

Kevin Hopkins

kevinhop@alumni.usc.edu | 630-442-4823 | kevdozer1.com

ML engineer building data pipelines for robotic foundation model training

EDUCATION

University of Southern California, Viterbi School of Engineering

Jan 2023 - Dec 2024

Master of Science in Computer Science

University of Wisconsin, Madison, College of Letters & Sciences

Aug 2019 - May 2022

B.S. Computer Science; B.S. Communication Arts (Film)

EXPERIENCE

Scout Space, Machine Learning Engineer

Feb 2024 - Present

- Founded Scout's research team and led TALON, a sim2real testbed recording isomorphic real and synthetic imagesets, including the synthetic render and data-capture pipeline used to validate sim2real performance
- Led an on-orbit Gaussian splatting product for edge 3D reconstruction of spacecraft, the basis of robotic de-orbiting missions and the company's most-viewed technical report
- Implemented end-to-end architecture, training, and deployment of a large 6-DOF state estimation model for autonomous spacecraft rendezvous, fed by a novel synthetic data pipeline
- Built a VLM-based hierarchical segmentation tool producing masks of space objects and sub-objects with per-instance metadata, an annotation system feeding the training stack
- Created a multi-camera Structure-from-Motion technique deriving per-object point clouds without sequential captures, generating 3D supervision data for downstream models

USC Institute for Creative Technologies, ML Research Assistant

Jun 2023 - May 2024

- Built multimodal perception models and the associated data pipelines (facial, postural, gestural) for human behavior understanding on embodied AI systems
- Ran cross-demographic fairness evaluations and optimized architectures with weight-sharing for real-time inference on resource-constrained systems

INDEPENDENT RESEARCH

robolabel: open-source annotation pipeline for robotic learning data at scale

May 2026

- Built an automated, model-agnostic pipeline that generates pi0.7-style conditioning metadata from robot video at scale for \$0.02/episode: subtask boundaries, quality/speed bands, active effector labels, and more
- Wrote annotations to a versioned Parquet schema and exported them into LeRobot's subtask format, alongside a quality gate that flags weak or degenerate labels for review
- Added value-aware curation: scoring each episode by quality and novelty, then compressing each episode according to its score, keeping high-value data full-fidelity and the rest compact, for sub-linear storage growth

World Model Fine-tuning Testbed & Vision-Signal Ablation

Apr 2026

- Designed and open-sourced a fine-tuning testbed probing which vision signals (depth, motion tracks, masks, and others) improve LeWorldModel policy learning on real manipulation episodes
- Devised the dataset filter, perceptive data pipeline, ablation methodology, and evaluation infrastructure, releasing the full study as a comprehensive technical report
- Documented architectural bottlenecks on dense-signal reconstruction and identified the data-scale conditions under which fine-tuning on perceptive signals helps or hurts policy learning

PROJECTS

Gaussian Splat Robotic Manipulation RL Playground

- Built a PyBullet RL environment for manipulation with Gaussian-splat scene integration and an automated evaluation pipeline tracking 90% grasping and 70% lifting success across diverse geometries

Indoor Scene Perception

- Built a pipeline producing semantically-labeled 3D point clouds from household images, unifying pretrained segmentation and depth models for downstream scene understanding

SKILLS

Fundamentals: Python, C++, C, Linux, Git, AWS, CUDA

Machine Learning: PyTorch, JAX, RL, world models, VLAs, multimodal learning, diffusion models, cloud training

Data Engineering: Parquet, HDF5, Docker, Kubernetes

Computer Vision: VLM development, segmentation, 3D reconstruction, SfM, depth/pose estimation, OpenCV

Robotics: sim2real validation, ROS, MuJoCo, PyBullet, SLAM, URDF