

Demetri Pappas

demetripappas.com | dmetripappas@gmail.com | Git: demetripappas123 | linkedin.com/in/demetri-pappas

EDUCATION

University of Central Florida

May 2027

B.S. in Computer Science.

GPA 3.8

University of Miami

Transferred 2024

Presidential Scholar (National Merit Finalist), Benacquisto Scholar, Foote Fellow Honors Student.

EMPLOYMENT

Puretalk AI– ML Engineering Intern

May - Sep 2025

Orlando, FL

- Built Python-based evaluation pipelines to benchmark proprietary LLM (RUTH©) performance across reasoning, retrieval, and instruction-following tasks, enabling rapid regression testing between model versions.
- Performed large-scale error analysis across thousands of model outputs, developing failure-mode taxonomies that guided improvements to prompting and retrieval systems.
- Conducted comparative benchmarking of proprietary transformer models and retrieval-augmented generation systems, analyzing tradeoffs between factual accuracy, inference latency, and operational cost.

Puretalk AI– Software Engineering Intern

Jan - May 2025

Remote, NY

- Led frontend development of a production SaaS platform using React, Next.js, Tailwind, and Node.js.
- Implemented responsive UI/UX improvements, API-driven components, dashboards, and AI-agent interfaces that supported growth from 40 to 150+ clients post-launch.
- Managed agile development cycles through Scrum/Jira, collaborating with CTO, CEO, and design teams while owning frontend implementation, CI/CD updates, and iterative product improvements.

FlexPT, LLC– Founder, Fullstack Developer

May 2025 - Current

Orlando, FL

- Built the frontend with React using dynamic UI components and state management for realtime data.
- Designed backend services with Supabase (PostgreSQL), implementing authentication, database operations, and server-side business logic.
- Deployed the application on Vercel, optimizing request routing and helper functions for scale, and deploying serverless functions for low-latency business logic operations.
- Implemented AI workout program generation and editing with the Open AI API, designing custom system prompts and a context-retrieval pipeline to minimize token usage and reduce latency.

PROJECTS

CertForge– *Microsoft Agents League Hackathon*

Jun 2026

- Engineered a typed multi-agent orchestration framework in Python that coordinates project intake, knowledge retrieval, planning, assessment, and validation through modular Pydantic v2 data models and deterministic execution pipelines.
- Implemented retrieval-augmented generation infrastructure using Azure AI Search and local semantic knowledge stores, with fail-safe fallback mechanisms, lazy-loaded cloud dependencies, and managed-identity authentication.
- Containerized and deployed an application on Azure Container Apps with CI/CD-ready infrastructure, structured telemetry logging, RBAC, and Azure OpenAI integration for intelligent workflows.

High Performance Memory Allocator– C

Mar 2026

- Built a custom memory allocator that reclassifies blocks between size buckets and a red-black tree for larger blocks.
- Designed test suites demonstrating 97.8% reuse rates, 1.66x faster speeds than glibc's malloc for small and random allocations, and perfect 1.0 internal fragmentation and 0.91 external fragmentation.
- Added a union between heap pointers for small blocks and tree metadata for larger blocks in the block header, and bit-packed free and color flags for tree blocks reducing overhead from 20% to 14% without increasing fragmentation.

SKILLS

ML/AI: PyTorch, Orchestration, LLMs, RAG, LLM Evaluation, Python, EDA, Transformer Language Models.

Software: Full-Stack Development, REST APIs, CI/CD, System Design, Database Design, Git, UI/UX.