

◆ Pierre Rognion

Applied AI Architect / Solutions Architect · Production LLM systems

Paris, France · Bilingual French / English · pierre@atomly.ai · linkedin.com/in/pierreroignon · atomly.ai · Systems in production (detail): atomly.ai/en/work

I take a business problem to a production AI system, and then get people to actually use it. Head of AI at one of France's largest freelance marketplaces, where I run the discovery, design the architecture, write the code and ship it, for every team that uses AI: Sales, Product, Engineering, Finance, Management. Before the code, twelve years in front of clients: an AI practice co-built from scratch, GenAI proofs of concept taken into production at Bpifrance and Arkéa, and 1,000+ professionals trained in applied GenAI at Chanel, Stellantis, Bouygues Telecom, Rakuten, Generali and Institut Pasteur. I would rather ship a system whose failure modes I can name than one that only demos well.

EXPERIENCE

Crème de la Crème · Head of AI / Applied AI Architect

May 2025 to present

Technical owner of the AI portfolio at a digital-native marketplace where 50,000+ freelancers meet the companies hiring them, engaged through my own company. Five business owners across Sales, Product, Engineering, Finance and Management, each use case carrying its own owner, success measure and stop criteria. Business problem, system design, build, production, adoption. Also shipped a licence-locked expert sourcing tool as a self-service capability for the whole sales team, a marketplace of packaged business skills, and an MCP server over the design system. Retired two early builds once the CRM already in place shipped equivalents.

Atomly AI · Founder & AI Engineer

May 2024 to present

My own company (SASU), through which I run my engagements and my own builds: custom agents, internal prototypes, and hands-on training programmes.

Mozza · GenAI Training Manager

Jun 2024 to Jun 2025

Hands-on GenAI programmes for Product and Tech teams: prompt design, AI-assisted ideation, workflow automation.

Twelve Consulting · Manager, Innovation & AI (Senior Consultant 2016 to 2021)

Sep 2016 to Apr 2024

Co-built the firm's AI practice from scratch: offering design, technical discovery with client executives, use-case scoping, and GenAI proofs of concept taken into production pipelines at Bpifrance and others. Carried several client accounts in parallel and led the consultants embedded in them, around ten across four concurrent projects at Bpifrance alone. Earlier, Product Owner of a Salesforce pricing engine.

Earlier

2012 to 2019

Le Wagon teaching assistant (2019) · CRM integration consulting, Oracle then Salesforce · CGI (2015 to 2016) · DISAPI (2014) · marketing internship in the San Francisco Bay Area (2012).

TECHNICAL SKILLS

How I build: I own the architecture and write the systems front and back, in TypeScript and Python, driving the build through coding agents (Codex, Claude Code) and shipping through the client's engineering review. When a build needs more hands or expertise I do not have, I bring someone in and write down who owns what afterwards. Full-stack foundation from a Le Wagon bootcamp taken in 2019 alongside the job, kept current since.

Languages & platform: TypeScript/JavaScript, Python, React, Node, SQL, PostgreSQL with pgvector, Vercel, PostHog, Git.

OpenAI platform in production: Responses API and schema-validated structured outputs, embeddings behind the retrieval layer, running daily on mission parsing and on part of the matching engine. Also built with: Agent Builder, ChatKit, Codex, image generation, and the Realtime and audio APIs on earlier voice work.

LLM systems: MCP server design (OAuth 2.1 + PKCE, multi-layer authorization, tool allowlists, audit), agentic workflows and sub-agents, packaged skills, structured outputs validated by schema, LLM-as-a-judge, eval harnesses and golden datasets, RAG, cost and latency control. Several frontier providers running side by side, picked per task rather than by default. Knowing where each one wins is the comparison customers ask for before they commit.

Retrieval & delivery: Hybrid search, vector kNN, BM25, Reciprocal Rank Fusion, text-to-SQL. Human-in-the-loop design, guardrails and governance on sensitive data, build-versus-buy calls, adoption by non-technical teams.

EDUCATION & CERTIFICATIONS

Education: Le Wagon, full-stack web development (2019) · Sciences Po Aix, Master's in International Relations and Business (2010 to 2014) · Truman State University, exchange programme in Business Administration (2011 to 2012) · Classes préparatoires A/L (2007 to 2010).

Certifications: Introduction to Model Context Protocol · Salesforce Certified AI Associate · Concepteur-Développeur d'Applications Web (RNCP, French professional qualification in software development).

Built inside a marketplace where thousands of freelancers meet the companies looking for them, and a Sales team has to make that match fast and without seeing data it has no business seeing. Code is client-owned and the repositories are private; architecture, demos and measured results can be walked through in an interview.

Opening a company back office to natural language, without loosening a single permission

Jul to Aug 2026 · pilot ahead of production · MCP · OAuth 2.1 + PKCE · Streamable HTTP · PostgreSQL · 49 tools · sole architect and engineer

- ♦ **What it unlocks.** Anyone in the company asks in one sentence for what used to take a trip through several screens, and the action is carried out. Work too small to be worth delegating can go to an agent, because the agent inherits the rights of the person asking rather than a service account. The same question from two people returns two different answers, and neither learns what the other can see.
- ♦ **How, and the call.** Three layers of authorization checked server side, personas whose scopes are computed from the data rather than transcribed by hand, field-by-field payload projection, and a queryable audit trail. Most of the work was not the API but explicitly answering the security questions an interface carries implicitly: how access is granted, what a departing employee loses and when, what happens the first time a tool is asked to do something irreversible. When a missing OAuth client access blocked the integration campaign I wrote a script through the real authorization loop rather than a bypass. It cost two days.

Hybrid retrieval, ranking and bounded LLM judgment

May to Jun 2026 · in production, daily · OpenAI Responses API and embeddings · PostgreSQL/pgvector · RRF · LLM-as-a-judge · TypeScript

- ♦ **What it changed.** A brief that used to mean a manual trawl now returns a shortlist a salesperson can defend to a client. It returns what matches and stops, instead of padding every brief out to the same fixed number of profiles, and it surfaces strong candidates that keyword ranking leaves below the fold. Rolled out to Sales ahead of general availability.
- ♦ **How, and the call.** Four-signal recall (BM25, skill overlap, vector kNN, an experience index) fused by Reciprocal Rank Fusion, deterministic weighted scoring, then a bounded LLM-as-a-judge capped at one call per candidate, so cost and latency per brief stay flat as the pool grows. The eval harness came before the engine, and the engine was not allowed to ship until it beat the incumbent on a golden dataset rather than on a demo.

Human-in-the-loop decisioning on inbound applications

Dec 2025 · in production, daily · built end to end, then hardened for production with a developer · schema-validated structured outputs · TypeScript

- ♦ **What it changed.** 90 to 95% of the inbound flow is settled without a human, about seven hours a week went back to a Product Manager, and intake stopped depending on one person being at their desk. Validated before go-live against fifty applications already decided by hand, with no disagreement.
- ♦ **How, and the call.** Three-state verdicts scored axis by axis, structured outputs validated by schema, and rule-based routing of every doubt to a human. Triage beats automating to 100%: keeping a person on the hard cases is what made automation acceptable on a decision this sensitive. The schema is the contract, not the model, so the pipeline has crossed model generations without a change to the decision format.

Plain-language access to a production database, under governance

Feb to Mar 2026 · in production, daily · database MCP server and packaged skill · PostgreSQL · TypeScript

- ♦ **What it changed.** Extracts that took from thirty minutes to several hours now run in seconds, self-serve, for a non-technical team. Routine questions no longer queue behind an engineer, and the ones that matter still reach one.
- ♦ **How, and the call.** A database MCP server with schema discovery plus a packaged skill carrying the business context the schema does not, and business rules wired into mandatory query templates rather than left to memory. The agent originally executed on its own; without a safe environment to execute against I restricted it to generating SQL for a human to review and run. A governance decision, not a technical limit, with a documented exit path.