

Wenyu Chiou

AI Agent Infrastructure & Evaluation Engineer

Bethlehem, PA · wec324@lehigh.edu · +1 (484) 937-2528 · github.com/WenyuChiou · linkedin.com/in/wenyu-chiou · wenyuchiou.github.io

I build AI-agent and research infrastructure — MCP tooling, multi-agent orchestration, and agent harnesses — and I apply LLMs to agent-based models of complex human–water systems, with the governance that keeps agent behavior trustworthy.

Open to Summer 2027 internships in AI evaluation, research engineering, and risk analytics.

Work eligibility: F-1 student, CPT-eligible for internships (no sponsorship required).

— EXPERIENCE

Graduate Research Assistant — Complex Adaptive Water Systems Lab, Lehigh University Aug 2024 – present

- Apply LLMs as the **behavioral engine of agent-based models** of complex human–water systems, grounded in a **937-household survey I designed and fielded**; co-authored the published *Water Resources Research* (2026) study on this LLM-driven agent-based simulation.
- Building **WAGF**, an **AI-governance framework** that screens each LLM-agent decision against domain rules and behavioral theory before it can change simulation state (research software in preparation for release).
- Built **FLOODABM**, a **52,141-household** coupled agent-based × catastrophe simulation, statistically calibrated and validated against real OpenFEMA insurance-claims data.

Independent Open-Source AI Infrastructure

2024 – present

- Built an **evaluation suite** for research-hub's own literature search that checks whether the papers that should be found actually appear in the top results — it measured my own system missing nearly all of them, so I **kept that failing result in the repo as a recorded audit baseline** instead of deleting it.
- Built **research-hub**, an **MCP server + CLI** that makes Zotero, Obsidian, and NotebookLM AI-operable — searching scholarly databases and syncing results. **PyPI research-hub-pipeline v1.1.1**; listed in **Awesome MCP Servers**; tested on Windows, macOS, and Linux.
- Built a multi-agent tooling stack — **ai-research-skills**, **fable-method-harness** (cost-routing that matches strong-model quality at **~2.5× lower cost** in a pre-registered benchmark), **agent-collab-skills**, and **codex-delegate** — on the **OpenAI and Anthropic APIs, MCP, and Ollama**; each agent's work is verified from git state, tested on Windows and Linux.
- **18 merged pull requests** into third-party open source — including **10 into a 58k-star LLM stock-analysis platform** (end-to-end Taiwan-market support: market routing, institutional-flows fetcher, circuit breaker, drift tests), a Windows file-permissions security fix ([langchain-ai/openwiki](#)), and an engine-selection fix with a regression test that fails on main ([BuilderIO/agent-native](#) ×4).
- **awesome-agentic-ai-zh** — a trilingual agentic-AI learning curriculum with automated checks that keep all three languages in sync; **4.7k+ GitHub stars**, 622 forks.

— PUBLICATIONS & PRESENTATIONS

Yang, Y. C. E., & Chiou, W. (2026). Leveraging Large Language Models for Agent-Based Simulation of Human–Water System Interactions. *Water Resources Research*, 62(6), e2025WR042111. doi.org/10.1029/2025WR042111

Published

Chiou, W., & Chiou, H. — “What Makes AI Believable? Comparing AI and Human Responses in Mental Health Risk Assessment.” *ISDSA 2026 Annual Meeting*; recorded presentation.

— EDUCATION

Ph.D., Civil & Environmental Engineering, Lehigh University

2024 – present · expected 2028

M.S., Hydrological and Oceanic Sciences, National Central University

Sep 2021 – Jun 2023

B.S., Earth Sciences, National Central University

Sep 2017 – Jun 2021

— SKILLS

AI & agent systems: OpenAI & Anthropic APIs · Model Context Protocol (MCP) · multi-agent orchestration · agent governance & validation · retrieval & search · Ollama · research automation (Zotero / Obsidian / NotebookLM)

Programming & tooling: Python · pytest & CI · Git · cross-platform (Windows / POSIX)