

# YIZHI (ETHAN) HAO

New York, NY | 914-882-6633 | [yh2222@cornell.edu](mailto:yh2222@cornell.edu) | <https://linkedin.com/in/ethan-yz-hao/> | <https://ethanhao.org/>

## EDUCATION

**Cornell University**, New York, NY May 2025

*Master of Science in Information Systems* | GPA: 4.065 | Merit Scholarship

**Coursework:** Algorithms and Data Structures, Web Development, Networked and Distributed Systems, Natural Language Processing, Applied Machine Learning, Computer Vision, Virtual and Augmented Reality, Cloud Computing and Big Data, Applied Data Science.

**Beijing Foreign Studies University**, Beijing, China Jun 2023

*Bachelor of Management in Information Systems and Information Management* | GPA: 3.93 | summa cum laude

**Honors/Awards:** Beijing Municipal Outstanding Graduate (Top 5%), Beijing Municipal Outstanding Thesis (Top 0.5%)

## TECHNICAL SKILLS

**Language/Software:** Python, Java, SQL, JavaScript, Go, HTML, CSS, C#, C, R, Shell, SPSS, Mplus, Stata, Git, Jira, Postman, Unity.

**Framework/Platform:** Next.js, React, Django, Django REST, SpringBoot, Spring Data JPA, Express, Vercel, AWS, GCP, Firebase.

**Database/Other Tools:** MySQL, PostgreSQL, Docker, Kubernetes, Redis, Spark, Hadoop, Kafka, NoSQL, PyTorch, Cloudflare, Linux.

## EXPERIENCE

**Rakuten, Rakuten Advertising, Software Engineer Intern**, New York City, United States Jun 2024 - Present

- Built **Go** plugin for **Kong** to log timestamped API usage, caching in **Redis** Sorted Set, aggregating 1M+ entries/day in **BigQuery**.
- Containerized **microservices** with **Docker**. Deployed to **Kubernetes (GKE)** using **Helm**, 50%+ cost saving from VM architecture.
- Automated **CI/CD** pipeline with **CircleCI**, managing image with **JFrog Artifactory** and secret with **Vault**, aiding agile development.

**Beijing Daxing International Airport Terminal, Administration Dept., Data Analyst Intern**, Beijing, China Oct 2022 - Jun 2023

- Developed a **queueing theory**-based system using **Python/MySQL** to forecast passenger check-in wait times, providing real-time estimates every 10 minutes and overnight backlog forecasts, covering 12,000+ flights, and 165,000+ passengers monthly.
- Implemented **multiprocessing** to speed up the integration of the **KDE** curve 4x during queue length estimation.

**Lenovo, Solution & Service Group, Data Analyst Intern**, Beijing, China Jul 2022 - Oct 2022

- Reviewed code, maintained Confluence document, and conducted testing of MARS (chipset inventory forecasting using **GLPK**), fixing bugs in linear programming constraints, and streamlining constraint matrix to reduce solution time by 40%.
- Developed optimization system demos for PCBA assembly planning and production scheduling using **Python/GLPK** and **Python/Geatpy**, solving multi-objective mixed integer programming problems with **genetic algorithm**.

## PROJECTS

**Zoning GPT** (Cornell × National Zoning Atlas), Research Assistant Fall 2024

Retrieval-Augmented Generation (**RAG**) pipeline to extract regulations from U.S. Zoning Codes, applying **LLMs** and **NLP** techniques.

- Used **AWS Textract** for OCR, **Elasticsearch** for indexing and keyword searching, **OpenAI GPT-4 API** for contextual extraction.

**MiniTorch**, Back-End Developer Fall 2024

**PyTorch** from scratch, with core features like computational graph, autograd, backpropagation, and **CUDA** support, PyTorch compatible.

**Trip Rover**, Technical Lead Spring 2024

Interactive trip planning application utilizing **Google Map Javascript API** with **vis.gl**, integrating **Places/Autocomplete/Routes APIs**.

- React** Front end with **Chakra UI** and **react-beautiful-dnd** for interactive components. Manage state via **Redux/Context API**.
- Utilized **Java SpringBoot** as the backend framework, and **Spring Data JPA** as **ORM** framework to manage the database.

**Issue Tracker**, Full-Stack Developer Spring 2024

Issue tracking system allowing users to post and update issues in markdown format, built with **Next.js**, and deployed using **Vercel**.

- Used **Radix UI** and **tailwind** for front-end component, **React-Hook-Form** for form handling, **Zod** for schema validation, and **Prisma** as **ORM** framework for **CRUD** operations. Built **RESTful** API and used **Axios** and **TanStack Query** for data fetching.
- Configured **AWS EC2** to deploy **MySQL** database using **Docker**. Used **Sentry** for performance monitoring and error-tracking.
- Integrated **OAuth2.0** with **JSON Web Token**, delivering third-party login and optimizing user-login experience.

**Equation OCR**, Full-Stack Developer Spring 2024

OCR application for converting equations into **LaTeX** code using **LLMs** and **NLP** techniques, specifically with **OpenAI GPT-4v API**.

**StudyBud**, Full-Stack Developer Fall 2023

Social Learning Platform developed with **Django** and productionized via multi-container **Docker compose** on **AWS EC2**.

- Implemented Django view and template mechanics for user interface. Integrated **PostgreSQL** for database operations
- Utilized **Django REST Framework** for API communication. Configure **Nginx** to serve static files efficiently.

**Identifying Supply Chain Risk Based on Agent-Based Modeling and Simulation** (BFSU), Research Assistant Fall 2023

Research implementing **ABMS** method to identify high-risk nodes in SCs, using **Taguchi** method to study factors affecting SC resilience.

- Utilized **Agentpy** and **Networkx** to construct Python modules simulating the firm's behaviors during SC disruption.
- Created a **parallel testing** process, achieving a 24x speed increase. Improved **MySQL** schema, reducing data redundancy by 60%.

## PUBLICATIONS

Ma, Xiaoyu; **Hao, Yizhi**; Li, Xiao; Liu, Jun; Qi, Jiasen (2023): Evaluating global intelligence innovation: An index based on machine learning methods. In *Technological Forecasting and Social Change* 194, p. 122736. DOI: [10.1016/j.techfore.2023.122736](https://doi.org/10.1016/j.techfore.2023.122736).