Vincent Philavong

+1 (847) 483-4501 | vphilavong@wisc.edu | GitHub | LinkedIn

Education

University of Wisconsin – Madison | Madison, WI Aug 2022 – May 2026
Bachelor of Science Computer Science, Certificate in Data Science
Relevant Coursework: Data Management, Optimization, Data Structures and Algorithms, Operating Systems.

Tech Career Accelerator Charles University | Prague, Czech Republic (Incoming) Sep 2025 – Dec 2025

Experience

SAIL Summer AI Laboratory Internship (Sponsored by UW–Madison CDIS & OpenAI) Jun 2025 – Current
AI Laboratory Intern *Madison, WI*

- Selected for a highly competitive AI internship (<25% acceptance) and worked hands-on with OpenAI tools (GPT-4o, DALL·E, Codex) in applied research projects.
- Building an AI copilot that uses OpenAI embeddings and prompt-driven UI generation; improved dev tool UX by 30% via rapid prototyping with React and Tailwind.

UW Tech Exploration Lab Jun 2025 - Current
Volunteer Contributor *Madison, WI*

- Partnered with a UW researcher to fine-tune a custom LLM for qualitative data analysis using Hugging Face Transformers, PyTorch, and Elasticsearch, enhancing thematic accuracy and reducing inference latency by ~20%.
- Led lab outreach efforts including alumni newsletters, marketing assets, and industry engagement campaigns, contributing to a 40% increase in partner interactions and event participation.

Projects

Streamlit AI Chatbot with Fine-Tuning and RAG Spring 2025

- Developed a Streamlit-based chatbot integrating Meta's LLaMA-3.2-1B-Instruct with fine-tuning and Retrieval-Augmented Generation (RAG) to support course-specific Q&A for lecture-based learning.
- Fine-tuned a 4-bit quantized LLaMA model using LoRA and lecture transcripts; monitored loss trends via Weights & Biases to iteratively refine topic-specific response quality.
- Engineered a RAG pipeline with Elasticsearch and Haystack for dynamic top-3 document retrieval, improving response grounding compared to fine-tuned-only model.

Jammy, Interactive Music Visualization Web App Spring 2025

- Built a full-stack music analytics platform using React.js, Django REST API, PostgreSQL, and Redis, reducing data load time from 12s to 1.5s (87.5% improvement); containerized deployment via Docker ensured consistent local and cloud execution.
- Implemented emotion classification ML pipeline using PyTorch, DistilBERT, and scikit-learn, enabling 8-category music mood prediction with confidence scoring for user-selected songs.
- Integrated Spotify Web API and Genius lyrics scraping with sentiment analysis libraries and custom ML models to deliver end-to-end music mood analytics across lyrics, tempo, and emotion profiles.

Skills

Languages: Python, JavaScript/TypeScript, SQL
Frameworks & Tools: Django, REST API, Streamlit, Next.js, Tailwind CSS, Docker, PostgreSQL, Weights & Biases, Elasticsearch, pandas, NumPy, ARIMA, PyTorch, SkLearn.
Concepts: Machine Learning, LLM Fine-Tuning, Retrieval-Augmented Generation, Time Series Analysis, Optimization