

ATS-Optimized Resume - AI Software Engineer

ALEX COSTA

AI Software Engineer | Solutions Architect | Full-Stack Lead

Lisbon, Portugal | +351 910 695 306 | lekodeveloperweb@gmail.com | [LinkedIn Profile](#) | [Personal Portfolio](#) | [Career Projects](#)

SUMMARY

AI Software Engineer and Solutions Architect with 14 years building enterprise systems (.NET, Node.js, React) and 2 years architecting production GenAI platforms. Expert in deploying AI-native platforms through Model Context Protocol (MCP) integrations, Retrieval-Augmented Generation (RAG) pipelines, and custom autonomous agents, scaling LLM orchestration across OpenAI, Anthropic, and Gemini APIs. Proven track record leading cross-functional engineering teams, modernizing legacy architectures, and accelerating delivery velocity by up to 95%.

EXPERIENCE

LDW Solutions / Axians – Technical Lead & Solutions Architect – Lisbon, Portugal

Jan 2026 – Jun 2026

- Led a 4-person cross-functional team for a government judicial client, bridging technical squads and stakeholders as Scrum Master to deliver predictable sprints via Azure DevOps.
- Spearheaded the modernization of legacy applications, migrating from .NET Framework to .NET 8 and React 19 using Clean Architecture principles.
- Established testing standards with Vitest (unit) and Playwright (E2E), reducing production defects by 90% across judicial case workflows.
- Architected frontend state management using Jotai and optimized SQL Server database layers to ensure high-security, reliable data handling for court applications.

LDW Solutions / CoolPeople / PwC – AI Platform Tech Lead & Solutions Architect – Prague (Remote)

Jul 2025 – Jan 2026

- Led a multidisciplinary squad of 8 engineers (Seniors, QA, DevOps) in an R&D initiative building an AI-native engineering platform using .NET Core/Node.js, PostgreSQL, and Azure.
- Architected Model Context Protocol (MCP) integrations connecting enterprise tools to multi-model LLM APIs (OpenAI, Gemini, Anthropic), unifying AI workflows and reducing context-switching overhead by 70%.
- Built custom AI agents automating code generation and test suite creation, reducing manual workflow overhead by 80%.
- Established a Spec-Driven Development workflow utilizing Cursor AI IDE and automated code scaffolding, increasing squad delivery velocity by 95%.
- Orchestrated scalable cloud deployments on Azure Cloud using Docker and Kubernetes, ensuring enterprise security compliance.

LDW Solutions / Irysan / Condé Nast – Senior Full-Stack Developer (Node.js / React) – London, UK (Remote)

Nov 2024 – Oct 2025

- Architected high-concurrency microservices for Vogue's editorial platform using Node.js (NestJS) and React 18 with Zustand for state management, integrating secure content delivery APIs (Spectra API).
- Designed robust REST APIs and MongoDB schemas supporting real-time collaborative workflows during high-traffic global launches (e.g., Vogue Fashion Week), handling 5,000+ concurrent users with <10 ms latency.
- Managed containerized AWS deployments utilizing Kubernetes and Terraform to ensure high scalability and zero-downtime releases.

LDW Solutions / QueensLab – Senior Full-Stack Engineer (.NET Core / React) – Sweden (Remote)

Nov 2024 – Jun 2025

- Architected a greenfield autonomous store management and EV charging platform using .NET Core 8, React (TanStack), and an Event-Driven Architecture powered by Azure Service Bus, PostgreSQL, and Redis.
- Built a real-time RTSP video processing pipeline in .NET using FFmpeg for store monitoring, achieving 80% incident

detection accuracy with <6s latency.

- Integrated complex third-party hardware APIs (Loxone environmental controls, LanCom/Teltonika networking devices, and Eneo surveillance systems) with secure RESTful APIs.
- Enforced end-to-end quality gates by implementing automated testing suites using xUnit, Reqroll, and Playwright, while deploying containerized services via Azure DevOps CI/CD and Docker.

LDW Solutions / GFT / Benchmark Capital – Senior Full-Stack Engineer (.NET / React) – London, UK (Remote)

Jan 2023 – Nov 2024

- Co-led a 5-engineer agile team delivering a greenfield wealth risk management platform for Benchmark Capital using .NET Core 8 microservices, SQL Server, and React 18 / TypeScript.
- Implemented enterprise security controls (JWT, OBO Tokens, Okta, Azure APIM) to secure third-party integrations with Oxford Risk and Medallia APIs.
- Engineered frontend architectures using Redux Toolkit for state management, enforcing high code quality through automated unit testing with Jest and Vitest.
- Managed enterprise CI/CD pipelines utilizing TeamCity, Octopus Deploy, Docker, and Azure to ensure reliable, zero-downtime production deployments.

LDW Solutions / Capgemini Engineering & Various Clients – Senior Software Engineer / Tech Lead / Mobile Architect – Brazil / Portugal

Mar 2015 – Dec 2022

- Led cross-functional teams of up to 5 developers across enterprise, edtech, and SaaS sectors, accelerating client service performance by 30% and cutting manual customer success overhead by 90% via custom workflows built with React, Node.js, and Redis.
- Engineered scalable, offline-first mobile applications using React Native, Flutter, and Ionic, automating iOS/Android build pipelines via macOS Jenkins slaves.
- Spearheaded enterprise migrations from legacy .NET Framework 4.8 to modern .NET 6 microservices alongside React 18 and Angular 13 frontends.
- Designed high-throughput REST APIs, Kafka event streams, and real-time C++/Node.js Linux embedded interfaces deployed via Docker, Jenkins, and Azure DevOps.
- Enforced enterprise-grade security (Keycloak, Azure AD, PCI-DSS) and automated testing pipelines using Jest, xUnit, Karma/Jasmine, and SonarQube.

SKILLS

Key Technologies & Frameworks: LLM Integration (OpenAI API, Anthropic Claude, Gemini), Model Context Protocol (MCP), Custom AI Agents, RAG Pipelines, Vector DBs, LangChain, Cursor AI Workflows, Prompt Engineering.

Languages: C# (.NET Core / .NET 8), JavaScript / TypeScript (Node.js), Python (AI Integration & Automation), SQL, HTML5/CSS3.

Backend & Architecture: Microservices, Event-Driven Architecture, RESTful APIs, Clean Architecture, NestJS, Flask, FastAPI, [ASP.NET](#) Core.

Frontend & Mobile: React (18/19), Next.js, Vue.js, Angular, Tailwind CSS, Jotai, Zustand, React Native, Flutter, Ionic Framework.

Cloud & DevOps: Microsoft Azure (Functions, Service Bus, APIM), AWS, Docker, Kubernetes, Terraform, Azure DevOps, Jenkins, CI/CD Pipelines.

Databases & Caching: PostgreSQL, SQL Server, MongoDB, Redis, Oracle.

Testing & Quality: Vitest, Playwright, Jest, xUnit, Reqroll, SonarQube.

EDUCATION

Postgraduate Degree: Project Management | UniCesumar (2025–2026)

Postgraduate Degree: Marketing and Digital Media | UniCesumar (2025)

Technologist Degree: Systems Analysis and Development | Estácio (2012–2015)

Certifications: Microsoft Certified Professional (MCPS) — [2013], MS: Programming in HTML5 with JavaScript & CSS3 — [2013]

PROJECTS

L-Coder: Terminal Coding Agent System for Locally-Served Models (Python / GenAI)

Role: Creator & Architect | <https://apps.ldw.solutions/l-coder>

- Built a terminal-based autonomous coding agent system orchestrating locally-served LLMs (Ollama, LM Studio, oMLX) for automated code generation and review.
- Architected a RAG pipeline with local vector indexing to embed project-specific knowledge directly into retrieval workflows.
- Engineered modular agent tools for task orchestration, file manipulation, and automated command execution.
- Developed a web interface providing real-time monitoring of agent reasoning chains, execution logs, and live session approvals.

Whisper-Code: Hands-Free Voice Dictation Utility for macOS (Python / AI / MLX)

Role: Creator & Developer | <https://github.com/lekodeveloperweb/whisper-code>

- Built a hands-free voice dictation utility for macOS using Whisper (MLX-optimized for Apple Silicon) that transcribes speech directly into active applications at the cursor position.
- Engineered a "Ghost HUD" overlay and smart text injection pipeline, maintaining focus stability across apps while delivering real-time transcription.
- Integrated local API support (oMLX, OpenAI-compliant servers) ensuring 100% on-device, privacy-first audio processing with zero disk persistence.
- Designed a modern glassmorphic visual interface with system tray controls, custom shortcut triggers, and native performance tuning.