Chelsea (Yuqing) Wu
+1 773-392-6893 | wuyuqing0327@gmail.com | https://github.com/wuyuqing0327

PUBLICATIONS

- Jiajun Luo, **Yuqing Wu**, *et al.* "The Role of Oral Microbiome in Circulating Metabolic Biomarkers and the Influence of Air Pollution." *Oral Diseases*, **January 2026**.
- Hui Yi Leong*, **Yuqing Wu*,** *et al.* "AMAS: Adaptively Determining Communication Topology for LLM-based Multi-agent System." **EMNLP of 2025: Industry Track, 2025**. (***Co-first authors***)

RESEARCH EXPERIENCES

University of Illinois Chicago **Chicago, IL, USA**

Research Area: Causal Reinforcement Learning, Multi-Agent Learning, Network Structure Inference, LLMs *Jan 2026 – Present*

- Construct user-level interaction/exposure signals from social context (e.g., neighbor/peer information flow), using LLM-based text representations to study how network influence relates to **stance switching**.
- Define **RL-style objectives for stance dynamics** (e.g., stability vs. switching, cost/constraints) to support **counterfactual evaluation of sequential interventions**.
- Model users as agents in a networked **multi-agent setting**, where each agent's behavior can be influenced by neighbors, aligning the task with **causal RL under spillovers**.

Northwestern University (On-site) **Chicago, IL, USA**

Researcher (Network Analysis, NLP, LLMs) *May 2025 – Dec 2025*

- Conducting NLP-based research to **map knowledge gaps in scientific fields across time** (e.g., shifts in brain/DNA/RNA research focus between 2019 and 2020), supporting researchers in identifying underexplored directions.
- Fine-tuned **SciBERT** on article sections to generate embeddings and built **graph networks** of papers; extracted structural features such as community clusters (Louvain), centrality, and connectivity to analyze field evolution.
- Generated **concept evolution trees** from section-level keywords identified via **BERTopic**, capturing how concepts emerge, expand, or decline over time within a field.
- Integrated graph features and evolution-tree features into an **FCI causal discovery model** to identify which structural and conceptual factors most directly explain shifts and gaps in research domains.
- Co-author of "Measuring Scientific Novelty as Network Rewiring." Under review in ***NetSci***, 2026.

Greaten AI (Engineering and Research Team, Part-time, Remote) **Albany, CA, USA**

Software Development Engineer (Multi-Agent, Reinforcement Learning, LLMs) *Dec 2024 – May 2025*

- Developed an **LLM-based multi-agent system** that **chooses the best agent collaboration structure per query** (dynamic communication topology), enabled by a **LoRA-tuned graph selector**.
- Implemented and evaluated **RL-based topology optimization** (actor–critic / policy-gradient training) to generate high-performing candidate graphs; benchmarked across multiple LLM backbones (e.g., GPT-3.5, LLaMA-3, DeepSeek-R1 distilled) and tasks including QA, reasoning, and code generation.
- **Co-first author** of an **EMNLP 2025** paper on adaptive topology selection for LLM multi-agent systems.
- Built and deployed production services for a zoning-ordinance **document Q&A** and **voice agents** based on AMAS, including PDF parsing (text/tables/figures), retrieval + generation APIs, and AWS-based telephony workflows (Connect/Lambda + ASR + email automation), with containerized deployment and monitoring for low-latency reliability.

University of Chicago (Part&Full-time, On-site) **Chicago, IL, USA**

Research Assistant (Bioscience, Causal Analysis, Biomarker Analysis) *Jan 2024 – Feb 2025*

- **Investigated causal relationships** between environmental exposures, gut microbiome composition, and biomarkers linked to **cardiometabolic diseases** using matched observational data.
- Processed and aggregated large-scale user-level microbiome, biomarker, and air pollution data; **implemented propensity score matching to reduce confounding and selection bias**.
- Built **regression-based mediation models** to analyze the **direct and indirect effects** of PM2.5 exposure on biomarkers via microbiome mediators, **incorporating FDR correction and bootstrap resampling** to ensure robustness.
- **Visualized** mediation effects using **chord and alluvial plots**; used heatmaps, bar plots, and boxplots to display microbiome distribution, biomarker significance, and conjunct effects.
- Co-authored a manuscript on PM2.5–microbiome–biomarker pathways, showing associations between air pollution exposure and higher Insulin/Resistin and lower Ghrelin— key biomarkers associated with type 2 diabetes, cardiometabolic risk, and appetite regulation; published in ***Oral Diseases*** (2026).

WORK EXPERIENCES

OOIN Media (Full-time, Remote) **Long Island, NY, USA**

SDE-AI Agent Engineer *Aug 2025 – Jan 2026*

➢ Creator Campaign Outreach

- Designed a **LangGraph-based multi-agent architecture** (Demand Agent / Plan Agent / Execute Agent) with tool-calling, memory/state, guardrails, and workflow orchestration; integrated multi-LLM routing across **GPT-4/5, Gemini, DeepSeek** based on task type and cost/latency constraints.
- Launched an end-to-end **merchant–creator collaboration automation pipeline** that covers outreach → multi-turn negotiation (budget/pricing rules) → sample request & shipping → draft reminders/review → publishing to social platforms, enabling scalable "creator BD agent" capabilities for merchants.
- Engineered **high-throughput async execution** for campaign-scale workloads with reliable retries and idempotency; implemented **Redis-based ticket/state management** to maintain low-latency multi-turn conversations and consistent workflow progression.
- Implemented an **algorithmic influencer matching & ranking service** using creator profile + engagement features; combined rules + model-driven scoring to optimize brand–creator pairing and improve campaign relevance.
- Added a **feedback learning loop** by logging interaction outcomes (e.g., negotiation results, drop-off points, review feedback) to iteratively refine prompts, policies, and routing rules for higher task success rates.

➢ Automated account nurturing

- Built an **intent-to-action agent pipeline**: a "Business Analyzer" interprets user prompts, a planning agent decomposes goals into executable steps, and an execution agent coordinates with **MCP** to run actions (tool back-and-forth), with robust error handling and step-level validation.
- Engineered **low-latency, high-throughput orchestration** under concurrent nurturing demands via **async task scheduling + queue-based execution**, per-account **rate limiting / concurrency caps**, and **Redis-backed state checkpoints** to resume workflows without re-planning; reduced tail latency by caching reusable context and short-circuiting no-op steps.
- Integrated **GPT-4o** for instruction-following and action planning, enabling reliable automation flows that translate high-level growth intents into repeatable operational behaviors.

➢ Prompt-to-Viral Video Generation

- Built a **trend mining + feature extraction pipeline** to analyze viral videos and distill reusable patterns (hooks, pacing, structure, keywords), injecting them into prompts as controllable constraints for content generation.
- Developed **VISTA-inspired agent framework** with self-improvement loops (critique/review agents + iterative prompt refinement) to address quality issues such as "not smooth" videos; added automated review gates that generate actionable optimization advice before re-generation.
- Optimized generation performance by integrating **Qwen** and applying **KV cache** strategies to reduce inference latency/cost for iterative agent loops; supported deep content analysis to improve consistency and adherence to promotional intent.

Hillhouse Capital (Big Health, Full-time, On-site) **Beijing,China**
Senior Machine Learning Engineer *May 2022 – Jul 2023*

➢ In-app Medicine Recommendation System

- Spearheaded the development of an **in-app recommendation system** by building a **Two-Tower Neural Network model**, resulting in a 2% increase in user conversion rate through **A/B testing**.
- One tower of the model leveraged **user basic features and item latent factors** derived from **matrix factorization** within **DeepFM**, capturing both **low- and high-order feature interactions**. One tower concatenates the most frequent and **latest item-content embeddings** from a **pre-trained GPT-3 model** for medicine instructions into a single vector, which is then processed **through dense layers** to enable **content-based filtering** aligned with individual health profiles.
- Deployed **data pipeline with Spark for real-time recommendations**, serving results via **API, supporting 1000 requests per minutes**.

DiDi (User Growth Department, Full-time, On-site) **Beijing, China**
Senior Machine Learning Engineer *Sep 2020 – May 2022*

➢ Intelligent Campaign for Retaining User Engagement and Revenue

- Utilized previous observed **multi-arm quasi-RCT samples** to develop uplift model, applying **CausalRandomForest with different criteria** and tried **Double-ML** to estimate the **ATE** and **CATE**, benchmarking these results against ATE estimates from **Difference-in-Differences** analysis to evaluate coupon effectiveness.
- Utilized **convex optimization to optimize the ROI** formula based on uplift score with budget constraints to determine coupon allocation for each promotion week, which led to a 20% increase in adoption rates and GMV, and achieved an 85% reduction in costs with a 4-6% improvement in ROI.
- Developed a Python package to automate the causal analysis and model deployment process for modeling functional team.

FinUp (Credit Risk Department, Full-time, On-site) **Beijing, China**
Model Analyst *Jan 2018 – Sep 2020*

- Developed an integrated credit risk and fraud model framework using **XGBoost, Logistic Regression** for pre-loan approval, **Seq2Seq/RNN** for in-loan management and post-loan collection, analyzing AUC, KS, recall, precision, and monotonicity by each decile each segment.
- Launched **decision-making strategies** to refine **loan lifecycle**, including customer credit lines, interest rates, and loan terms based on model score, resulting in a 15% reduction in overdue rates for the 18% APR product.

EDUCATION

University of Illinois Chicago, *Doctoral of Computer Science* Jan 2026
University of Chicago, *MS in Applied Data Science, GPA 4.0* *Sep 2023 – Dec 2024*
Yanshan University*, Master in Engineering* *Sep 2012 – Jul 2015*
Shenyang University of Chemical Technology, *Bachelor in Engineering* *Sep 2007 – Jul 2011*

SKILLS

- **Programming**: Python (TensorFlow, PyTorch), R, SQL, Scala, Java
- **Full-Stack Development**: Frontend (Java, HTML, React, Vite, JavaScript); Bankend (Python, Java, HTTP, TCP)
- **Dev Tools & Platforms:** Spark(PySpark), Hive, HDFS, Docker, RESTful APIs, Flask/FastAPI, Git, Linux, GCP, AWS, CRM system
- **Machine Learning/Deep Learning**: LLM, AI-Agent, Generative AI, RL, RAG, NLP, Recommendation Systems, Causal Inference, Optimization Techniques, Anomaly Detection, Multi-task Learning
- **Analytics:** Product analytics, Marketing strategy, Risk analysis, Supply-demand optimization, Healthcare management
- **Visualization**: Tableau, Power BI, Superset