YUNYA (YOYO) LIN

Champaign, IL | +1 217-791-3992 | yoyolin2@illinois.edu
linkedin.com/in/yoyo-lin-631889299 | github.com/YoyoLin008

EDUCATION

B.S. in Information Sciences + Data Science, Minor in Computer Science
University of Illinois Urbana-Champaign Expected Dec. 2026
GPA: 4.00/4.00 | Dean's List: all completed semesters
Relevant coursework: Probability, Statistics, Linear Algebra, Calculus I–III, Data Structures, Database Systems, Machine Learning, Data Visualization, Information Modeling, and Data Management.

RESEARCH EXPERIENCE

Undergraduate Researcher | Evaluating LLMs as Peer Reviewers
UIUC School of Information Sciences Nov. 2025 - May 2026
Advisor: Prof. Nitin Verma

- Investigated whether fine-tuning can reduce overly positive manuscript acceptance recommendations from large language models and support more selective peer-review judgments.
- Implemented fine-tuning and modified Python and API workflows in the research codebase to generate and inspect peer-review outputs across OpenAI models.
- Organized model outputs and repeated experiments to support the team's comparison of model recommendations with human judgments and analysis of acceptance tendencies.

Undergraduate Researcher | AI-Assisted Conceptual Modeling
UIUC School of Information Sciences Fall 2026–Present
Supervisors: Prof. Bertram Ludäscher and Yuanxi Fu

- Compare AI-generated and human-created conceptual models on the same topic: funding opportunities for university researchers.
- Examine how the models represent domain concepts and relationships to understand differences between AI-assisted and human modeling.

Visiting Research Assistant | BIRD-OnChain
CUHK-Shenzhen / Shenzhen Institute of Big Data May–Jul. 2026
Supervisor: Prof. Chenhao Ma

- Collected and labeled blockchain transaction data under domain-expert guidance to distinguish user intents, including closing positions and mitigating risk, for the BIRD-OnChain benchmark.
- Validated SQL queries and on-chain records retrieved through Dune, checking whether query results accurately represented the underlying transaction data.

PROFESSIONAL EXPERIENCE

Quantitative Research Intern
Ipsos | *Guangzhou, China* Jul. 2026–Oct. 2026

- Prepare and validate Chinese consumer research data for a Fortune Global 500 automotive client, covering brand perceptions, buyer age distributions, sales, and regional differences.
- Use Python and SQL to support market analysis and summarize patterns for internal research discussions on the client's positioning in China.

SELECTED DATA SCIENCE PROJECTS

Chicago Public Health and Community Violence Analysis

University of Illinois Urbana-Champaign Jan.–May 2026

- Analyzed associations between health and socioeconomic indicators and gun-violence victim counts across 77 Chicago community areas (2005–2010), using a team-developed Python workflow.
- Led interpretation, figure selection, and report writing; identified teen-birth rate as the strongest selected correlate of mean annual victim counts (r = 0.719), while distinguishing association from causation.

Crop Recommendation System

Illinois Data Science Club Jan.–May 2025

- Compared Random Forest, XGBoost, LightGBM, SVM, KNN, and MLP classifiers to recommend crops from soil and environmental features.
- In the team's six-model comparison, Random Forest achieved the highest accuracy on a 440-sample held-out test set: 99.32% accuracy and 0.993 weighted F1 with an 80/20 train–test split.

PUBLICATIONS AND MANUSCRIPTS

Lin, Y. (2026). When Evidence Sets Become Relevance Lists: A Controlled Audit of Scientific Retrieval Evaluation. *Submitted to Information Retrieval Research (IRRJ), October 3, 2026.*
Manuscript and reproducibility materials (v1.1.1): Zenodo DOI: 10.5281/zenodo.23126392 | GitHub release

Lin, Y. (2023). Domestic pension industry: Trends, themes and research trends—Visual analysis based on CiteSpace knowledge graph. *Advances in Economics and Management Research, 7*(1), 677–689.
https://doi.org/10.56028/aemr.7.1.677.2023

LEADERSHIP

Project Lead

Illinois Data Science Club 2025–Present

- Coordinate an eight-week Data Dive across 10 technical teams and 50+ students, including project milestones and final presentations to industry judges from Caterpillar and P&G. The club receives semester sponsorship from John Deere.
- Mentor teams in Python and data analysis, review GitHub repositories, and guide project scoping, troubleshooting, documentation, and presentations.
- Progressed from Data Dive participant to Project Lead over four semesters of club involvement.

TECHNICAL SKILLS

Programming and libraries: Python (pandas, NumPy, scikit-learn, XGBoost, LightGBM), SQL, R, C++
Databases and tools: Neo4j, Cypher, Git/GitHub, OpenRefine

HONORS AND ADDITIONAL INFORMATION

Honors: Illinois for Illinois Scholarship (2025–2026)
Certification: IBM Data Science Professional Certificate (2025)
Languages: Mandarin Chinese (native); English (native)