Soufien HAJJI



Lead Software & AI Engineer · Enterprise AI Copilots in Production · Azure

haji.soufien@gmail.com • +33 6 69 04 92 63 • Thionville, France, Luxembourg border • linkedin.com/in/hajji-soufien • github.com/SFHAJJI • **soufien.lu**

PROFILE

Lead Software & AI Engineer with 9 years designing and operating business-critical C#/.NET systems on Azure, from architecture through production. I take AI assistants **from prototype to production operations**: deterministic logic in typed code, grounded and cited answers or explicit refusals, automated evaluations that block regressions, and observable deployments. Architecture decisions are documented with rejected alternatives and reversal criteria. ENSTA ParisTech (Information Systems Architecture & Security). **Microsoft Certified: Azure AI Apps and Agents Developer Associate.** Currently preparing for TOGAF certification.

PROFESSIONAL EXPERIENCE

Software & AI Engineer · Natixis Investment Managers

Oct 2024 - Present · Paris

Own both regulatory-credit workflows, productionized the market-risk integration, then built and deployed the AI layer across all three workflows. The platform covers 8 investment portfolios.

Phase 1 · Solution engineering: regulatory-credit ownership and market-risk industrialization

Own the design, development and production support of two regulatory-credit workflows within a platform covering 8 investment portfolios.

- Designed and built an audited credit-inventory workflow covering ingestion, normalization, historical matching, counterparty identification and data-quality controls, with reproducible checkpoints, rollback and validated enriched outputs.
- Built the internal and securitized look-through reporting workflow, taking exposures through enrichment, contracts, outstanding amounts and securities to auditable regulatory deliverables.
- Delivered the application end to end in **.NET 8**, from workflow and data-processing services to the Razor and SignalR user experience, backed by EF Core and SQL Server.
- Brought a specialist Python market-risk workflow into the platform behind a **FastAPI** service, defining the .NET–Python boundary, orchestration and failure isolation while leaving the quantitative calculations under the specialist component's ownership.
- Took the solution across development, UAT and production on IIS, with AD/LDAP access control, environment-specific configuration, automated build, test and deployment pipelines, and correlated Serilog logs.

`.NET 8 / ASP.NET Core MVC` `Razor / SignalR` `EF Core / SQL Server` `Python / FastAPI` `Full-stack delivery` `Automated testing` `CI/CD / IIS` `AD / LDAP` `Serilog` `RWA · FRTB/SBM`

Phase 2 · AI engineering: internal copilot in production across existing workflows

Designed and deployed an internal copilot across three business workflows within a platform covering 8 investment portfolios.

- Enabled user guidance, pipeline explanations and natural-language analysis over production data. Built a Document Intelligence and RAG knowledge system grounded in technical and functional specifications, with curated tools, mandatory citations, SELECT-only analytical access, guardrails, human review, accuracy and faithfulness evaluations, and production observability.
- The model orchestrates and explains; deterministic services retain control of workflow state and financial calculations.

`LLM agents · tool-use` `RAG · embeddings · vector store` `text-to-SQL` `Responsible AI · guardrails` `prompt engineering` `evals` `Azure App Insights` `Azure AI Document Intelligence` `Python · SQL · C#`

.NET Software Engineer, Loans & Securities · Société Générale

Jul 2021 - Sep 2024 · Paris La Défense

- Developed X-One, a latency- and correctness-sensitive front-office trading platform for the securities lending and borrowing desk, working with trading and operational users in a distributed team of approximately 30 engineers in Paris and 20 in Bangalore.
- Designed and built the deal position-calculation engine and collateral-valuation features used as key inputs to risk exposure. Delivered prepay-offset processing with runtime-selectable strategies for anticipated interest flows.
- Led the migration from .NET Framework 4.8 to .NET 6, introduced dependency injection and strengthened SOLID and test-driven practices. Rotated through front-office production support and a quality task force resolving structural defects before delivery freezes.

`C# / .NET 6` `Front-office trading` `Position engine` `Collateral valuation` `TDD / SOLID` `Jenkins`

.NET / Azure Consultant · SNCF Réseau

May 2019 - Jun 2021 · Paris

- Built a real-time operational train-monitoring application connecting stations with maintenance centres, giving each centre a dynamically configured view of its trains, events and operational workflow.
- Developed the ASP.NET MVC 5 application and ASP.NET Core 3.1 business APIs, using SignalR for live updates and integrated chat. Improved responsiveness and capacity through asynchronous processing, lazy loading and IIS load balancing while bridging legacy and modern .NET components.

ASP.NET MVC / Core · SignalR · Real-time systems · IIS load balancing · Async processing

Azure Consultant / .NET Developer · Orange Business Services

Jan 2018 - Apr 2019 · Massy

- Built an internal event-ticketing platform for concerts, exhibitions and corporate events, including ticket chaining and secure payment, as .NET Core microservices packaged with Docker and exposed through Azure API Management.
- Set up automated build, integration and deployment pipelines in Team Foundation Server, and delivered document-management components integrated with the Box API and Entity Framework.

.NET microservices · APIM · Docker · Team Foundation Server · CI/CD · Box API / EF

PERSONAL PROJECTS · DESIGNED, BUILT & DEPLOYED SOLO

law.soufien.lu · Lex, point-in-time Luxembourg & EU law

2026 · live

Point-in-time Luxembourg and EU law with grounded AI, verifiable provenance and an extensible publisher architecture.

- Built and operate **2,647 Luxembourg and EU works and collections across 9,374 verified dated versions**, behind a publisher-neutral temporal platform.
- Built and operate a **grounded legal research assistant** that resolves professional law names and ambiguous intent, plans bounded point-in-time, search, timeline and comparison calls, and asks for clarification rather than silently selecting a law. An evidence ledger and conditional judge verify model-written claims before citations and deep links are rendered.
- Published a **public MCP server** exposing ten read-only tools for search, point-in-time retrieval, history, comparison, provenance and coverage.
- Put local hybrid retrieval behind public relevance and latency holdout gates; the measured benchmark rejected the semantic candidate, so keyword retrieval remains the production default. Recorded the evidence, rejected alternatives and reversal criteria publicly.

C# / .NET 10 · Legal AI assistant · Tool orchestration · Evidence judge · Temporal retrieval · Signed artifacts · Public benchmarks

energy.soufien.lu · Luxembourg Energy Transition Co-Pilot

2025 · live

Grounded solar-PV prospecting combining geospatial analytics, deterministic services and a production AI assistant.

- Built and operate a **production AI assistant** as one bounded agent over five typed tools: NL-to-SQL, document RAG, legal retrieval, tariff calculations and tender discovery. The model plans and explains while deterministic services retain ownership of calculations, read-only data access and rendered actions.
- Designed the orchestrator as a safety host: it enforces execution limits and typed UI contracts, while a groundedness judge checks model-written prose against the documents and rows actually returned by tools. Direct result cards remain deterministic rather than being rewritten by the model.
- Implemented hybrid RAG using PostgreSQL, pgvector, full-text search and rank fusion. Used ADRs and evaluations to defer Azure AI Search and reject a premium model that increased cost without improving relevance or groundedness.

Azure OpenAI · Production AI assistant · Tool orchestration · Evidence judge · RAG · pgvector / FTS / RRF · NL-to-SQL · PostGIS · Evaluations · Terraform / OIDC

TECHNICAL SKILLS

AI / LLM	LLM agents & tool-use (function calling) · **Microsoft Agent Framework** · RAG (chunking, embeddings, pgvector) · NL-to-SQL · **MCP** (Model Context Protocol) servers · prompt engineering · evals · Responsible AI / guardrails · Azure OpenAI · Content Understanding · Document Intelligence
Languages	C# · Python · TypeScript / JavaScript · SQL
Backend	ASP.NET Core 8 / MVC · EF Core · FastAPI · SignalR · REST & SOAP
Data	PostgreSQL / PostGIS / pgvector · SQL Server · Oracle · SQLite / FTS5 · pandas · reproducible pipelines & content-addressed provenance
Frontend	React · Vue 3 / PrimeVue · MapLibre · Vite · Bootstrap
Cloud / DevOps	Azure (Container Apps, Static Web Apps, OpenAI, AI Foundry, Functions, Key Vault, App Insights, APIM) · Terraform (IaC) · GitHub Actions & Azure DevOps · Docker · IIS
Practices	SOLID / clean architecture · TDD & unit testing · AI Act & RGPD governance · mentoring & documentation

CERTIFICATION

Microsoft Certified: Azure AI Apps and Agents Developer Associate

LLM integration · Generative AI & agentic solutions · Planning Azure AI solutions · Information extraction · Text analysis · Computer vision

Earned Jul 2026 · exam AI-103 · verify on Microsoft Learn ↗

In preparation: TOGAF Enterprise Architecture.

EDUCATION

ENSTA ParisTech, Engineering degree (Bac+5)

Architecture & Security of Information Systems · 2015-2018

LANGUAGES & INTERESTS

French, bilingual · English, fluent (TOEIC 895/990) · German, basics

Coding challenges (9th / 148, Spring Challenge 2020) · Unity game dev · strategy games