Sydney, Australia
jaeyow@gmail.com
0402226022
LinkedIn

Jose 'JO' Reyes

Github
Fullstack ML
Fullstack Developer Tips

Experienced software engineer specialising in **Machine Learning** and **Generative AI**, with expertise in building full-stack, end-to-end **machine learning systems**. Backed by over a decade of cloud-native development experience, I have a proven track record across diverse industries—including digital twins, industrial automation, and e-commerce particularly on AWS. Skilled at leading cross-functional teams and delivering comprehensive systems including machine learning applications supported by robust **MLOps** pipelines.

Work Experience

AI/ML Technical Specialist   **Cevo Australia**   **Jan 2023–Present**

- Cevo Australia is an AWS specific cloud computing consultancy that specialises in cloud adoption, digital transformation, modernisation, and data & AI.
- As an AI/ML Consultant, I am responsible for the design and build of AI/ML projects in the AWS ecosystem, with noteworthy projects such as:

Productionisation of a Generative AI Project
A large Australian company in the mergers and acquisitions (M&A) space required assistance to transition their data scientists' code to a production-ready state. This involved porting their code to a more robust and asynchronous format, enabling concurrent large language model (LLM) calls rather than sequential calls. Consequently, the client application received insights within a shorter timeframe, significantly improving the application's reliability and user experience.

RAG-based Internal Project
An internal project that allows internal Cevo users to upload past Statements of Work (SOWs), which are proposal documents. This enables the sales team to review past proposals, study them, ask questions, and identify patterns for creating new SOWs. The project utilises an AWS service called Amazon Bedrock Knowledge Bases, a generative AI managed service that facilitates building Generative AI systems without developing them from scratch, resulting in faster development, and reduced time and cost.

Intelligent Document Processing Project
A leading strata management Australian company required assistance in building intelligent document processing systems for files received via their email system. I led a team that designed, built, and delivered the proof of concept (POC) within the AWS ecosystem. The POC enabled automatic classification, data extraction, and downstream handling, saving time and allowing the client to focus on higher-value tasks, while also allowing them to scale to acquire more business.

Image Similarity Search Engine
A large Australian manufacturer of timber mouldings required assistance in building a system to expedite the search of their database for similar and matching designs. We designed, built, and delivered an image similarity search system, a custom machine learning project based on ResNet and vector databases. This involved using PyTorch to fine-tune a deep learning image model complete with end-to-end MLOps. The new system reduced manual search time from several hours to mere seconds.

Senior Software Engineer   **Willow - Mining Digital Twin**   **Nov 2021–Dec 2022**

- Willow is an Australian software developer of Willow Digital Twin - a platform that enables the management and operation of buildings and infrastructure.

- As an individual contributor, I have developed full-stack web applications for a Digital Twin platform for the mining industry. This involved intuitive and responsive user interfaces, as well as robust backends to support the platform's complex data visualisations.This enabled faster, data-driven decisions.

- As a part-time member of the data engineering team, I built and maintained data pipelines that collect, transform and deliver data required from mining operations. IOT related technologies such as Azure Data Explorer, Azure Data Factory and Azure IOT Hub were used. These activities enabled the transformation of raw, real-time data into actionable insights for improved operational efficiency, safety, and decision-making.

Technical Lead   **Inchcape Australia - Subaru**   **Feb 2019–Nov 2021**

- Inchcape is a multinational automotive distribution, retail and services company, and in Australia, distributes Subaru, a brand well known for their all wheel drive vehicles.
- I was responsible for the direction of technical strategy and development for Subaru's e-commerce Content Management System (CMS) on AWS.
- Team leader to a distributed international team, performing architecture, development and operations, ensuring a resilient and high-performance web experience.

Senior Consultant   **Readify (Telstra)**   **Oct 2016–Feb 2019**

- Readify was an Australian IT consultancy known for its expertise in software development and cloud solutions, particularly within the Microsoft/Azure ecosystem.
- I specialised in building enterprise web/mobile solutions, focusing on full-stack and cross-platform development on Azure.
- I worked closely with clients to tailor web application solutions, ensuring alignment with business goals.

Education and Certifications

- **Graduate Certificate of Data Science,** University of New South Wales   **Jan 2023-Jan 2024**
- **AWS Machine Learning Specialty**, AWS   **Sep 2022**
- **Deep Learning Specialty (4-Months)**, DeepLearning.AI, Coursera   **Jun 2022**
- **Data Science Bootcamp (3-Months)**, General Assembly, Sydney   **Jan 2022**
- **BSc Computer Science,** University of San Carlos, Philippines

Technologies and Languages

- Cloud - AWS, Azure, Modal, Baseten
- Data Science/Machine Learning - Python, Statistics, AWS SageMaker, Metaflow, TensorFlow, PyTorch, Jupyter, MLFlow
- AWS AI/ML - SageMaker, Personalize, Rekognition, Comprehend, Textract, pgVector, LanceDB
- Generative AI - RAG Architecture on AWS, LlamaIndex, Langchain, Haystack, AWS Bedrock, Vector Databases, Hugging Face
- Data - SQL, DynamoDB, Azure DataFactory, Snowflake, Azure Data Explorer, Kusto, AWS Glue
- Frontend/SPA - ReactJS, TypeScript, NextJS
- Backend - .Net Core, Python, TypeScript, Docker, FastAPI

Personal Projects and Blog Articles

Community Involvement

- ***Retrieval Augmented Generation Tutorial Sessions for Code.Sydney***
  Delivered a 7-part series introducing Generative AI and RAG techniques, covering LLMs, vector databases, and advanced RAG concepts. Youtube playlist   Github Repo

AI/ML Related Projects

- *Modelling with Metaflow and MLFlow*
  Using modern MLOps toolset, this project demonstrates using Metaflow to build an end-to-end pipeline on AWS, and MLFlow for machine learning experimentation, enabling the predictable orchestration and repeatability of experiments. Technologies: Metaflow, MLFlow, PyTorch, Model registry, Hyperparameter optimisation, Model checkpointing.

Productivity Improvement

- *Democratise Database Understanding using Generative AI*
  Article on using Generative AI to democratise database understanding through Text2SQL

- *Generative AI: Building systems, not just chatbots*
  Article on using Generative AI to build software systems and not just limited to chatbot-like applications

Recommenders on the AWS ecosystem

- *How to Build, Train and Deploy Your Own Recommender System – Part 2*
  Deploy the best model to serverless (Lambda) using SAM and Lambda container images
  Source code for Frontend and MLOps

- *How to Build, Train and Deploy Your Own Recommender System – Part 1*
  An implicit feedback (ALS) recommender system on AWS using Metaflow and Comet for MLOps - Source code

- *Build Recommender Systems the Easy Way in AWS*
  A production ready recommender system using Amazon Personalize
  Source code

Image Recognition System on AWS

- *Accelerate ML Application Development in AWS*
  An image classifier using Amazon Rekognition and AWS Amplify

- *Going to Production with Github Actions, Metaflow and AWS SageMaker*
  An image classifier using modern MLOps toolset - Source code

Formula 1 Race Prediction System

- *Small to Reasonable Scale MLOps*
  An MLOps solution (Metaflow) for reasonable scale systems when you're not a giant like Google - Source code

- *Data Science Bootcamp - MLOps on the cheap!*
  A race prediction system complete with MLOps using GitHub Actions and Python - Source code