BAIYI WANG (EMOO)

(1)215-596-9449 | e: emoowang392@gmail.com | Philadelphia, PA |
LinkedIn: www.linkedin.com/in/baiyi-wang-8819262bb | GitHub: https://github.com/emoow |

EDUCATION

University of Pennsylvania — Philadelphia, United States
Master of Computer and Information Technology — Aug 2025 – May 2027

- **Courses**: Software Development (Java), Computer Systems (C), Applied Machine Learning (PyTorch).

Nanjing University — Nanjing, China
Bachelor of Laws in Sociology — Sept 2021 – June 2025

- **Selected awards**: Ruli Scholarship (3%), National Scholarship for the 2022-2023 & 2023-2024 school year (10%).

TECHNICAL SKILLS

- **Programming Language:** Python, Java, C, C#, Swift
- **Skills:** PyTorch, RLHF, LangChain, LangGraph OpenAI API, Elasticsearch, SQL, React.

WORK EXPERIENCE

Starways AI — Beijing, China
AI Engineer — Apr 2025 – Aug 2025

- Co-founded and scaled Starways AI by securing $0.7M in first-round funding from Sequoia China and growing the team from 4 to 12 members. Led AI development initiatives in the initial phase.
- Developed an AI-powered recruiting agent platform with three other founders by designing a planner-worker agent structure, creating 6 functional agent tools and 3 benchmark workflows (resume parsing, HTML-based rewriting, and interactive editing) using Python, Langchain, and FastAPI.
- Enhanced AI recruiter-candidate matching accuracy by 25% through training agents to perform SQL and Elasticsearch-based search operations.
- Built a first-of-its-kind internal evaluation platform across 6 core AI workflows by architecting integrated database systems, prompt libraries, and debugging environments with cross-functional team delivery.

Moonshot AI — Beijing, China
Intern, Product Manager — Sep 2024 – Apr 2025
Moonshot AI is China's leading AI unicorn ($3.3B valuation) backed by Alibaba and Sequoia China, pioneering long-context AI with their Kimi chatbot that processes 200K+ characters per conversation.

- Led prompt engineering for search tool and MCP framework, enabling multimodal content processing (images, PDFs) with multi-search capabilities in a single round to overcome N2S model limitations. Designed frontend UI and ChatBot interfaces in Figma and HTML, authored PRD documentation.
- Independently planned and led the "image-based search" feature, solving technical bottlenecks where the original search model failed to handle image and document scenarios; completed early-stage research, integrated image-based search interfaces, and fine-tuned prompts for small-model search decision-making.
- Enhanced the KimiV multimodal AI model to over 90% accuracy across image similarity, common sense understanding, and intent analysis by developing benchmark standards and evaluation guidelines.
- Increased OCR model accuracy to over 90% and eliminated $3M in annual API costs by replacing legacy interfaces and implementing Markdown/Mermaid visualization formats for enhanced data extraction.

Deloitte China — Shanghai, China
Intern, Risk Management Consulting — June 2024 – Aug 2024

- Built an AI legislation checker achieving around 80% accuracy and a research report generator reducing creation time from 10 hours to 1 minute using LLM+RAG architecture.
- Created to B interactive risk management demos in Axure and improved system performance by 10% through comprehensive testing supervision and optimization of Y identified issues.

PROJECTS

Image2GPS: CNN-Based Geolocation Model (Team of 3) HuggingFace: https://huggingface.co/cis519FinalProject| Dec 2025

- Built ResNet-50 regression model predicting GPS coordinates from images with 38.9m average error (95% over baseline)
- Collected 2,500+ multi-angle street view images and engineered a two-stage classification-regression pipeline

Feynman - Hack Princeton 2025 (Team of 2) GitHub: https://github.com/danleeaj/feynman **|** Nov 2025
Multi-modal AI tutoring platform with real-time vision, voice, and interactive drawing in 36 hours

- Designed multi-agent vision pipeline enabling live canvas handwriting, text input, and voice transcription. Integrated a voice agent that delivers skeptical Socratic critique and visual tool calls to parse the handwriting canvas into structured text and Mermaid diagrams.
- Implemented real-time UI updates with Next.js (TypeScript) and custom canvas components, synchronizing canvas state, vision inference, and spoken feedback to create a fluid interactive teaching loop.