RICARDO JOSÉ MENEZES MAIA

Phone: +55 61 982068043
E-mail: ricardo.jmm@gmail.com
LinkedIn: linkedin.com/in/ricardojmm
Brasília – Brazil

Senior Data Science Architect

TWO DECADES DELIVERING HIGH IMPACT SOFTWARE SOLUTIONS

Seasoned, results-driven, and highly motivated senior software engineer with a solid record of delivering high-impact, state-of-the-art data science and machine learning solutions. Passionate about learning, challenging projects, and always striving to exceed expectations and achieve established goals.
Currently, part of the ChatGPT Project team at OpenAI as a Data Scientist, working on the accuracy enhancement for the python code generation model, and former software architect on the Brazilian Government's project called Electronic Judicial Process, the most impactful digital transformation initiative in the history of Justice in Brazil, adopted nationally as a unified solution for all Brazilian Courts.
Fluent in several programming languages and cloud platforms, and concluding a Ph.D. degree in Computer Science, focusing on guaranteeing privacy-preserving in machine learning models, having earned several prizes on correlated published papers and international competitions.

CORE SKILLS

- Software Development
- Data Science
- Machine Learning
- Cloud Computing
- Privacy-preserving
- Cryptography
- Detail oriented
- Problem Solver
- Team work

PROFESSIONAL EXPERIENCE

OpenAI, USA **08/2022 – Present**

- **Senior Python developer -** Team member of the ChatGPT project (https://openai.com/blog/chatgpt/), developing strategies and python components to improve the accuracy of the ChatGPT model.

Superior Labor Court, Brasília-DF, Brazil **03/2017 – Present**

- **Data Scientist and Data Engineer –** Participated in operationalizing a machine learning solution called Bem-Te-Vi and Sabiá, whose focus is to support the court servers to find similar processes and thus allow the same decision to be applied to many processes. This tool has won many awards in the Brazilian judiciary. Technologies: Language Model with 18 million legal documents: Word2Vec; Supervised Model: XGBoost; Unsupervised Model: KMeans, DBScan, Agglomerative hierarchical Clustering; Framework Big Data: Spark; Framework Automl: H2O; Frontend: React; ML Microservice: Python, Sklearn; Business Rule Microservice: Java 11; Data Extraction: Pentaho; Source Data: Soap API, Rest API, Postgres, Oracle; Deploy Models and Application: Jenkins, Docker, Kubernetes, Microservices Architecture.
- **Software Architect –** Main Java Architect of the Electronic Judicial Process project of the Court **(3.5k internal users, 200k external users, 1 M Cases/year)**. Designing and building software integration components with legacy systems and the new Electronic Court Process via Microservices. Technologies: JPA, JBoss, Angular, React, Spring Boot, Java 8, Docker, Kubernetes, PostgreSQL, Oracle, Elastic Search, Microservices, Maven, Jenkins, Jira, GitLab, Microservices Architecture.

Superior Electoral Court, Brasília – DF, Brazil **04/2016 – 03/2021**

- **Software Security Specialist** – Responsible for securing the code that counted votes in Brazil's elections. Brazil has **147.9 million voters**, and this software took care of elections for president, deputy, senator, mayor, and governor. It is one of the most delicate software, where Brazil elections every two years, and due to political criticality, it requires constant security improvements. Relevant to **code security** improvements. Technologies: Java, JPA, JBoss, Angular, Spring Boot, Oracle, Microservices, Maven, Jenkins, Redmine, GitLab, Sonar, libsodium library, C, C++, Sonar, Microservices Architecture. I created libraries to avoid changing data inserted in the database, manipulation of software in production and automations to evaluate security and code quality.

Labor Justice Superior Council, Brasília – DF, Brazil. **09/2010 – 03/2017**

- **Software Architect -** Leader technical the first Electronic Process System Program, the most critical digital transformation initiative in the history of Labor Justice in Brazil, adopted nationally as a unified solution for all 25 Brazilian Labor Courts (current figures: **4k** Judges, **43k** Servants, **500k** Lawyers, and over **40M** Electronic Cases). Technologies: Java 8, Spring Boot, JPA, JBoss, Angular, Docker, Kubernetes, PostgreSQL, Oracle, Elastic Search, Maven, Jenkins, Jira, GitLab, Microservices Architecture.
- I participated in this project, acting in technical leadership and **performance improvements**, as the system was not designed to be used nationally. Actions were monitoring bottlenecks with AppDynamics, switching to a microservices architecture, optimizing SQL queries, and creating a tool to automate code analysis via GitLab merge request. Relevant

to **code security** improvements using Sonar.

Regional Labor Court of Amazonas, Manaus – AM, Brazil. **12/2008 – 09/2010**

- **Software Architect -** Led the team and architected and implemented solutions for the administrative and legal areas of the court. This work was essential after an incident where the court lost all its software. We needed to rebuild a portal and an application for processing administrative proceedings in record time. Technologies: Java 5, JPA, JBoss, JSF, PostgreSQL, Oracle, Maven, SVN, Monolithic Architecture.

EDUCATIONAL BACKGROUND

Ph.D. degree, Computer Science (University of Brasília – UNB – Brazil) **2019 – Expected March 2023**

- **Thesis theme: Privacy-Preserving Machine Learning**
- Developing new methods (Secure Multi-Party Computation, Differential Privacy, Federated Learning) that allow training and inference using machine learning and deep learning algorithms while maintaining data privacy, which is essential for us to have the benefits of artificial intelligence and maintain the privacy of users' data. I'm creating an application to use privacy-preserving machine learning to resolve the problems related to Collaborative Intrusion Detection System, Malware, Phishing, Covid-19, and Recommendation Systems.
- Working on a project to classify Domain Generation Algorithm (DGA) using privacy-preserving machine learning with LSTM, MLP, and CNN-1D models. Training with differential privacy using DP-SGD and Inference with MPC.
- Prototype to detect malwares: http://www.detectormalware.com . Backend(Python, Flask), Machine Learning Models (Tensorflow,Keras, Pytorch, Sklearn), AWS(EC2 instances for train and inference), GCP(AppEngine, Datastore to use NoSQL, Memcache, VM instances), Frontend(React)

Master's degree, Computer Science (University of São Paulo – USP – Brazil) **2008 – 2010**

- Dissertation: Analysis of the feasibility of implementing post-quantum algorithms based on quadratic multivariate quasi-groups in limited processing platforms
- https://teses.usp.br/teses/disponiveis/3/3141/tde-30112010-151111/pt-br.php

Bachelor's degree, Computer Science (University of Amazonas, Brazil) **1999–2003**

ACADEMIC ACHIEVEMENTS AND PUBLICATIONS

- **Second place in the U.S. PETs Prize Challenge: Phase 2 (Financial Crime–Federated), Team PPMLHuskies** https://www.drivendata.org/competitions/105/nist-federated-learning-2-financial-crime-federated/leaderboard/
- **First Place in IDASH, Team PPMLRobots (University of Washington)** – PRIVACY & SECURITY WORKSHOP 2021 – Secure Genome Analysis Competition - Track III. http://www.humangenomeprivacy.org/2021/, USA, 2021.
- **8th position**(accuracy) and **7th position**(explainability) with my Team PPMLHuskies in **Hackathon about Artificial Intelligence for Covid-19 prognosis: aiming at accuracy and explainability** https://ai4covid-hackathon.it/submissions_results/ February 2022.
- Sikha Pentyala; Davis Railsback; Ricardo Maia; Rafael Dowsley; David Melanson; Anderson Nascimento; Martine De Cock. **Training Differentially Private Models with Secure Multiparty Computation**. Computer Science -> Cryptography and Security. 2022. https://arxiv.org/abs/2202.02625**.**
- David Melanson; Ricardo Maia; Hee-Seok Kim; Martine De Cock; Anderson Nascimento. **"Secure Multi-Party Computation for Personalized Human Activity Recognition".** 2022. Neural Processing Letters – Springer Journals. https://www.springer.com/journal/11063
- Liriam Enamoto; Andre Santos; Ricardo Maia; Li Weigang; Geraldo Pereira Rocha Filho. **Multi-Label Legal Text Classification with BiLSTM and Attention.** International Journal of Computer Applications in Technology (IJCAT). 2021. https://www.inderscience.com/
- MAIA, Ricardo José Menezes; BARRETO, Paulo Sérgio Licciardi Messeder; OLIVEIRA, Bruno Trevizan. **Implementation of Multivariate Quadratic Quasigroup for Wireless Sensor Network.** Springer Transactions on Computational Science (Print), v.6480, p.64 - 78, 2010. https://doi.org/10.1007/978-3-642-17697-5_4