Aiman Najjar | AIMANNAJJAR.COM | LINKEDIN GITHUB | **AI ENGINEER**

Introduction

I am a Platform Engineer with a strong focus on AI, GenAI, MLOps and technologies such as HuggingFace Transformers, LangChain (including LangGraph for Agentic AI), Apache Airflow, AWS SageMaker, Vector Databases (Weaviate) and Semantic Search. I have experience in both developing software and building infrastructure using modern practices.

I bring a breadth of experience spanning software, platform, and AI engineering. As a **Software Engineer**, I developed scalable APIs and backend systems for companies such as John Wiley & Sons, NBCUniversal, Rakuten Marketing, and Electronic Arts. In my **Platform Engineer** roles at Pythian, Couchbase, and VMware, I automated cloud infrastructure, built CI/CD pipelines, and improved system reliability and performance. More recently, I've focused on **AI engineering**, building data pipelines with Airflow, fine-tuning LLMs with Transformers, and developing custom AI agents with LangChain and LangGraph.

Selected Project

TripGPT.io - AI-Powered Travel Planning Agent

Developed a custom AI agent that leverages Large Language Models (LLMs) with Retrieval-Augmented Generation (RAG) and a custom-built LangGraph-based agent (written in Python) with advanced technology such as short-term memory, LLM response streaming, RAG nodes and others.

Technology Stack:

- **AI/ML:** LangChain framework, LangGraph, HuggingFace Transformers, SageMaker / Jupyter, vLLM
- **Data Infrastructure:** Airflow, Weaviate vector database for semantic search, RAG implementation
- **Backend (API):** Python, Redis, SQLAlchemy, FastAPI, SSE streaming
- **Compute Infrastructure:** Containerized environment, nginx, Docker

Professional Experience

As Platform / DevOps Engineer

Senior DevOps Engineer, Propel Inc
April 2023 - April 2024

Propel builds very high-traffic mobile applications (5M active monthly users) for Food Stamps benefit recipients to access their benefits using machine learning applied on automated phone calls and other information retrieval and processing technologies

- Started as a consultant when the then-early stage startup started experiencing some growth pains in 2017 to help the company adapt then as a full-time engineer in 2023 to help it grow the DevOps team
- Built and migrated the legacy compute infrastructure to a highly scalable infrastructure for the startup and helped the organization adopt new and modern DevOps practices

- Introduced IaC, Pipeline as Code and GitOps to the organization using tools such as Terraform, Terragrunt, Jenkinsfiles/JCasC and Atlantis
- Introduced Kubernetes to the organization with strong ecosystem consisting of EKS, external-dns, certificate-manager and Istio service mesh along with strong CD pipelines
- The infrastructure supported well-over 7000 operations per second (OPS) with demonstrable autoscaling success when the traffic surged significantly and unexpectedly as various states introduced various COVID-19-era benefits
- Built repeatable and auditable infrastructure change process driven by GitOps using Atlantis for Terraform
- **Technologies include:** AWS, Ansible, Kubernetes, Helm (custom charts and 3d-party), Python, Ansible, Terraform, Jenkins (containerized-workers), RDS (Postgres), Elasticsearch/Fluent/Kibana, Istio, Prometheus/Grafana, Datadog

Senior DevOps Engineer, VMware
September 2021 - April 2023

Worked in a large team of DevOps engineers responsible for creating and supporting infrastructure of VMware Workspace One (EUC) enterprise cloud products

- Worked on developing automation processes for various infrastructure processes such as OS and DB patching, servers rip and replace, automated datacenter migrations,..etc
- Development the automated processes involved working on Ansible playbooks and custom Ansible modules (Python), Powershell, Bash/Python scripting and interacting with large numbers of technologies such as SQL servers, Windows networks, F5 load-balancers,..etc
- Took the lead in on-call rotations responding to critical incidents when they occur whose resolution involved working directly on both Linux and Windows systems leveraging knowledge in low-level networking, TCIP/IP, F5 load balancers, bash scripting, Powershell and low-level OS debugging skills
- Took the lead on new initiative to migrate the organization Jenkins bare-metal infrastructure to Kubernetes-based one with containerized workers
- **Technologies include**: Bare-metal, VMware virtualization, Ansible, Powershell, Bash, Python, Jenkins, JCasC

Independent Contractor
October 2015 - Present (on and off)

- Contracted by a TravelClick hospitality firm to develop Chef cookbooks and Ansible playbooks for their large data-center-hosted infrastructure.

- Contracted by Northwestern University in the Chicago area to plan and execute the migration of legacy .NET web services into containerized Java/Spring microservices in the AWS Cloud (ECS).

- Assisted and mentored data scientists and developers in refactoring Spark jobs into testable code by removing environmental dependencies and replacing them with mocks for HDFS and other external dependencies.

- Worked closely and provided operational support for engineers using Hadoop clusters, including ensuring accessibility

and security, troubleshooting system and applications issues, installing and configuring new services, and other operational support responsibilities.

- Setup Jenkins CI/CD pipelines for complex Big Data applications, such as Storm topologies and Spark jobs.

Principal Site Reliability Engineer, Couchbase
Dec 2019 - June 2020

- Designed plan for automated testing of the DBaaS platform on mock AWS environments based on Localstack, k3s, and a Kubernetes controller I specifically developed to mimic behavior of ALB ingress controllers and PVC ELB provisioner locally. Effectively allowing the product to be tested ad-hoc in build pipelines without the need for real target AWS environments.

- Designed plan for Overall Reliability Enhancements, including introducing GitFlow strategy, BDD testing and Ansible for automated runbooks.

Engineer V, Cloud, Research & Development , Cray Inc
April 2019 - August 2019

- Engineered operational process enhancements for offline deployments, streamlining procedures and improving deployment accuracy
- Engineered Kubernetes automation processes such as writing advanced scripts to deploy Helm charts and setup new bare-metal supercomputer clusters.
- Engineered novel automation solutions at Cray to generate offline media containing Helm charts and Linux applications running in containers, enabling the deployment and upgrades of supercomputer clusters in an air-gapped setting.

DevOps Principal Architect, Pythian
July 2017 - March 2019

As a DevOps architect with Pythian, I was responsible for:

- Conducting thorough information discovery with clients to understand problems and paint points.
- Identifying targeted as well as end-to-end automation opportunities or working with clients to capture their desired end-state in structured and actionable format.
- Making architectural recommendations, including short-term and long-term roadmap and technology selections.
- Creating backlogs of tasks to achieve desired end-state.

As Software Engineer

Senior Software & DevOps Engineer, Rakuten Marketing
September 2013 – November 2015

Started as Backend Engineer and transitioned into a more DevOps-focused role in 2015.

DevOps Engineering:

- Wrote Chef cookbooks and Python/Shell scripts to fully automate the provisioning of a complex, interdependent 100 terabyte CDH cluster
- Built custom CSDs and Cloudera Parcels for Redis, Storm and Pentaho BI server (opensource available at http://github.com/linkshare)
- Wrote deployment scripts to deploy Oozie workflows and HDFS scripts using Jenkins
- Automated deployment of Oozie workflows and HDFS scripts
- Automated SOPs to manage/operate Storm streams using Jenkins
- Worked on Lambda Architecture, with focus on real-time layer and streaming
- Wrote Chef cookbooks to provision various application servers both on AWS and data centers, leveraging advanced concepts such as encrypted data bags, Chef Search API, LWRPs, Berkshelf, etc.

Back-End/Full Stack Development:

- Developed strictly compliant RESTful API using PHP/Symfony
- Developed Node.JS real-time data streaming application, capable of streaming up to 5k records/second
- Worked with MongoDB, including concepts such as sharding and replica sets

Big Data / Java Engineering

- Wrote performant Storm topology to stream high-volume data (roughly 7 - 10k rows per second) from Oracle to HDFS in real time (some portions are open source now, available at http://github.com/linkshare)

Software Engineer Intern, Electronic Arts Inc *June 2013 – September 2013*

Software Engineer - Manager, Software Development, NBCUniversal *April 2010 – December 2011*

Web Development Consultant, John Wiley & Sons, Inc *November 2008 – April 2010*

Freelance Web Developer *2005 – November 2008*

Please see my customers reviews at: https://www.freelancer.com/u/aimannajjar.html

Technical Skills

- **Software Engineering:** Python, Flask, FastAPI, SQLAlchemy, ORMs,
- **Infrastructure as Code:** Terraform, Terragrunt, Cloudformation, Terraform Cloud
- **AI, ML and MLOps:** Apache Airflow, Kafka, LangChain, LangGraph Hugging Face, SageMaker, Weaviate
- **Configuration Management**: Ansible, Chef, Puppet
- **System Administration:** RHEL/Fedora/CentOS, Debian and Arch-based, Container-optimized flavors such as Alpine
- **Containerization:** Docker, Podman, Kubernetes, Istio, Helm, Docker Swarm
- **Container Orchestration:** Kubernetes (Bare-metal, AWS EKS; GCP GKE,..etc), AWS ECS, Fargate
- **Continuous Integration / Continuous Delivery (CI/CD) / GitOps:** Jenkins, ArgoCD, Spinnaker, Github Actions, CircleCI
- **Cloud Platforms:** AWS (Advanced), GCP (Certified Architect), Azure (certified fundamentals, entry-level)
- **Programming/Scripting Languages:**Bash, Python, Golang (to build custom DevOps tools and also gRPC for microservices)
- **Networking:** VPC (Virtual Private Cloud), NLB, Route53, F5 Load Balancers, Good understanding of TCP/IP, CCNA certified
- **Monitoring and Observability:** Prometheus, Grafana, Elasticsearch / OpenSearch, ELK, Datadog
- **Web Servers/L7 Load-Balancers:** Nginx, Traefik, HAProxy, ALB, ELB, Tomcat, Apache