**Objective**

"Highly experienced Engineering leader with over 15 years of expertise in AI/ML and software development, aiming to leverage a strong background in managing AI-driven projects. Passionate about driving technological innovation and applying AI to solve real-world problems."

**Skills:**

* **Programming Languages:** C, C++, Java, HTML, CSS, Python, AngularJS, Node.js, MeteorJS, React, GraphQL.
* **Key Specialized Areas:** Leadership, Software Engineering, Architecture, Distributed System, AI/ML, Large Language Model (LLM).
* **Certifications:** Amazon Certified Solution Architect Professional (AWS), Machine Learning Certified.
* **Platforms &Tools:** AWS, Google Cloud, Azure, Open AI, NVIDIA Infrastructure, Linux.
* **Framework and Analytics:** Docker/Containers, Kubernetes, Serverless, Apache Spark, AWS tools (Redshift, Glue, SageMaker), Flink, DevOps, MLOps, AI/LLM Ops.
* **AI/ML Technologies:** TensorFlow, PyTorch, Ray, LangChain, LlamaIndex, NVIDIA AI Sdk (Rapids, Nvidia Nemo), Large Language Models, Natural Language Processing (NLP), Computer Vision, Recommender Systems, AISS, Pinecone, Vector Databases, MLflow, Weights & Biases (W&B).

**Experience:**

**Senior Principal Engineer/Architect Amazon Web Services: Oct 20 – Present**

* **Built** AI/ML-driven ad personalization at [Amazon Ads](https://aws.amazon.com/blogs/machine-learning/automating-product-description-generation-with-amazon-bedrock/), leveraging LLMs (Claude, Falcon, Llama) for real-time ad targeting, increasing engagement by 30%.
  + Developed LLM-powered ad ranking models using **Reinforcement** **Learning** (RLHF), Deep Speed, and Hugging Face Transformers.
  + Optimized scalable AI inference using PyTorch, NVIDIA Triton, SageMaker, and TensorRT, reducing latency by 50%.
  + Adopted by Uber Eats & major marketplaces, supporting millions of real-time AI-driven ad recommendations per day**.**
  + **Published technical blogs & open-source contributions**:
    - [Blog on the Amazon Ads solution](https://aws.amazon.com/blogs/machine-learning/automating-product-description-generation-with-amazon-bedrock/)
    - [Codebase for the solution](https://github.com/aws-solutions-library-samples/guidance-for-generating-product-descriptions-with-amazon-bedrock)
* **Designed & deployed** **RAG**-based AI chatbots for financial services, healthcare, and customer support automation:
  + Integrated LLMs with Vector Databases (OpenSearch, FAISS, Pinecone, pgvector) to enhance chatbot retrieval accuracy.
  + Developed LangChain-based document retrieval pipelines, improving knowledge retrieval speeds by 40%.
  + Deployed serverless AI inference on AWS Lambda, API Gateway, Step Functions, cutting infrastructure costs by 40%.  
    **Open-Source code and blog** 
    - [RAG based chatbots used in Finanical, healthcare …](https://github.com/aws-samples/aws-bedrock-with-rag-and-react)
    - [Using Large Language model for solving problems](https://github.com/shadhav/bedrockDemo)
    - [LLM for driver safety](https://github.com/aws-samples/sample-ai-driver-safety-helper)
    - [LLM customization](https://github.com/aws-samples/amazon-bedrock-customization-workshop)
* **Architected** large-scale AI/ML infrastructure for LLMs, and real-time AI inference:
  + Built multi-GPU training pipelines using DeepSpeed, Horovod, FSDP, and PyTorch, improving AI model efficiency.
  + Implemented scalable inference with NVIDIA Triton, TorchServe, KServe, reducing inference latency by 50%.
  + Enabled real-time feature engineering with Apache Flink, Kafka Streams, and SageMaker Feature Store.
  + Cost-optimized AI/ML workloads, reducing cloud infrastructure costs by 50% via Spot Instances & AI pipeline optimizations.
* **Scaled** AWS SageMaker Feature Store adoption by 30%, improving AI observability and real-time feature engineering capabilities.
  + Integrated Feature Store with DynamoDB, Redis, and Aurora for high-speed lookups.
  + **Blogs and** **Open-Source Contribution**:
    - [Unveiled machine learning insights by leveraging the Amazon SageMaker Feature Store](https://aws.amazon.com/blogs/machine-learning/unlock-ml-insights-using-the-amazon-sagemaker-feature-store-feature-processor/)
    - [Amazon SageMaker Feature Store Examples on GitHub](https://github.com/aws-samples/amazon-sagemaker-feature-store-examples)
    - [Cross account Sagemaker Feature Store](https://aws.amazon.com/blogs/machine-learning/amazon-sagemaker-feature-store-now-supports-cross-account-sharing-discovery-and-access/)
    - [Amazon SageMaker Feature Store Examples](https://github.com/aws-samples/amazon-sagemaker-feature-store-examples)
* **Led** cross-functional teams of **7 engineers** to integrate AI models with **Agents** into Amazon’s product suite, significantly enhancing system integrity and accelerating project delivery timelines.
* **Devised** executed strategic technology roadmaps and customer-driven enhancements.
* **Engineered** fiscal strategies that slashed operational costs by half while amplifying efficiency by 75%, influencing and advising C-suite executives on pivotal tech-driven cost-saving measures.
* **Optimized** third-party software assessments, yielding a 20% budget alleviation coupled with a 15% leap in performance metrics.

**Software Engineer Manager - JP Morgan Chase Oct 17 – Oct 2020**

* **Managed** a high-performing agile team of 9 software engineers.
* **Orchestrated** a complex, multi-stage overhaul of Chase’s enterprise payment system, culminating in a seamless transition to a modern Spring Boot architecture and the successful launch of a digital wallet platform, managing over $2 billion in transactions annually.
* **Integrated** agile frameworks and backlog refinement techniques, boosting team velocity by 30% and crafting an 18-month strategic roadmap for critical payment platform initiatives and the introduction of a virtual assistant service.
* **Synchronized** cross-functional data, machine learning, and software development teams to ensure timely feature rollouts, reinforcing the company's competitive edge in digital payment solutions.
* **Built** a mentorship program, culminating in a 30% skill enhancement for junior engineers and a surge in job satisfaction exceeding 50%.
* **Refined** requirements elicitation practices, leading to a 42% decrease in engineering change requests, and instituted stringent code review and testing protocols, achieving a 22% reduction in production defects.

**Projects:**

1. **Migration Project:** Migrating legacy code from Struts/WebLogic to Spring/Tomcat, breaking it into granular microservices for cloud readiness, and deploying to Cloud Foundry.
2. **Payment System:** Handling the back-end system for card payments, managing Java, Spring, Struts, and Angular technologies.
3. **Redesign Payment System:** Upgrading old systems to Angular 4, migrating from WebLogic to Spring Boot.
4. **Mobile Wallet:** Building a future-ready wallet capable of holding various cards and wallets (Apple Pay, Samsung Pay) [Link1](https://www.chase.com/business/payments/virtual-terminal), [Link2](https://www.jpmorgan.com/merchant-services/solutions/payment-gateway).

**Software Developer Lead (Verizon) Sept 16 –Oct 2017**

* **Spearheaded** a team of 15 in the creation of the 'Endless Aisle' product for Verizon store used by customer to buy products, which generated $2M within the first 10 months post-launch.
* **Led** the adoption of agile methodologies elevated deployment frequency by 25% and cut defect density by over 20%.
* **Architected** and led the design of Verizon's Chatbot, an AI-driven virtual assistant, by developing a unique pattern matching algorithm and an AIML engine that enhanced responsiveness and user interaction. This initiative managed over 500K daily customer inquiries within a 6-month period on a $2M budget.
* **Revolutionized** Verizon’s legacy payment system by converting SOAP APIs to a robust microservices architecture, coupled with a decentralized database model, resulting in a 40% performance boost and guaranteeing future scalability.

**Projects:**

1. **Endless Aisle:** Designed the Front-end UI with AngularJS 2 and played a major role in payment system integration for kiosks in all Verizon stores.
2. **Chabot:** Led the development of a new Veriza (Chabot) for buying products from Verizon bots, ensuring a simple yet rich customer experience.
3. Involved in designing UI, writing REST APIs, developing a distinct pattern matching algorithm, and implementing NLP processing and AIML mapping for an interactive and humorous Veriza.
4. **Redesigning Backend System:** Implemented a microservices approach using Java, Spring Boot, Spring Data, Angular 2, AIML, and machine learning for backend system redesign at Verizon.

**Founding Engineer (Leap doctor) May 15 – Aug 16**

* **Pioneered** the development of a groundbreaking doctor job portal and real-time scheduling SaaS product, steering an 8-member team from concept to launch. Skillfully translated the product vision into actionable user stories, delineated requirements, and architected pivotal features.
* **Engineered** a sophisticated job listing search engine using the Meteor framework and MongoDB, integrating a smart recommendation engine that matched jobs with doctors' credentials, enhancing user engagement.
* **Core Technologies Utilized:** Leveraged the Meteor JavaScript framework, MongoDB, Java, Spring, and Hibernate to deliver a seamless and scalable user experience.

**Research Scientist (USF). Sep 14 – May 15**

* **Developed** a data visualization research tool from initial requirements to first production release. Defined product scope and managed the build-out of backend Java APIs, libraries, and frontend using JavaScript and Fits standards.
* **Utilized** an in-house Java print capture app to ingest scanned research datasets, implementing storage and querying of large datasets (50K+ records) using an SQL database.

**Java Developer Nvistec Jun 12 – May 14**

* **Contributed** significantly as a Junior Java Developer on a healthcare system, focusing on improving user experience.
* **Developed** a responsive design to enhance user interaction with minimal reading, scrolling, and panning.
* **Designed** UI pages using HTML5, DHTML, CSS3, JSP, JSTL, JSF/Prime Faces, AJAX, JavaScript, DOJO, and Struts taglibs.
* **Developed** the middle tier using Session Bean, Entity Bean/DAO, and Business Objects.

Blog:

* [Optimize Software development with Amazon CodeWhisperer](https://aws.amazon.com/blogs/devops/optimize-software-development-with-amazon-codewhisperer/)
* [Unlock ML insights using the Amazon SageMaker Feature Store Feature Processor](https://aws.amazon.com/blogs/machine-learning/unlock-ml-insights-using-the-amazon-sagemaker-feature-store-feature-processor/)
* [Product Description with GenAI](https://aws.amazon.com/blogs/machine-learning/automating-product-description-generation-with-amazon-bedrock/)
* [Sagemaker Feature cross account support](https://aws.amazon.com/blogs/machine-learning/amazon-sagemaker-feature-store-now-supports-cross-account-sharing-discovery-and-access/)

GitHub:

* [AWS Bedrock with Rag and React](https://github.com/aws-samples/aws-bedrock-with-rag-and-react)
* [Bedrock Demo - Generative AI Code in Python](https://github.com/shadhav/bedrockDemo)
* [Amazon SageMaker Feature Store Examples](https://github.com/aws-samples/amazon-sagemaker-feature-store-examples)
* [AWS Custom Solutions](https://github.com/aws-samples/aws-custom-solutions)
* [Personal GitHub Repository](https://github.com/dhaval09/) , [Fine-tuning Bedrock](https://github.com/aws-samples/amazon-bedrock-customization-workshop)
* [AI to help driver’s safety](https://github.com/aws-samples/sample-ai-driver-safety-helper)

Research paper published:  
[Tracking Learning and Detection of Multiple Objects using Static Camera](https://www.ijcsit.com/docs/Volume%205/vol5issue02/ijcsit20140502298.pdf)

**Education** Master’s of Science in Computer Science: University of South Florida May- 2016

Bachelors of Information Technology - May - 2011