

Ryan T. Gordon, Ph.D.

Software Engineer | 9 YOE | Python | AI/ML | Pipelines | Backend Development
Software Engineering Portfolio Website Link: <https://ryan-gordon-314159.github.io/>
Contact Information: (309) 255-6871 • rtg314159@gmail.com • Yorktown Heights, NY

Versatile software engineer with nearly 9 years of experience developing ML pipelines, managing model training infrastructure at scale (including GPU-backed clusters), building deployment-ready tooling for large-scale AI systems, and constructing general Python automation frameworks at IBM Research.

Skills

Languages:	Python, Java, C++, SQL, HTML
DevOps & Infrastructure:	Linux, UNIX Tools, GPU/CPU Clusters, Docker, Kubernetes, Ray, Spark, Open Data Hub, CI/CD, Git, GitHub, Cloud (AWS, IBM Cloud, Google Cloud), Data Lakes, OpenShift, Resource Management, Fault Tolerance, and MLflow
Machine Learning:	NLP, LLMs, Model Training (Monitoring and Inference), Fine Tuning, Generative AI, PyTorch, Hugging Face, Prompt Engineering, Retrieval Augmented Generation (RAG), and Tokenization
General Software:	Distributed Computing, Data Processing Pipelines, Data Engineering, Lakehouse, Data Analysis, Visualization Tools, Automation, Debugging, Software Test Engineering, Data Structures, Algorithms, Linting, and Logging

Work Experience

Software Engineer (Research Staff Member), IBM Artificial Intelligence Program, T. J. Watson Research Center, Yorktown Heights, NY (September 2023-March 2025)

- Developed Pipelines on Distributed Cluster for AI Model Training
 - Critical for making IBM's Granite 3.0 and Bamba AI models
 - Processed 100s of TB of data at-scale for 12 different languages across dozens of datasets with GPU/CPU-based training for LLM development
 - Pipelines ran on distributed cluster using Open Data Hub and Ray for scalable data processing and fault tolerance with Kubernetes for container orchestration
 - Collaborated with Data Acquisition and Model Training Teams across IBM AI Program
- Constructed Super-Pipelines
 - Reduced data processing time by 50% and improved accuracy by 10% by developing a robust super-pipeline system with automated data validation and checkpointing for ML training at-scale (100s of TBs processed across GPU-backed clusters)
- Built Python SDK for IBM-Internal Retrieval Augmented Generation (RAG) System
 - Increased performance of fine-tuned model by 10% by implementing log-prob parameter distribution analysis with generative AI system
- Evaluated Tokenizers by Ablation Experiments
 - Tokenizer evaluation work allowed IBM AI Program to see direct performance differences among commonly used tokenizers and rank them
- Contributor to IBM Product for AI Model Data Preparation: data-prep-kit
 - Contributor to IBM's internal ML infrastructure tooling for data preparation and tokenization performance benchmarking
 - Integrated into training workflows for LLMs (e.g., Granite 3.0, Bamba AI), supporting evaluation of MLflow-like workflows

Work Experience (Continued)

Software Engineer/Physicist (Research Staff Member), IBM Quantum Computing Program, TJ Watson Research Center, Yorktown Heights, NY (August 2016-September 2023)

- Project Manager: Studying Two-Qubit Gate Fidelity vs Repetition Rate in Multi-Qubit Processors
 - Obtained critical parameters for calculating quantum volume, a necessary milestone in the Quantum Computing Program's roadmap for IBM. These parameters showed how much performance drop occurs when quantum circuits are run faster on these processors
 - Developed Python software for running quantum circuits, collecting two-qubit fidelity data, and analyzing it to produce data visualizations
- Project Manager: Environmental Impact on Qubit Coherence
 - Produced a qubit packaging where world record T_1 time was recorded for a transmon qubit. See: <https://x.com/jaygambetta/status/1395347923123245056>
 - Careful study looking at how environmental factors can influence qubit performance. Factors such as temperatures fluctuations, vibrations, magnetism, and radioactivity were varied in a controlled way and the qubit performance impact was measured
 - Collaboration between 15 scientists and engineers for this project
- Developing Python Automation for Single Qubit Calibration and Characterization
 - This work led to a factor of 10 increase in qubit relaxation time (produced world record)
 - This automation work allowed our team to measure hundreds of thousands of devices and gave us a fast turnaround time for trying new experiments to improve them
 - Software developer/test physicist as part of a cross-functional team to improve qubit coherence. Created software for automating qubit characterization and calibration
 - Cross-functional team collaboration between Device Fabrication and Qubit Measurement teams

Assistant Professor of Physics (Tenure-Track), Western Illinois University (August 2012-August 2016)

- Research and teaching at both the undergraduate and graduate levels

Postdoctoral Research Fellowship, University of Sherbrooke, Canada (May 2011-June 2012)

- Low-temperature investigation of thermal and electrical properties of novel superconducting and magnetic materials in high magnetic fields

Education

- **Ph.D., Physics, Iowa State University**
- **B.S., Double Major in Physics and Mathematics, Western Illinois University**

Patents, Presentations, and Publications

- 11 US patents granted
- 22 invited conference talks
- 19 contributed conference talks
- 27 accepted publications in peer-reviewed journals