Carlos Contreras

Applied AI, Agents & Big Data FDE at AWS, since 2018 · Hands-on builder since 2006

I'm a Prototyping Architect at AWS, aka FDE (Forward Deployed Engineer) externally.

At AWS, I design and build AI agents, knowledge graphs and voice AI with enterprise teams, from the first workshop to a production-grade pilot.

- Location: Cambridge, United Kingdom
- Work rights: Spanish citizen · UK Indefinite Leave to Remain (ILR)
- Website: https://holacarlos.dev/
- LinkedIn (contact): https://www.linkedin.com/in/carloscontreras

Summary

Applied AI, agents and big data FDE at Amazon Web Services since 2018. I design and build AI agents, knowledge graphs, voice AI and data platforms with enterprise teams, from the first workshop to a production-grade pilot, and speak about it at SQLBits and AWS Summits.

Ways to Work Together

In my free time and in a personal capacity. I work full time at AWS, so availability is limited.

- On stage (Keynotes, Panels): Agents, voice AI and knowledge graphs, usually with a live demo.
- Hands-on (Half-day workshops): Your team builds a first working agent, side by side.
- Advisory (Architecture reviews, Pilot sprints): A second pair of eyes on an AI architecture, or a short sprint to a working pilot on your real data.

What I Work On

With customers, end to end: from shaping the idea with C-level leaders to building it side by side with their engineers, as one team.

Agentic AI

Multi-agent systems with Strands Agents and LangGraph on Amazon Bedrock AgentCore, with guardrails, memory and an evaluation set that gates every release.

Keywords: Multi-agent & subagents, MCP & RAG, Guardrails & evals

Voice AI

Real-time speech-to-speech agents over phone and web: Amazon Nova Sonic, ElevenLabs, and recent work with gpt-live-1.

Keywords: Voice agents on real phones, Speech-to-speech, Voice auth & barge-in

Knowledge graphs

Graphs and ontologies that give LLMs structured, explainable context, from engineering requirements to natural-language queries over the graph.

Keywords: Graph + LLM assistants, Ontologies, Text-to-SQL

Big data platforms

From Oracle Exadata warehouses to serverless data lakes and streaming on AWS: the data foundations the AI runs on.

Keywords: Streaming, SQL & NoSQL

AI-native engineering

Spec-driven development with coding agents (Claude Code, Kiro, Codex, Cline), agents that test the UI like a person (Playwright, AgentCore Browser Use), and fitness functions that keep it safe.

Keywords: Spec-driven coding agents, UI/UX testing via Browser Use, Infrastructure as code

My current favourite stack

- Coding agents: Claude Code, Codex, Kiro, Cline
- IDE: Orca
- Agents: Strands Agents, LangGraph, Amazon Bedrock AgentCore, MCP
- Voice: Amazon Nova Sonic, ElevenLabs, gpt-live-1, Twilio
- Offline, on the plane: Ollama + Qwen
- Build & test: AWS CDK, Playwright

Experience

Senior Generative AI & Big Data Prototype Architect, Amazon Web Services (AWS)

2018 – Present · London, UK · Prototyping and AI Cloud Engineering (PACE)

Prototype Architects work like forward-deployed engineers: embedded with customer teams, from executives to engineers, co-building solutions to undefined or evolving problems.

- Design and deliver multi-agent root-cause-analysis systems on Amazon Bedrock AgentCore and LangGraph, cutting incident diagnosis from hours to seconds.
- Build GenAI assistants for business teams that combine LLM agents with existing ML models, SQL databases, knowledge bases and live data.
- Knowledge graphs for vehicle-engineering companies: graphs that link product signals to engineering requirements, with GenAI assistants on top, and a reusable demo deployed end to end as code.
- Build real-time speech-to-speech voice agents on Amazon Nova Sonic, extending what the service offers out of the box.
- Run GenAI bootcamps and hands-on agentic workshops for enterprise customers and partners across Europe.

Senior Data Warehouse Engineer & Nagios Lead, WorldPay

2016 – 2018 · Cambridge, UK

Senior Data Tech Lead, Vix Technology (payments startup)

2015 – 2016 · Cambridge, UK

Data Warehouse Tech Lead & Oracle Exadata SME, Inditex (Zara, Bershka, Pull&Bear…)

2011 – 2015 · Global HQ, Spain

Database Architect, Tecnocom / Indra

2008 – 2011 · Spain

Consultant Database Engineer & Migration Specialist, General Electric

2006 – 2008 · US & Spain

Database Developer & DBA, Mexican Stock Exchange (Bolsa Mexicana de Valores)

2004 – 2006 · Mexico

Technology Path

- AI agents: since 2024
- Voice AI: since 2025
- Knowledge graphs & RAG: since 2021
- Generative AI: since 2023
- Big data & streaming: since 2018
- Cloud on AWS: since 2018
- Distributed data platforms, HA & DR: since 2008
- Data architecture & warehousing: since 2004

Speaking

- 2026, SQLBits, main auditorium: "AI Agent Orchestration in Action". Live root-cause-analysis demo with Amazon Bedrock AgentCore and LangGraph.
- 2026, AWS Summit Milan, AI Village: "From Alert to Action: The AI That Calls You Back". Agents diagnose a database failure, then phone the on-call engineer and fix it after voice authentication.
- AWS Summit: "Get quicker and more valuable insights from your data, without managing your own data warehouse". 900+ attendees
- 2026, SQL community webinar: "Automating SQL Server DB Ops & Health Checks with AI Agents on AWS"
- 2025, SQLBits: "RCA Agents · Natural Language to T-SQL: Minimising LLM Hallucinations"
- 2025, Internal AWS prototyping conference. Speaker and award winner.
- AWS Summits London and Madrid
- 2026, Executive briefings and partner events. Live agentic demos.

Alongside regular bootcamps, hands-on workshops, customer briefings and AWS Developers videos.

Recent Projects

- **Lumo360** (https://lumo360.io): A side project, built in my free time — voice agents that run first-round job interviews.

- **Agentic Platforms Guidelines**: An engineering playbook for shipping production software with AI coding agents. This site was built with it.
- **Voice RCA agents**: Agents find why a database failed, then phone the on-call engineer (Amazon Nova Sonic speech-to-speech, Twilio). Open source.

Teaching and Writing

- Course author: Generative AI with AWS Bedrock (Spanish), Udemy · bestseller (2025)
- Course author: Graph databases, real-time streaming, infrastructure as code, ProjectPro.io courses (2021–2024)
- Courses: On pause: not maintained at the moment. Available on request.
- How Formula 1 uses generative AI to accelerate race-day issue resolution, AWS Blog
- Hobbies: Formula 1 is a passion I often get to work on: with F1 itself, and onsite with two F1 teams.

Commit Activity

2381 commits between 2026-01-27 and 2026-11-12, across private and public repositories.

Industries

Motorsports; Fashion, retail & luxury; Banking, finance & fintech; Insurance & telecoms; Transport & logistics; Vehicle engineering & manufacturing; Real estate & proptech; Energy & utilities; Gaming & entertainment

Education

- MSc in Computer Science (2017–2018)
- BSc in Computer Information Systems (2000–2004)

Certifications

AWS Certified AI Practitioner, AWS; AWS Certified Security – Specialty, AWS; AWS Certified Data Analytics – Specialty, AWS; AWS Certified Solutions Architect – Associate, AWS; AWS Certified Cloud Practitioner, AWS; Earlier: Oracle Database Certified Professional, Oracle.

Languages

Spanish (Native); English (Full professional)

Last updated 2026-11-15.