Jinglin (Ollie) Jian

(+1) 217-819-7884   jj50@illinois.edu   jianjinglin.github.io   https://www.linkedin.com/in/olliejian1999/

EDUCATION

University of Illinois Urbana-Champaign, School of Information Sciences Illinois, US
M.S. in Information Sciences/Management, GPA: 4.0/4.0 *Aug 2023 - Dec 2024*
Course: Large Language Models, Machine Learning, Text Mining, Database, Data Mining, Data Warehousing, Sociotechnical Info Systems, Independent Study

Peking University, National School of Development Beijing, China
B.Econ. (Double Major) *Sep 2021 - Jul 2023*

Beijing Normal University, Faculty of Education Beijing, China
B.S. in Educational Technology, GPA: 3.76/4.0 *Sep 2017 - Jul 2021*
Course: Data Structure, Database, Software Engineering, WebDev, Computer Network, Intelligent System

PUBLICATIONS

[1] Haoyang Liu, Yijiang Li, **Jinglin Jian**, Yuxuan Cheng, Jianrong Lu, Shuyi Guo, Jinglei Zhu, Mianchen Zhang, Miantong Zhang, and Haohan Wang. **Toward a Team of AI-made Scientists for Scientific Discovery from Gene Expression Data.** arXiv:2402.12391 (2024).[Paper] [Code]

RESEARCH EXPERIENCE

Distributed Aptamer-Protein Binding Computation System *Jan 2024 – Present*
Advisor: Professor Yang Zhang (UIUC) and Dr. Jin Chen (Cleveland Clinic)

– Built a high-concurrency, fault-tolerant system on **AWS** for aptamer-protein binding, employing RESTful API (**AWS ApiGateway**), computation and load balancing (**EC2**) and parallel processing (**MPI**).

– Utilized **AWS RDS** (task state management), **AWS S3** (file storage), and Docker (containerization & computation task modularization).

– Managed distribution of heavy computation tasks across stateless **EC2 workers** for asynchronous processing and implemented a **Cron Job** for monitoring and recovery of siloed tasks.

– Achieved a significant reduction in processing time (from 583h to 48h) and cost by optimizing resource scaling and usage, resulting in a 92% decrease in time and enhanced cost efficiency.

Team of AI-made Scientists (TAIS) [Paper] [Code] *Aug 2023 – Feb 2024*
Advisor: Prof. Haohan Wang (UIUC)

– Used **Entrez API** to fetch **GEO data**, stored meta-/raw-data in **HDF5 format** for **efficient I/O**.

– Streamlined analysis via aligning gene symbols (**mygene lib.**), merging data (**pd.merge**), auto-logging (**JSON**), statistical corrections (**multipletests** for p-value), and ML (**LASSO**, **linear mixed model**).

– Created an **Action** class for autonomous code refinement, via creation (**template-based prompting**), execution (**subprocess**), outputs/errors capture (**logger**), and feedback (**Code Reviewer** agent).

– Contributed **TAIS**, a **multi-agent system** pioneering LLM-based scientific discovery, benchmarked against **gold standard** (human-made) and AI systems (**GPT-4/MetaGPT/AutoGPT**).

Semi-automatic Knowledge Graph Construction [Web] *Sep 2020 - Jul 2021*
Advisor: Prof. Qinhua Zheng (Beijing Normal University)

– Built a script for video-to-text transcription (NetEase API), designed metadata schema, and organized multi-person text annotation (**BIO tagging**).

– Iterated a supervised **BiLSTM-CRF ML model** for **entity recognition** with expert feedback and dynamic term re-ranking (using **cross-entropy**), improving F1-score from 0.54 to **0.76**.

– Implemented entity vectorization (**Word2Vec**), text segmentation (**sliding window**), Relation Extraction (template-based), and stored knowledge graphs (**Neo4j**).

Using Hypervideo to Facilitate Online Interactions [Paper]   *Sep 2019 – Aug 2020*
Advisor: Prof. Jingjing Zhang (Beijing Normal University)

– Developed an **MVP** for an online learning platform featuring video streaming and real-time commenting, using **React** (frontend), **Node.js/Express** (backend), and **flv.js** for scalable video playback.
– Developed a real-time commenting feature with storage (**MongoDB**)/interactions (**WebSocket**) and enhanced it by clustering comments to **hotspot video timestamps**.
– Built a user **behavior tracking** mechanism with a **RESTful API** (endpoints receiving **JSON**), trackable elements (in **React**), HTTP (**axios**), routing (**Express**), and data storage (**MongoDB**).
– Assisted in **pre-test** and **post-test** experiments, and contributed to data-driven behavior pattern visualization (**Gephi**), resulting in a publication.

PROFESSIONAL EXPERIENCE

Research Assistant   *Aug 2023 – Present*
Supervisor: Prof. Mackenzie Alston, University of Illinois Urbana-Champaign

– Conducted literature review (**randomized controlled trials**) using **Zotero**.
– Scraped Google Scholar (**Selenium**) extracted 2,000+ email URLs (**re**).

Head on Online Learning Department   *May 2019 – Aug 2022*
China Starry Night (NGO / Startup) [Web]

– Reported directly to the CTO; scaled the technical team from 5 to 30, achieving a 22k viewer increase on our Bilibili channel within 4 months.
– Secured many **national award (Top 1%)** and presented at the 5th China Education Innovation Expo.

SELECTED PROJECTS

– **Demo of ChemTutor: AI Q&A system with Inorganic Chemistry Textbooks** [Code]   *2024*
  `LLM` `Q&A system` `Question a 700-page Textbook` `LangChain`

– **HMM-DRL Model for Data-driven Auto-Trading** [PDF]   *2022*
  `Reinforcement Learning` `Hidden Markov Model` `Time Series Data` `Financial Index`

– **Evolution of Key Themes in Learning Sciences** [Web]   *2020*
  `Topic Model: LDA` `TF-IDF` `Text Mining` `Data Visualization`

HONORS AND AWARDS

– **National Innovation and Entrepreneurship Training Program**, Ministry of Education   *2021*
– **Jianghaiziqiang Scholarship (1%)**, Beijing Normal University   *2020*
– **First-class Scholarship for Competition Excellence (1%)**, Beijing Normal University   *2019*
– **Outstanding for Academic Excellence**, Beijing Normal University   *2017 - 2021*

SKILL SET

Machine Learning & NLP	LangChain, TensorFlow, PyTorch, SciPy, sklearn, gensim, NLTK
Programming Language	Python, Java, C, JavaScript, HTML, CSS, Matlab, Stata
Framework & Database	React, Node.js, RESTful API, MySQL, MongoDB, Neo4j
Cloud	AWS - EC2, S3, Lambda FaaS, API Gateway
Code Management & Others	Git/Github, Docker, Tableau, Unit Testing, LaTeX