

# Michael Rodriguez

## AI Model Engineer | Generative AI & NLP Expert

Austin, TX | (512) 555-3698 | michael.rodriguez@example.com  
linkedin.com/in/mrodriguez-ai | github.com/mrodriguez-gen-ai

---

### Professional Summary

Dynamic AI Model Engineer with 8+ years of experience specializing in generative AI systems and large language models. Expert in model deployment, optimization, and scaling strategies. Strong background in developing models for content generation, conversation, and creative applications. Proven track record of enhancing model performance through innovative fine-tuning approaches.

---

### Professional Experience

#### Lead AI Engineer

##### Creative AI Systems, Austin, TX (Sep 2020 - Present)

- Architected and deployed a 40B-parameter multimodal generative model for text-to-image applications. - Reduced inference costs by 65% through model quantization and serving optimizations. - Developed custom RLHF pipelines for aligning model outputs with creative content guidelines. - Led team of 7 engineers across model development, evaluation, and deployment workstreams.

#### Senior Machine Learning Engineer

##### LanguageTech Corp, Seattle, WA (May 2017 - Aug 2020)

- Implemented efficient fine-tuning strategies for domain adaptation of large language models. - Built reinforcement learning systems for training dialogue agents with human preferences. - Developed specialized datasets and pre-training methodologies for creative text generation. - Created evaluation frameworks to assess creative quality and grammatical correctness of outputs.

#### AI Research Engineer

##### Cognitive Systems Inc., San Francisco, CA (Feb 2014 - Apr 2017)

- Researched and implemented early transformer-based architectures for text generation. - Developed custom loss functions for improving semantic coherence in long-form generation. - Created automated evaluation metrics correlating with human judgments of text quality.

---

### Technical Skills

- **Programming Languages:** Python, Scala, C++
  - **ML Frameworks:** PyTorch, TensorFlow, HuggingFace, FastAPI
  - **Model Architectures:** GPT variants, DALL-E, Stable Diffusion, LLaMA, PaLM
  - **Distributed Computing:** DeepSpeed, FSDP, Megatron-LM, Ray
  - **Serving & Inference:** TorchServe, TensorRT, ONNX Runtime, vLLM
  - **Evaluation & Alignment:** RLHF, DPO, Constitutional AI methods
  - **Vector Databases:** Pinecone, Milvus, FAISS
  - **Cloud & MLOps:** Azure ML, GCP Vertex AI, Kubernetes, Kubeflow
- 

### Education

#### MS, Artificial Intelligence

University of Washington, Seattle, WA (Graduated: Dec 2013)

- Thesis: "Generative Models for Creative Text Applications"

**BS, Computer Science**

University of Texas, Austin, TX (Graduated: May 2011)

- Minor in Linguistics - Senior project: Neural approaches to text generation

---

**Patents & Publications**

- US Patent 11,432,567: “Method for Efficient Inference of Generative Language Models” (2022)
  - Rodriguez, M., et al. (2019). “Optimizing Transformer Models for Creative Applications.” *ACL Workshop on Creative NLP*.
  - Rodriguez, M., et al. (2018). “Quantitative Evaluation Metrics for Creative Text Generation.” *EMNLP*.
- 

**Certifications**

- AWS Solutions Architect - Professional (2021)
  - NVIDIA DLI Certified Instructor - Generative AI (2022)
- 

**Open Source Contributions**

- Core contributor to “GenerativeToolkit” - open-source library for generative model deployment
  - Maintainer of evaluation frameworks for creative text applications
- 

**Languages:** English (native), Spanish (native), Portuguese (conversational)