional-memory
▪ https://doi.org/10.5281/zenodo.19972258
▪ https://pypi.org/project/emotional-memory/
▪ https://github.com/gianlucamazza/emotional-memory/tree/main/benchmarks
▪ https://github.com/gianlucamazza/emotional-memory#validation--benchmarks
▪ https://github.com/gianlucamazza/emotional-memory/releases

orka
Self-hosted Rust runtime that takes work from Telegram, Discord, Slack, WhatsApp, and HTTP, assigns priority, and routes it into bounded LLM workflows with MCP/A2A support. Pre-release, no SLA. An optional loop that turns past runs into prompt rules is off by default.
▪ https://github.com/gianlucamazza/orka

mklang
Declarative language for LLM-driven state machines. Spec 0.4; the 1.4.0 reference interpreter is stable, with check, lint, test, doctor, and a conformance suite. Gate accuracy is measured, not guaranteed, and runs are not guaranteed to be deterministic. The optional hosted judge is experimental and is the only part marked not production-ready.
▪ https://github.com/gianlucamazza/mklang
▪ https://docs.mklang.dev/
▪ https://huggingface.co/spaces/gianlucamazza/mklang-playground

Technical background

AI systems
LangChain, LangGraph, MCP (Model Context Protocol), A2A (Agent-to-Agent), OpenAI / Anthropic / Gemini, RAG systems, Vector databases, Prompt engineering, Function calling, CrewAI, Agent pipelines, MCP servers, WASM, Local inference (Ollama, llama.cpp), Redis

Data & ML
PyTorch, Transformers, Hugging Face, Embeddings and semantic search, pandas / NumPy, Jupyter, PostgreSQL

Infrastructure & security
End-to-end encryption (X25519 / Ed25519), Double Ratchet, AWS, GitHub Actions, Docker, Linux / Unix

Languages & frameworks
Python, TypeScript, Rust, Go, JavaScript, C++, SQL, Bash, React, Node.js, GraphQL, Git / GitHub

Education
▪ Computer Science, Faculty of Mathematical, Physical and Natural Sciences — Università degli Studi di Milano-Bicocca, 2009 – 2011, Milan, Italy, University studies, degree not completed
▪ Liceo scientifico — Liceo Luzzago, 2003 – 2009, Brescia, Italy

Languages
▪ Italian — Native
▪ English — Professional working proficiency