

Karter Barnes

📞 (650) 555-0184 ✉ karter.barnes@example.com 🔗 linkedin.com/example/karterbarnes 📍 San Mateo, CA 94401

SUMMARY

Early-career ML and **cloud infrastructure** engineer pursuing a Bachelor of Science in **Computer Science**, with hands-on experience building Python data, training, and evaluation systems for robotics and computer vision. Developed reproducible pipelines that curate rare events from simulated fleet logs, prepare datasets, and replay **field scenarios** against model releases. Applied AWS, Docker, Kubernetes, Terraform, Ray, Spark, GitHub Actions, Prometheus, Grafana, MLflow, pytest, and GPU-enabled environments across supervised lab and academic work. Comfortable owning defined deliverables from design through testing, documentation, observability, and team demonstration with mentor guidance. Motivated by physical AI that supports safer, more efficient construction of critical **energy infrastructure** and eager to contribute practical infrastructure work alongside robotics and software teams.

EDUCATION

Bachelor's Degree in Computer Science
Stanford University GPA: 4.0
Coursework: Distributed Systems, Machine Learning, Cloud Computing, Autonomous Robotics

2027
Stanford, CA

TECHNICAL SKILLS

- **Programming Languages:** Python, Bash, SQL
- **Cloud Platforms:** AWS, Google Cloud Platform, EC2
- **Containers:** Docker, Kubernetes, CUDA
- **Infrastructure as Code:** Terraform, HCL, AWS IAM
- **CI/CD Tools:** GitHub Actions, Git, pytest
- **Distributed Computing:** Ray, Apache Spark, Parquet
- **ML Experiment Tools:** MLflow, model checkpoints, experiment tracking
- **Observability Tools:** Prometheus, Grafana, structured logging
- **Data Storage:** Amazon S3, Parquet, partitioned datasets
- **Robotics Evaluation:** scenario replay, regression harnesses, mobile robot testing
- **Data Engineering:** data curation, schema validation, data-quality checks
- **GPU Infrastructure:** GPU clusters, CUDA servers, resource scheduling

SKILLS

- | | | | |
|----------|--------------|------------------|-----------------------|
| • Python | • GCP | • GitHub Actions | • MLflow |
| • Bash | • Docker | • CI/CD | • Data pipelines |
| • SQL | • Kubernetes | • Ray | • Experiment tracking |
| • AWS | • Terraform | • Spark | |

EXPERIENCE

ML & Cloud Infrastructure Engineer Intern
University Robotics Infrastructure Lab

June 2026 - Present
Stanford, CA

Support supervised robotics infrastructure work spanning Python data curation, GPU-enabled evaluation, AWS environments, Kubernetes orchestration, and CI/CD. Own a defined prototype from design through testing, observability, documentation, robot validation, and mentor demonstration, giving research teams a repeatable foundation for model development.

- Built a **Python** curation pipeline that mined 180,000 simulated fleet-log records for gust, occlusion, and near-miss events, supplying rare-event data for robotics training.
- Orchestrated Ray preprocessing and evaluation jobs on a GPU-enabled **Kubernetes** cluster, documenting resource settings and recovery steps so researchers could rerun experiments reliably.
- Created **Terraform** modules for AWS storage, IAM roles, and compute configuration, giving weekly lab experiments reproducible **infrastructure** with clearer access boundaries.
- Added Prometheus metrics and Grafana dashboards for queue time, failed tasks, GPU utilization, and incomplete data, helping mentors identify pipeline issues before demonstrations.
- Established **GitHub Actions** checks for unit tests, container builds, schema validation, and smoke evaluations, increasing confidence in changes submitted by four student researchers.
- Presented an **end-to-end** system to lab mentors and tested replay outputs on a mobile robot in an indoor obstacle course, using findings to strengthen a final technical report.

ML & Cloud Infrastructure Engineer

1/3

January 2026 - May 2026

Delivered a supervised course project focused on regression evaluation for autonomous navigation. Combined Python, Docker, MLflow, Spark, Parquet, pytest, and CUDA infrastructure into a traceable workflow that helped a student team compare model releases with consistent evidence.

- Designed a Python **regression harness** that replayed 420 labeled navigation scenarios against successive model checkpoints, recording accuracy, latency, and collision-proxy results for release decisions.
- Packaged evaluator and scenario runner with **Docker**, enabling consistent execution across student workstations and a shared CUDA development server.
- Used **MLflow** for datasets, configuration parameters, model artifacts, and evaluation summaries, preserving experiment lineage during model iteration.
- Implemented **Spark** transformations that normalized camera, pose, and event metadata into partitioned **Parquet** datasets, making scenario selection and analysis faster.
- Added **pytest** coverage for scenario loading, metric aggregation, and malformed-record handling, catching edge cases during instructor review and release rehearsals.
- Demonstrated the harness for a 24-person project class and wrote runbooks covering setup, CI checks, experiment comparison, and limits of simulated testing.

Cloud Infrastructure Research Assistant

September 2025 - December 2025

Distributed Computing Research Assignment

Stanford, CA

Investigated containerized batch processing for robotics perception data under academic supervision. Compared distributed execution patterns, built quality controls and staging workflows, and translated findings into practical recommendations for scalable **evaluation infrastructure**.

- Investigated containerized batch processing for 62,000 annotated image frames and sensor records, clarifying resource needs for **robotics** perception research.
- Compared Spark and **Ray** execution on **AWS** EC2 instances, recording throughput, memory behavior, and retry characteristics to guide framework selection.
- Wrote Python data-quality checks for missing timestamps, invalid bounding boxes, duplicate frame identifiers, and inconsistent train-validation splits.
- Built an **S3** staging workflow with Terraform configuration and documented least-privilege access assumptions for a safer course research environment.
- Exposed job duration, record counts, and exception categories through structured logs and a **Prometheus**-compatible metrics endpoint, improving experiment visibility.
- Presented findings at an undergraduate research poster session, recommending Ray for parallel evaluation and Spark for larger tabular transformations.

PROJECTS

Fleet Rare-Event Curation Pipeline 📅 2026

Created a Python and Ray workflow that mines simulated fleet logs for rare gusts, occlusions, and near-misses. Added schema checks, cloud storage, metrics, and repeatable evaluation steps so curated data could support safer robotics model development.

Model Release Regression Harness 📅 2026

Built a Dockerized evaluator that replays labeled navigation scenarios against model checkpoints. MLflow tracking, Spark preparation, Parquet storage, and pytest checks made results easier to compare before simulated release reviews.

LEADERSHIP & AWARDS

- Outstanding Undergraduate Research Poster, Stanford Computer Science Research Showcase, 2026.
- Dean's List, Stanford University, 2025-2026.
- Second Place, Stanford Autonomous Systems Hackathon, 2025.
- Computer Science Department Scholarship, Stanford University, 2025-2026.

CERTIFICATIONS

- AWS Academy Cloud Foundations 📅 2026
- Google Cloud Computing Foundations Certificate 📅 2026
- Kubernetes Fundamentals, Linux Foundation Training 📅 2026

PROFESSIONAL AFFILIATIONS

- Cloud Infrastructure Study Group Coordinator, Stanford Computer Science Association, 2025-2026.
- Peer Mentor, Distributed Systems Course, Stanford University, 2025-2026.
- Student Research Poster Committee Volunteer, Stanford Robotics Center, 2026.

LANGUAGES

- English (Native)
- Spanish (Intermediate)

ADDITIONAL INFORMATION

Work Status : Authorized to work in United States. No sponsorship required.

REFERENCES

AVAILABLE ON REQUEST