

Dr. Evelyn Reed

San Francisco, CA | (555) 901-3456 | evelyn.reed.ai.arch@progen.com

PROFESSIONAL SUMMARY

Highly accomplished Senior AI Architect with 10+ years of experience leading cross-functional teams in the design, development, and deployment of large-scale generative models and robust MLOps pipelines. Expert in deep learning frameworks, specialized LLMs, and leveraging cloud computing for high-throughput inference systems. Seeking to drive innovative AI strategy in a challenging leadership role.

Skills

- Programming: Python (Expert), R, C++, Go
- ML Frameworks: TensorFlow, PyTorch, Keras, Hugging Face, Scikit-learn
- Cloud MLOps: AWS (SageMaker, Lambda, EC2), Azure ML, Docker, Kubernetes, Terraform, MLflow, CI/CD
- AI Domains: Large Language Models (LLMs), Generative AI (GANs, VAEs, Diffusion Models), Reinforcement Learning (RL), Computer Vision
- Databases Tools: SQL, NoSQL, Spark, Kafka, Distributed Systems

Experience

Senior Principal AI Architect — Synapse Dynamics Labs (2018 – Present)

Palo Alto, CA

- Led the architectural design and deployment of a proprietary 70B parameter LLM, reducing inference latency by 45% using optimized quantization techniques and distributed GPU clusters.
- Developed and implemented robust, scalable MLOps pipelines on AWS utilizing Kubernetes, Terraform, and MLflow, resulting in 99.9% model uptime and 8x faster iteration cycles for research teams.
- Managed a team of 12 Machine Learning Engineers and Data Scientists, overseeing the entire product lifecycle of 5 critical generative AI applications targeting content creation and synthetic data generation.
- Pioneered research into multimodal diffusion models for rapid prototyping, achieving a 30% increase in prototype acceptance rates from business stakeholders and securing \$5M in follow-on

funding.

Lead Machine Learning Engineer — Quantum Insight Systems (2013 – 2018)

Boston, MA

- Designed and implemented high-frequency trading algorithms based on Deep Reinforcement Learning (DRL), resulting in a 12% increase in portfolio alpha year-over-year.
- Established the cloud infrastructure (Azure) for large-scale data ingestion and model training, handling petabytes of financial time-series data.
- Mentored junior engineers and standardized coding practices for the ML team, promoting maintainability and scalability across all research projects.

PROJECTS

Project Chimera: High-Fidelity Synthetic Data Generator — (Tech: PyTorch, AWS SageMaker, Distributed Training, Variational Autoencoders (VAEs), Differential Privacy)

Designed a scalable platform capable of generating millions of high-quality, privacy-preserving synthetic records matching complex statistical distributions of real-world sensitive datasets.

Impact: Enabled secure model training across stringent compliance barriers (HIPAA, GDPR) and reduced external data acquisition costs by \$1.5M annually.

Contextual Language Compression Engine — (Tech: Python, Hugging Face Transformers, Transfer Learning, C++ Backend)

Developed a custom LLM fine-tuning process to compress contextual representations for fast execution on edge devices without significant loss of semantic accuracy.

Impact: Improved model deployment efficiency on mobile platforms by 60% and secured key contracts in the telecommunications sector.

Education

Ph.D. in Computer Science (Specialization: Deep Learning and Distributed Systems),

Massachusetts Institute of Technology (MIT) (2013)

Magna Cum Laude

M.S. in Electrical Engineering, Stanford University (2009)

4.0/4.0 GPA