

Baiyi Lian

📍 Helsinki, Finland 📩 echolian97@gmail.com ☎ +358 413142124 ⚡ github.com/Echo-Lian
LinkedIn: linkedin.com/in/baiyi-lian-10244b28a 🌐 echo-lian.github.io

SUMMARY

I am an AI Developer with hands-on experience in LLMs and cloud technologies, with a passion for designing AI agents and bridging business requirements with technical solutions. I am seeking to leverage AI engineering expertise in building cutting-edge AI applications.

EDUCATION

University of Helsinki, Helsinki, Finland

Master of Arts / Language Technology

Sep. 2023 – Jun. 2025

- GPA: 3.52/5.0
- Relevant Coursework: Deep Learning, Building NLP Application, Models and Algorithms in NLP, Advanced Programming, etc.

PROJECTS

MediCore-Agentic, Personal Project

Topic: Secure Decision Support System for Clinical Workflows

Jan. 2026

- Autonomous agent capable of solving complex clinical data tasks e.g. **clinical guideline retrieval**, achieving a 30%+ score on the **GAIA Benchmark**.
- **Code-Action Agent** using **smolagents** and **Azure OpenAI (GPT-4o)** without relying on LLM intuition.
- Secure execution environment using **E2B sandboxing** which runs in an isolated VM to prevent data leakage.
- Custom **RAG** tools for clinical guideline retrieval and automated form-filling logic using vector-based retrieval.
- **SKILLS:** AI agent, RAG, E2B sandboxing, Azure OpenAI.

Car-ing is sharing, DataCamp

Topic: Analyse car review with LLM

Sep. 2025

- Use a variety of pre-trained HuggingFace LLMs to respond to the tasks using **prompt engineering**, e.g. classifying the sentiment in a car's text review, answering a customer question, summarizing or translating text, etc.
- **SKILLS:** Prompt Engineering, sentiment analysis, text summarization, question-answering, machine translation.

DigiNsure Inc., DataCamp

Topic: OCR for insurance classification

Aug. 2025

- Used **multi-modal learning** with text and scanned insurance documents to train an **Optical Character Recognition (OCR)** model for insurance classification.
- **SKILLS:** Multi-modal learning, Optical Character Recognition, sentiment analysis.

ABC_search, University of Helsinki

Topic: Coursework: Building NLP Application

Jan. 2024 – Mar. 2024

- Developed a search engine using **Boolean search**, **TF-IDF search** or **semantic search** with sBERT, retrieving medical text data from MedicalNewsToday.
- Applied **Agile methodologies** in the workflows and collaborated using **Git**.
- **SKILLS:** NLP, Git, Software Development.

WORK EXPERIENCE

Language Trampoline Project (KONE foundation), Helsinki, Finland

Consecutive Interpreter

Dec. 2025 - Current

- I worked as a freelance English-Chinese interpreter on the Language Trampoline project (KONE Foundation), where I met parents of non-Finnish-native toddlers and preschoolers and collaborated with the PhD researcher to develop language support strategies for their children.
- **SKILLS:** OPI interpretation, Cross-cultural Communication, Active Listening.

CERTIFICATES

HuggingFace: Fundamentals of Agents (Dec. 2025)

DataCamp: Associate AI Engineer for Data Scientists (Dec. 2025)

IBM: Python for Data Science, AI & Development (Aug. 2023)

Microsoft Certified: Azure AI Fundamentals (Jun. 2023)

TECHNICAL SKILLS

AI & Machine Learning: Large Language Models (LLMs), RAG Systems, PyTorch, HuggingFace, Scikit-learn, Reinforcement Learning (Gymnasium, RLHF)

Agentic framework: smolagents, LangChain, LlamaIndex

Data Engineering: Distributed Computing (PySpark), SQL (Snowflake/Supabase), ETL Pipelines

MLOps & Cloud: Docker, Kubernetes, Azure DevOps, GitHub Actions, CI/CD for ML (CML)

Software & Languages: Python, R, Shell Scripting, High-Performance Computing (HPC), JavaScript, HTML, CSS, LaTeX

Infrastructure: Git/GitLab/GitHub, Azure OpenAI Service

LANGUAGE SKILLS

Mandarin/Cantonese/Hakka: Native

English: Working proficiency

Japanese: Working proficiency

Finnish: Basic
