Anton Vice

AI Infrastructure and ML Platform Engineer

Florida | anton96vice@gmail.com

github.com/antonvice | linkedin.com/in/anton-vice | antonvice.github.io

AI engineering leader and co-founder building data platforms and private AI infrastructure. Deployed an eight-H100 stack for ~200 NBCUniversal users. Owns architecture and delivery, with AI-assisted implementation, code review and testing.

Experience

Grably | Co-founder and Co-owner | May 2024–Present

CTO May 2024–Aug 2026 (part-time through Oct 2024); ongoing advisory work. Techstars; $1M raised; $7M total company sales.

- Built Python/Go ingestion and delivery infrastructure on-premises and on DigitalOcean, processing and delivering approximately 1–5 TB of client data monthly.
- Continued lakehouse storage optimization and indexing after the CTO tenure; Grably's combined stored datasets totaled approximately 120 PB.
- Built media pipelines, ClickHouse indexing and customer APIs for a platform with ~500,000 app users and ~10,000 Latin American egocentric-data contributors.
- Led eight technical staff across mobile, frontend, backend, data science and blockchain, plus regional collection/support managers in Latin America, Belarus and Nepal.

X-lab Robotics | Robotics Researcher | Jan–Jun 2026

Part-time contract. Built egocentric-data processing and VLA/world-model evaluation workflows, including simulation for SO-101 and ABB YuMi.

NBCUniversal | AI and Robotics Engineer | Feb–Oct 2024

Contract

- Deployed an eight-H100 private AI stack with Docker, OpenWebUI, LangChain and Arize for approximately 200 internal users; fine-tuned Llama 3.1 on the cluster.
- Led two senior engineers and two interns on private LLM, document AI, executive-assistant, media-avatar and robotics projects.
- Built retrieval and document-processing workflows and worked on world-action models for robotic dogs.

Earlier experience

Esmee.io | Staff ML Engineer, contract | Aug–Dec 2023
Python EEG/EMG workflows and prosthetic-arm prototypes with ESP boards.

NamelyAI | AI Engineer, contract | Jan–Aug 2023
Voice agents and FastAPI/Flask integrations with telephony APIs and PostgreSQL.

DeepVice | Data Scientist and Python Developer, freelance | Mar 2019–Dec 2023
Recommendation systems, NLP, chatbots and Python APIs for clients.

Technical skills

Python, Go, SQL • On-premises, DigitalOcean, Docker • ClickHouse, PostgreSQL • PyTorch • GPU inference, fine-tuning, RAG and evaluation

Education and training

Florida International University | BA Philosophy, minor in Psychology | 2019
SkillFactory online data science program • Machine learning coursework via Coursera