Abner Evandro Duarte

Engenheiro de IA | Desenvolvedor Backend

São José do Rio Preto, SP, Brasil | (17) 99181-6083 | abnerduartw@gmail.com
LinkedIn: linkedin.com/in/abner-evandro-duarte-79352523a | GitHub: github.com/Abner-Duarte
Portfólio: abner-duarte.vercel.app | ZapSync: zapsync-ia.com.br
Disponível para trabalho remoto, híbrido ou presencial

RESUMO PROFISSIONAL

Engenheiro de IA que coloca LLMs em produção. Fundei e opero sozinho o ZapSync, SaaS multi-tenant de atendimento com IA no WhatsApp, com RAG híbrido, saída estruturada, guardrails, evals e benchmark de modelos, observabilidade de custo e latência e servidor MCP. Desenvolvo backend com Node.js, Python e PostgreSQL, com testes automatizados e CI no GitHub Actions. Antes disso, quase 3 anos em suporte N2/N3 e implantação de ERP em produção, com mais de 10 mil chamados resolvidos e foco em causa raiz, SQL e logs. Domínio de sistemas fiscais brasileiros (NF-e, NFC-e, SEFAZ) e TEF. AWS Certified AI Practitioner, Bootcamp Santander Dados com Python e IA e inglês avançado (C1).

EXPERIÊNCIA PROFISSIONAL

Fundador e Engenheiro de IA | ZapSync

Out/2025 até o momento | São José do Rio Preto, SP | zapsync-ia.com.br (demo pública em /demo)

SaaS multi-tenant de atendimento com IA para WhatsApp que eu projetei, desenvolvi e opero sozinho: a IA atende clientes de restaurantes, pizzarias, barbearias e clínicas, fecha pedidos, marca horários e passa para um humano quando precisa.

- Construí o pipeline da mensagem: webhook com validação HMAC, deduplicação, agrupamento de mensagens em sequência, fila por conversa, áudio com Whisper, imagem com modelo de visão, chamada ao LLM, guardrails e validação no banco.
- Criei uma cascata de 11 tentativas entre OpenAI e Groq (gpt-5.4-mini, gpt-oss-120b e gpt-oss-20b) com saída estruturada em JSON Schema estrito, para o cliente nunca ficar sem resposta.
- Implementei RAG híbrido do cardápio no PostgreSQL: busca por palavra-chave mais embeddings com similaridade de cosseno, reindexação só do que mudou e fallback para a busca por palavra se a API cair.
- Escrevi guardrails com auto-correção e regras no código (preço e quantidade recalculados pelo banco, pedido só fecha após o resumo), com cada disparo registrado.
- Montei um conjunto de evals com 58 turnos em 7 categorias (12 vindos de bugs reais) e avaliador determinístico, e um benchmark por modelo: 97% de acerto, 1,6 s de resposta típica e US$ 5,45 por 1.000 respostas. O benchmark revelou uma falha de prompt injection, que corrigi no prompt e no código: segurança foi de 89% para 100%.
- Implementei observabilidade de LLM (custo, tokens, latência, fallback e guardrails por chamada, via AsyncLocalStorage) exibida numa aba do painel do cliente.
- Desenvolvi um servidor MCP com 16 ferramentas para o dono operar o negócio conversando com um assistente de IA, reaproveitando autenticação, isolamento entre negócios e auditoria da API.
- Garanti isolamento multi-tenant, segurança (scrypt, sessões assinadas, rate limit, LGPD) e 194 testes automatizados contra PostgreSQL real no GitHub Actions. Onboarding self-service com Embedded Signup da Meta e cobrança recorrente no Mercado Pago. Hospedado no Railway e na Vercel.

Analista de Suporte Técnico N2/N3 e Especialista em Implantação | Consumer Soluções em Tecnologia

Fev/2024 a Set/2026 | São José do Rio Preto, SP

Referência técnica N2 e N3 em plataforma de gestão (ERP) para food service, responsável por manter restaurantes operando em produção.

- Resolvi mais de 10 mil chamados técnicos, com média acima de 400 por mês, do atendimento emergencial a casos críticos escalados.
- Diagnostiquei falhas complexas com análise de logs, consultas SQL diretas no banco Firebird e rastreamento da comunicação entre sistemas, chegando à causa raiz (RCA).
- Criei o D3code, ferramenta interna com IA adotada pelo time de suporte: transcreve áudios de clientes (Groq Whisper) e extrai dados de prints (Gemini), com Cloudflare Workers, Google OAuth restrito ao domínio da empresa e controle de uso por usuário.
- Reproduzi, documentei e reportei bugs ao time de desenvolvimento, validando as correções no ambiente do cliente após cada atualização.
- Atuei em integrações críticas: NF-e e NFC-e junto à SEFAZ, certificados digitais, TEF e pinpads, e integrações de delivery (iFood, Keeta, 99Food).
- Conduzi implantações completas (requisitos, instalação, configuração fiscal, migração de dados e go live), padronizei procedimentos e treinei novos colaboradores e equipes de clientes.

Técnico Autônomo (MEI)

Prestação de serviços sob demanda

- Instalação e configuração de equipamentos e infraestrutura de TI, incluindo instalação de TVs em escolas, com emissão de nota fiscal.

PROJETOS

Data Pilot (projeto do Bootcamp Santander/DIO): painel de vendas que conversa. O LLM escolhe qual função Python chamar (tool calling) e o Python calcula, então o modelo não inventa números. Testes garantem que o chat e o painel mostram o mesmo valor.

Python, Streamlit, Plotly, OpenRouter, pytest · github.com/Abner-Duarte/data-pilot

Detecção de fraude em cartão (projeto do Bootcamp Santander/DIO): 284 mil transações com 0,17% de fraude. Comparei modelos por precisão, recall e PR-AUC, ajustei o limiar e expliquei as decisões com SHAP; o XGBoost final pega 83 de 98 fraudes com 15 alarmes falsos.

Python, scikit-learn, XGBoost, SHAP, Jupyter · github.com/Abner-Duarte/deteccao-fraude-cartao

BeHappier: app Android com IA multimodal (chat, visão e voz) que projetei e construí sozinho, com cerca de 8.700 linhas em TypeScript, backend no Firebase e chave da API protegida em Cloud Functions.

React Native, Expo, TypeScript, Firebase, OpenAI · github.com/Abner-Duarte/behappier-showcase

Teste técnico de QA (Verzel): plano, casos em Gherkin, testes manuais e exploratórios e automação com Playwright (interface em Chrome, Firefox, Safari e mobile, mais testes de API) rodando no GitHub Actions. Encontrei 4 bugs, 2 de severidade alta.

Playwright, TypeScript, Gherkin, GitHub Actions · github.com/Abner-Duarte/Teste-qa-verzel

Prospecção de clientes do ZapSync: script que busca negócios locais por tipo e cidade na Google Places API e gera a lista de contatos para abordagem; a primeira rodada levantou 1.500 leads em São José do Rio Preto e Mirassol.

Python, Google Places API, CSV

Quest Contratado: jogo web que criei para gamificar minha busca de emprego e meus estudos, com XP, missões diárias, trilhas de Python (60 lições), SQL e AWS com revisão espaçada, roadmap de cloud com 112 lições, entrevista simulada com IA e herói 3D animado.

JavaScript, HTML, CSS, three.js, three-vrm

Reinos Eternos: idle RPG de fantasia em 3D no navegador, com 4 classes animadas, combate com armadura, esquiva e bloqueio, árvores de talento com sinergias, forja, 10 zonas e 24 monstros, e um simulador próprio para balancear a progressão.

JavaScript, Three.js, WebAudio, Node.js

TT2 Otimizadores: conjunto de ferramentas que roda localmente os otimizadores de um jogo mobile a partir do export da conta, com motor próprio em Node.js e Python, tabelas oficiais do jogo em CSV e execução de páginas web sem navegador.

Node.js, Python, jsdom, CSV, JSON

Userscript do Chatwoot (Consumer): extensão no navegador que escuta os eventos em tempo real do Chatwoot e avisa o agente quando entra uma conversa nova na fila, sem mexer no servidor.

JavaScript, Tampermonkey, WebSocket

Simuladores de QA (Consumer): simuladores interativos em HTML que montei para estudar para a vaga interna de QA, cobrindo Azure DevOps, Selenium, Cypress, Robot Framework, Git, SQL e Postman.

HTML, CSS, JavaScript

Jogos em Python: dois jogos em Pygame Zero com código comentado para quem está aprendendo, com toda a arte e todos os sons gerados por código, e três scripts de terminal (estoque, comissão e juros).

Python, Pygame Zero, Pillow · github.com/Abner-Duarte/portfolio-python

Desafio técnico em Python: três exercícios de regra de negócio (comissão de vendas a partir de JSON, movimentação de estoque com validação e juros por atraso) usando só a biblioteca padrão.

Python, json, datetime, uuid · github.com/Abner-Duarte/Desafios-Dev

Simulador de portão com Arduino (Fatec, 2026): projeto acadêmico de portão automático simulado com Arduino, documentado em blog.

Arduino, eletrônica

Bot de WhatsApp com GPT: meu primeiro bot de WhatsApp integrado a um modelo GPT, que serviu de base para o ZapSync.

WhatsApp, OpenAI GPT

Barbearia do Pra (Fatec, em equipe): sistema web full stack de agendamento com painel administrativo de agenda, profissionais, serviços e horários.

Java 17, Spring Boot, SQL Server, JavaScript · github.com/Abner-Duarte/Barbearia-Site

Fatec Fighters (Fatec, em equipe): jogo de luta 2D para dois jogadores com login, placar e ranking no PostgreSQL. Evoluiu de um protótipo numa classe única para telas separadas com camada DAO e consultas parametrizadas.

Java, libGDX, PostgreSQL, JDBC, Gradle · github.com/Abner-Duarte/FatecFighters

SysOs (projeto final do curso de front-end, 2023): site institucional e de vendas de um sistema de ordens de serviço e pedidos para pequenos negócios, com planos, login e cadastro e layout responsivo.

HTML, CSS, JavaScript, Bootstrap · github.com/Abner-Duarte/Projeto-Final

COMPETÊNCIAS TÉCNICAS

IA e LLMs: OpenAI, Groq, Gemini, Whisper, engenharia de prompt, saída estruturada (JSON Schema), tool calling, agentes, RAG híbrido, embeddings, busca vetorial, guardrails, evals e benchmark de modelos, observabilidade de LLM, prompt injection, MCP (Model Context Protocol)

Machine Learning e dados: pandas, scikit-learn, XGBoost, SHAP, métricas para classes desbalanceadas, Streamlit, Plotly, Jupyter

Backend: Node.js, Express, Python, Java, Spring Boot, APIs REST, webhooks, WhatsApp Cloud API, arquitetura multi-tenant, autenticação e segurança de APIs

Banco de dados: PostgreSQL, SQL Server, Firebird, Firestore, modelagem de dados, SQL de diagnóstico em produção, migração de dados

Front-end e mobile: JavaScript, TypeScript, HTML, CSS, React Native, Expo

QA e testes: Playwright, Gherkin (BDD), testes de API, testes funcionais, de regressão e exploratórios, node:test, pytest, reprodução e documentação de bugs, critérios de aceite

Cloud e DevOps: GitHub Actions (CI/CD), Railway, Vercel, Firebase, Cloudflare Workers, Git, AWS (AI Practitioner)

Suporte e negócio: suporte N2/N3, análise de causa raiz, implantação de ERP, go live, NF-e, NFC-e, SEFAZ, certificado digital, TEF, pinpads, PDV, food service, treinamento de usuários

FORMAÇÃO ACADÊMICA

Tecnólogo em Informática para Negócios · Faculdade de Tecnologia do Estado de São Paulo (FATEC), São José do Rio Preto · Concluído em 2026
TCC sobre mecanismos de desligamento automático em IA.

CERTIFICAÇÕES E CURSOS

- **AWS Certified AI Practitioner (AIF-C01)**, Amazon Web Services
- **Bootcamp Santander Dados com Python e IA**, DIO, concluído em nov/2026: Python e POO, testes e boas práticas, tratamento, visualização e automação de dados, bancos de dados relacionais, machine learning, NLP, visão computacional, IA generativa, engenharia de prompt e agentes de IA
- **Bootcamp IBM Confluent: Dados em Tempo Real para Agentes de IA**, DIO
- EF SET English Certificate: 69/100 (C1 Advanced)
- NodeJS: The Complete Guide (incl. MVC, REST APIs, GraphQL), Udemy (Maximilian Schwarzmüller)
- 100 Days of Code: The Complete Python Pro Bootcamp, Udemy (Angela Yu)
- SQL and PostgreSQL: The Complete Developer's Guide, Udemy (Stephen Grider)
- Playwright JS/TS Automation Testing from Scratch & Framework (Cucumber BDD, CI/CD com GitHub Actions), Udemy (Rahul Shetty)
- The AI Engineer Course: Complete AI Engineer Bootcamp, Udemy (365 Careers)

IDIOMAS

Português: nativo · Inglês: avançado (C1)