

Jonathan Jin

Machine Learning Infrastructure Engineer

jonathanjin.in @ me@jonathanjin.in [linkedin.com/in/jinnovation](https://www.linkedin.com/in/jinnovation) github.com/jinnovation

SELECT EXPERIENCE

| | | |
|--------------------|---|----------|
| Present 03/2021 | Spotify SENIOR MACHINE LEARNING ENGINEER <i>Member of ML Platform. Working on Hendrix, Spotify's centralized ML platform. My focuses revolve around: ML governance; cloud-native infra; and SDK development for MLOps.</i> <ul style="list-style-type: none">▶ Leading the development of a centralized model inventory system for Hendrix as part of the ML Platform's Governance initiative;▶ Led multi-quarter promotion of ML Platform's flagship pipeline orchestration product to general availability; collaborated with product and engineering stakeholders across peer teams within ML Platform as well as "power user" teams – other applied ML teams at Spotify – to: define multi-quarter engineering roadmap; provide technical and project leadership/direction. Contributed to an increase in ML Platform adoption of 50% and increase in user satisfaction of 10%.▶ Providing ongoing mentorship and support to three junior IC team members, i.e. half of the team's IC makeup; holding bi-weekly one-on-ones with each to: provide technical and career mentorship; help identify potential future ownership areas and unblock existing ones; and address ongoing concerns.▶ Implemented formalized SLO tracking for multi-cluster, pipeline execution infrastructure; used Terraform to canonicalize SLO definitions for monitoring and violation reporting in Google Cloud Monitoring; implemented custom Kubernetes listener to implement nuanced and domain-specific SLIs.▶ Bootstrapping early-stage development of centralized, managed Ray infrastructure based on KubeRay for high-performance ML prototyping and experimentation; collaborating with product and engineering leadership to concretize our platform's long-term Ray strategy; <div>Ray Kubernetes Go GCP Terraform Helm TensorFlow TFX Kubeflow Prometheus gRPC</div> | New York |
| 01/2021 12/2019 | NVIDIA SENIOR SYSTEMS SOFTWARE ENGINEER, AI INFRASTRUCTURE <i>Member of AI Infrastructure. Contributor to MagLev, NVIDIA's AI infrastructure for autonomous vehicle development. Also contributed to Modulus, the deep learning SDK for autonomous vehicle R&D.</i> <ul style="list-style-type: none">▶ Initiated development of solution for "hybrid data/model parallelism" using a Ray-based parameter server design and Horovod to enable horizontally-scalable multi-task training;▶ Co-delivered a Kubernetes-based scheduling mechanism to enable priority access to cluster resources for select use cases, e.g. prep for upcoming external demos, via virtual "resource shares"; <div>Ray Horovod TensorFlow Kubernetes Helm gRPC Bazel SwiftStack</div> | New York |
| 12/2019 08/2018 | Twitter MACHINE LEARNING SOFTWARE ENGINEER <i>Member of Cortex, Twitter's central ML platform organization. Worked on: workflow orchestration; experiment management/iteration; and overall ML engineering productivity.</i> <ul style="list-style-type: none">▶ Spearheaded initial integration of TensorFlow Extended (TFX) with legacy Airflow-based orchestration platform to increase agility of workflow development, iterative execution/experimentation, etc.▶ Enabled distributed training of TensorFlow models in Apache Mesos from an Airflow pipeline via Deepbird, Twitter's TensorFlow-based model training/evaluating/serving framework <div>Apache Airflow Apache Aurora TensorFlow</div> | New York |

SPEAKING

- 2022 **TWIMLcon AI Platforms 2022**, "How Spotify is Navigating an Evolving ML Landscape with Hendrix Platform"
- 2022 **MLconf**, "Empowering Traceable and Auditable ML in Production at Spotify with Hendrix"
- 2021 **KubeCon + CloudNativeCon**, "Scaling Kubeflow for Multi-tenancy at Spotify"

SKILLS

| | |
|-------------------------------|---|
| Programming Languages | Python, Go, Bash, C++, Java |
| Machine Learning | Ray, Kubeflow, TensorFlow, TensorFlow Extended (TFX), |
| Distributed Systems | Kubernetes, gRPC, Docker |
| Infrastructure Tooling | Bazel, Prometheus, Grafana, M3, Cassandra, Apache Airflow |
| Cloud Infrastructure | Google Cloud Platform (GCP), Terraform |

EDUCATION

2015 **University of Chicago**, B.S. Computer Science, B.A. Economics