

João Rietra

Médico e AI Engineer · IA aplicada a saúde, validação clínica e dados sensíveis

Recife, PE · (81) 99243-3485

joaorietra@gmail.com · jhlr.github.io · github.com/jhlr · linkedin.com/in/joaorietra

CRM-PE 32980 · Medicina, FPS/IMIP (2022)

RESUMO

Médico com três anos de prática clínica — urgência pediátrica e atenção primária — que migrou para engenharia de IA sem abrir mão do lado clínico. Hoje construo e avalio modelos de aprendizado de máquina e aplicações de IA generativa, com mestrado profissional em Engenharia de Software (CESAR School) cuja dissertação trata exatamente de **confiabilidade de modelos em imagem médica multidomínio**: calibração, generalização fora-de-distribuição e fidelidade de explicação. Sou a interface rara entre os dois mundos — modelo o domínio, converso com o especialista na mesma língua, sei o que uma métrica significa à beira do leito e trato dado sensível com o cuidado que ele exige.

O que me diferencia de um engenheiro de ML com bom currículo: eu sei por que sensibilidade importa mais que acurácia numa triagem, por que um modelo mal calibrado é perigoso mesmo com AUC alta, e por que um mapa de saliência bonito não é evidência de nada. Não é vocabulário aprendido — é a régua com que fui treinado a decidir.

EXPERIÊNCIA

Médico — Urgência Pediátrica e Atenção Primária · 2022 – 2025

- Decisão diagnóstica sob alta demanda em pronto-socorro pediátrico; cuidado longitudinal e coordenação em atenção primária.
- Prática de triagem, estratificação de risco e comunicação com famílias — a mesma estrutura de decisão sob incerteza que hoje modelo em software.
- Leitura crítica de literatura e consensos como base de conduta, hoje reaproveitada na validação de sistemas de apoio à decisão.

AI Engineer — LetzGrow (Grow Group) · 2026 – atual

Plataforma de jornada de desenvolvimento de pessoas assistida por IA.

- **Agente conversacional sobre LLM** que conduz a jornada, interpreta respostas abertas e gera recomendações personalizadas; geração de plano de desenvolvimento com decomposição em micro-tarefas adaptativas.
- **RAG** para ancorar as recomendações no perfil do usuário e no domínio, reduzindo alucinação — o mesmo problema de *groundedness* que qualquer aplicação clínica de LLM enfrenta.
- **Dados sensíveis por design**: acesso condicional, trilha de auditoria e conformidade com a LGPD; avaliação de fairness do pipeline de recomendação, com relatório versionado.
- Backend NestJS + Prisma sobre PostgreSQL, LLM via Groq (Llama 3.3 70B), embeddings locais com Ollama, infraestrutura AWS com CDK.

Engenheiro de Software (SMO + QA) — GasoGraph (Mestrado / CESAR School) · 2025 – 2026

Sistema de apoio à decisão clínica sobre árvore de regras derivada de consensos publicados.

- **Tradução de consenso publicado em pipeline de decisão auditável** — a parte que costuma travar esse tipo de projeto, e que eu resolvo por ser dos dois lados.
- Validação rigorosa de entrada, tratamento de inconsistências e cobertura de teste nas regras críticas.
- Modelagem do domínio e validação das regras com background próprio, encurtando drasticamente o ciclo de iteração com especialistas.

ML Engineer — Veritas (School Innovation, CESAR School) · 2025 – 2026

Detecção de mídia manipulada com deep learning.

- Pipeline de visão computacional com Grad-CAM, análise sistemática de falsos positivos e avaliação de generalização fora-de-distribuição — o mesmo protocolo que aplico a imagem médica.

- Composição de evidência do modelo (score forense + heatmap) com explicação em linguagem natural por IA generativa.

Pesquisador — Neurociência Computacional, University of Alberta · 2015 – 2016

- Intercâmbio Ciências sem Fronteiras: modelagem de fluxos neurais e análise de datasets de alta dimensionalidade, no cruzamento entre Computação, Neurociência e Genética.
- Bioinformática aplicada: identificação de indels em arquivos BAM (indelMINER).

Monitor — CESAR School · 2025 – 2026

- Monitorias em Programação (C), Estatística (R), Cálculo I e Programação Competitiva.

PESQUISA — CONFIABILIDADE DE MODELOS EM IMAGEM MÉDICA

Dissertação de Mestrado, CESAR School (2026). Dois artigos conectados sobre **quando um modelo é confiável o bastante para chegar perto de uma decisão clínica**.

1. Reuso de treino para triagem binária em imagem médica multidomínio.

Comparação controlada de 4 origens de backbone sobre 5 domínios heterogêneos (MedMNIST), avaliando AUC, sensibilidade/especificidade, **calibração (ECE)** e **fidelidade de explicação por oclusão**. O fine-tuning intermediário forense venceu de forma consistente (AUC **0,842** vs 0,796 do ImageNet) com calibração muito superior (ECE **0,045** vs 0,219) — ou seja, além de acertar mais, a confiança declarada pelo modelo correspondia à realidade, que é o que permite usar um limiar com segurança.

Resultado negativo reportado: o Grad-CAM produzia mapas espacialmente concentrados e convincentes, com **fidelidade nula à decisão real do modelo**. Explicação bonita não é explicação.

2. Detecção de imagens manipuladas com redes convolucionais leves.

Avaliação empírica e adversarial de detectores próprios vs. externos (ViT, Swin) e ensembles, em 3 distribuições, com foco em OOD. Nenhum detector único generaliza entre tipos de manipulação; "leve em parâmetros" ≠ "leve na prática" — a métrica de peso relevante é latência. Proposta de arquitetura de fusão *gated* por sinal de domínio.

Infraestrutura experimental. MLflow com backend SQLite: métricas por época, matrizes de confusão como artefato, logging de inferência com deduplicação content-addressed (sha256) e agregação estatística (média, intervalo de confiança, Wilcoxon) sobre múltiplas seeds. Reprodutibilidade ponta a ponta do treino à avaliação.

PROJETOS

pedrita — código da dissertação · github.com/jhlr/pedrita

Transfer learning multidomínio sobre MedMNIST, avaliação com AUC/ECE/oclusão, infraestrutura MLflow e backend FastAPI servindo o modelo.

GlucoTrack · github.com/jhlr/GlucoTrack

Aplicativo de controle de diabetes: registro de glicemia e calorias.

indelMINER · github.com/jhlr/indelMINER

Bioinformática: identificação de indels em arquivos BAM.

klone-persona-rag — RAG com banco vetorial (Sentence-Transformers + ChromaDB), recuperação por similaridade com priorização de recência. Base de *groundedness* para aplicações de LLM.

Minnow · github.com/jhlr/minnow

Agente de coding que roda inteiro dentro do navegador do celular — sem conta e sem backend, com o workspace nunca saindo do dispositivo. Relevante para saúde pelo argumento de arquitetura: privacidade obtida por onde o dado roda, não por promessa contratual.

FORMAÇÃO

- **Mestrado Profissional em Engenharia de Software** — CESAR School, Recife · Out/2025 – em andamento
- **Tecnólogo em Ciência de Dados** — CESAR School, Recife · Jan/2025 – em andamento

- **Medicina** — FPS / IMIP, Recife · Concluído em 2022
- **Ciências sem Fronteiras (Neuroscience)** — University of Alberta, Canadá · 2015 – 2016
- **Ciência da Computação** — CIn / UFPE · Jan/2014 – interrompido

Reconhecimento: primeiro aluno classificado com dois projetos simultâneos no School Innovation da CESAR School (2026.2) — trilha de pré-incubação com seleção competitiva e orientação docente. Dos oito grupos classificados na edição, dois são meus.

COMPETÊNCIAS

Saúde e IA: modelagem de domínio clínico · validação com especialistas · tradução de consenso em regra de decisão · imagem médica (MedMNIST) · métricas clinicamente relevantes (sensibilidade, especificidade, calibração) · LGPD e dados sensíveis · leitura crítica de literatura

Machine learning: PyTorch · scikit-learn · Hugging Face Transformers · OpenCV · transfer learning e fine-tuning · calibração (ECE) · avaliação OOD · explicabilidade (Grad-CAM, oclusão) · estatística inferencial (IC, Wilcoxon, múltiplas seeds)

IA generativa: LLMs (OpenAI, Anthropic, Groq, Ollama) · RAG e bancos vetoriais · fluxos agênticos · avaliação de groundedness e alucinação

Engenharia: Python (FastAPI) · TypeScript (NestJS, React) · PostgreSQL · MLflow · Docker · AWS · Git · testes e CI · R, SQL

Idiomas: Português (nativo) · Inglês (nativo) · Espanhol (básico) · Francês (básico)