Anton Vice

Curriculum vitae · AI systems and applied AI engineering

Florida | anton96vice@gmail.com
github.com/antonvice | linkedin.com/in/anton-vice | antonvice.github.io

AI engineer and co-founder building private AI systems, data infrastructure and LLM-powered products. Led engineering and distributed data-collection teams, deployed private GPU infrastructure, and delivered systems used by technical and creative teams.

Experience

Grably | Co-founder and Co-owner | May 2024 to Present
CTO May 2024 to August 2026; part-time through October 2024. Subsequent lakehouse work and limited customer and partner advisory involvement.

- Led eight technical staff across mobile, frontend, backend, data science and blockchain, plus three Latin American support managers and regional collection managers in Belarus and Nepal.
- Built ingestion, processing, indexing and delivery infrastructure using Python, Go, ClickHouse, on-premises systems and DigitalOcean; processed and delivered 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.
- Supported a Telegram crowdsourcing app with approximately 500,000 users and coordinated camera distribution and collection operations for a network of approximately 10,000 Latin American egocentric-data contributors.
- Worked with data procurement teams and automated collection operations. Grably participated in Techstars, raised $1M and generated $7M in total company sales.

NBCUniversal | AI and Robotics Engineer | February to October 2024
Contract

- Deployed a private eight-H100 AI stack using 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 across private LLM, document-processing, executive-assistant, media-avatar and robotics projects.
- Built structured-output document parsing, comparison and document-generation workflows, alongside retrieval over internal documents.
- Worked on a robot-dog character project, world-action models, a fine-tuned executive assistant with tool access, and a speaking avatar workflow generating visual, voice and language assets.

Additional experience

X-lab Robotics | Robotics Researcher | January to June 2026
Part-time contract. Built egocentric-data processing and VLA/world-model evaluation workflows, including simulation for SO-101 and ABB YuMi.

Esmee.io | Staff ML Engineer | August to December 2023
Contract. Developed Python EEG/EMG workflows and prosthetic-arm prototypes using ESP boards.

NamelyAI | AI Engineer | January to August 2023
Contract. Built voice agents and FastAPI/Flask integrations with telephony APIs and PostgreSQL.

DeepVice | Data Scientist and Python Developer | March 2019 to December 2023
Freelance. Built recommendation systems, NLP applications, chatbots and Python APIs for clients.

Selected independent projects

Zelflayer hybrid retrieval
Built a hybrid retrieval system combining graph and vector search.

Telegram storage filesystem
Built a macFUSE filesystem using Telegram Saved Messages for file storage, PostgreSQL for the filename and location index, and a bot for upload and retrieval.

Mojo ML experiments
Ported Whisper to Mojo and built a data-marking model for tracing copies of datasets.

BCI experiments
Built an experimental OpenBCI interface for an LLM and explored whether a custom model could map recorded brain signals to words.

Technical skills

Python, Go, TypeScript, SQL. PostgreSQL and ClickHouse. Docker, DigitalOcean and on-premises GPU and storage infrastructure. PyTorch, LangChain, OpenWebUI and Arize. Retrieval, document ingestion, evaluations, fine-tuning, APIs, voice agents and experimental robotics.

Education and training

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