David Jackson

Orlando, FL | 407-405-5154 | davidtjackson1029@gmail.com
www.linkedin.com/in/david-jackson-89135727a | github.com/JylerTackson | jyler.dev

SUMMARY

AI/ML engineer and M.S. Computer Science candidate (AI/ML specialization) with hands-on experience in multi-modality, computer vision, object detection, contrastive representation learning, and deployment of models to edge GPU hardware (Jetson Orin AGX, Jetson Nano, DGX Spark).

EDUCATION

University of Central Florida, Orlando, FL

M.S. Computer Science, Specialization: AI/ML — Aug 2026 – May 2028

Coursework: AI Computing Infrastructure, ST: Applied Machine Learning, ST: AI Knowledge Systems

B.S. Computer Science, Minors: Mathematics, Physics — Aug 2023 – May 2026

TECHNICAL SKILLS

Languages: Python, C/C++, Java, TypeScript/JavaScript, SQL

Frameworks & Libraries: PyTorch, TensorFlow, Keras, React, React Native, Flutter, Node.js, Express.js, Flask

ML & Computer Vision: Object detection (YOLO), multi-modal contrastive learning (CLIP-style encoders), model evaluation and benchmarking, RAG, Light-RAG, LoRA / QLoRA fine-tuning, LangChain, LangGraph, transformer inference optimization (vLLM, SGLang, PagedAttention, FlashAttention)

Edge & Accelerated Compute: NVIDIA Jetson Orin AGX, Jetson Nano, DGX Spark, GPU inference profiling, offline/air-gapped deployment

Cloud & Infrastructure: Google Cloud Services, Azure, FastAPI, REST APIs, Linux, Git, Docker, Kubernetes, Neo4j, PostgreSQL, virtual machines

Mathematics: Linear algebra, matrix methods, probability and statistics, convolution and transform methods, trigonometry

Tools & Methodologies: Jira, Confluence, Agile, Linear, Claude Code, Cursor

PROFESSIONAL EXPERIENCE & RESEARCH

Graduate Research Assistant, Dr. Lou's UNBOUNDED AI Lab — Graduate Researcher | Aug 2026 – Present

- Present and critique current literature on multi-agent LLM orchestration and model routing; produce a cross-paper evaluation of their shared approaches and current open-source contributions, creating key comparisons between the baseline numbers, methods, and novelty of the approach.
- Investigating scheduling and memory-efficiency limitations in the vLLM serving and router path and contributing to a co-authorship with a PhD student for improvements on LLM serving infrastructure targeting tail latency and GPU utilization under contention.

Command Post Technologies — AI Solutions Architect — Orlando, FL | Aug 2026 – Present

- Designing a multimodal deep learning system that refines low-quality geospatial vector features against aerial and satellite imagery, using per-modality encoders (imagery, vector geometry, geocoordinates, depth) contrastively aligned with CLIP-style methods after GeoCLIP/UniGeoCLIP.
- Architecting the company's enterprise generative AI rollout on AWS Bedrock: model access and provisioning, Bedrock Guardrails for content and policy filtering, and the shared service layer other teams build against.
- Developing a CUI-compliant knowledge base and MCP server that exposes controlled document retrieval to Bedrock-hosted models, with an in-server ingestion pipeline handling document types such as PDF, DOCX, and PPTX.
- Packaged and automated the team's AWS deployment tooling as an internal Python distribution; diagnosed and remediated the distribution and access failures blocking multi-host rollout.

Command Post Technologies — AI/ML Engineer Intern — Orlando, FL | *Jun 2026 – Aug 2026*

- Project Lead and Full Stack Engineer on an Agentic Marketing AI Platform, owning Frontend, Backend, Database, MCP Deployment, and AI Deployment; cut manual creation of Social Media Posts, Job Postings, and Newsletter down by 70%.
- Interviewed the marketing team to scope pain points and source grounding content for a LangGraph RAG pipeline, and built in-UI rating and feedback capture that persists scored response pairs as a labeled preference dataset.
- Built the Backend on a company-hosted NVIDIA DGX Spark utilizing Python 3.12 with Fast API hosting vLLM running GPT-OSS-20b for RAG retrieval and main document drafting, as well as a Stateless MCP Server hosted on a P100 hosting Ollama running a LLaVA-3 b model for image analysis.

Growing Martin Minds — **Google Cloud Developer** (Part-Time) | Jan 2025 – Jun 2026

- Architected an automated backend for payroll, invoicing, and tutor log tracking; built Apps Script automations for document generation and processing.
- Designed data pipelines and validation logic that eliminated manual entry error paths.

Element Construction Services of Orlando — General Office Specialist (Part-Time) | Aug 2020 – Jan 2025

PROJECTS

Closed-Loop Perception-to-Action Agent (Independent) | 2026 – Present

- Building an end-to-end vision-to-action agent for a partially observable real-time strategy game, operating under a closed loop of Context-Action-Feedback (CAF) while interacting with an emulator of a real-time strategy card game.
- Sourcing and labeling several custom datasets to train a YOLO model for localization and classification of assets used by the interface.
- Designing a two-tier control architecture that keeps the real-time loop on a fast reactive policy while retrieval-backed reasoning runs at a coarse cadence to combat retrieval latency that comes with each decision made.
- Designing a custom Embedding model for efficient retrieval utilizing a MOE-based framework to improve the base accuracy of the MMEB-V2 model while attempting to reduce the TFLOPs.
- Benchmarking a hand-authored heuristic policy against a learned policy trained offline on replay data, holding the perception stack fixed so the comparison isolates the decision layer.

VITALS — Search-and-Rescue Agentic Base Station (Lockheed Martin–sponsored capstone) | Aug 2025 – May 2026

- Owned the mapping and planning agent in a 4-agent SAR architecture (vision, retrieval, reasoning, mapping/planning), coordinating over MCP and ACP message contracts, on a 6-engineer team under Lockheed Martin sponsorship.
- Inherited and re-architected the terrain preprocessing pipeline: PostGIS-backed OpenStreetMap queries through SQLAlchemy, EPSG:3857 to EPSG:4326 reprojection, R-tree spatial indexing over a tiled search grid, with per-tile feature classification driving search priority.
- Redesigned multi-drone task allocation from a single greedy heuristic to a two-layer planner: A* global routing over a live costmap plus a local avoidance layer for obstacles absent from the OSM basemap, with non-overlapping timestamped sector assignment for persistent coverage.
- Specified the retrieval subsystem: bge-small-en-v1.5 encoder producing 384-dimensional embeddings into ChromaDB with cosine ranking and metadata filters, and ParquetDB cold archive for events aged out of hot storage.
- Targeted deployment on Jetson AGX Orin 64GB inside a 15–60 W configurable power envelope for disconnected field operation for the base station, and a Jetson Nano configured onto the drone for the vision edge.

Kestrel — Autonomous Videography Drone (Blue Origin–sponsored club project) | Jun 2024 – Aug 2025

- Defined system architecture and data pipelines across sensing, perception, and control; applied systems engineering practice to coordinate subteams and delivery.
- Designed a MobileNetV3 CNN integrated into a DeepSORT model, achieving 23.64% mAP, 35.34% Rank-1 Accuracy, and 48.99% Rank-5 Accuracy.
- Designed and simulated the custom Kestrel drone, achieving a best-case flight time of 40 min and a maximum speed of 40 mph. Sensor suite included a camera on a dual-pivot mount, 16 short-range ToF sensors, and 4 long-range LiDAR sensors for obstacle detection.

LEADERSHIP & VOLUNTEERING

- Foundation Exam review sessions for UCF CS students covering C, data structures, and algorithms.
- Annually deliver an introductory AI presentation to a 3rd-grade class on modern-day algorithms.
- LightOrlandoHungerProject: Assembly line to create a million meals for those in need.