

Yidan Sun

Updated September 7, 2026

Email: yidans@usc.edu

Homepage: yidans.github.io

LinkedIn: linkedin.com/in/yidansun07

Research Group: Artificial Intelligence and Complex Systems (AICS) (Website: aicomplex.github.io)

Research Interests

Network Science, Complex Systems, Computational Social Science, Text Analytics, Statistical Modeling of Network Data

Education

University of Southern California

Los Angeles, CA, USA

Ph.D. Student, Industrial and Systems Engineering

Aug. 2023 - Present

MSc, Analytics

Aug. 2021 – Dec. 2022

Advisor: Mayank Kejriwal

Key Courses: Design of Experiments and Quality Engineering, Stochastic Processes, Decision Analysis, Reinforcement Learning, Linear Optimization, Stochastic Optimization, Data Management, Data Mining, Text Analytics

The Chinese University of Hong Kong, Shenzhen

Shenzhen, China

BSc, Statistics

Sep. 2016 – May 2020

Key Courses: Probability and Statistics, Regression Analysis, Statistical Inference, Survey Sampling, Nonparametric Statistics, Categorical Data Analysis, Statistical Modelling in Financial Markets, Stochastic Calculus

Awards

Jenny Wang Excellence in Teaching Award (USC)

2024

ISE Department Outstanding Teaching Assistant of the Year (USC)

2024

Publications

Sun, Y., & Kejriwal, M. (2025). Modeling complex associations between climate change policies and sociodemographic factors using conditional inference trees. Under review at *Humanities and Social Sciences Communications*.

Sun, Y., He, D., & Kejriwal, M. (2026). Examining persistent inequities in financial complaint resolution for older Americans and veterans in the United States. *PNAS Nexus*.

Sun, Y. (2025). Scalable estimation of exponential random graph models on large networks. *The Web Conference (TheWebConf)*.

Gawin, C., Sun, Y., & Kejriwal, M. (2025). Navigating semantic relations: Challenges for large language models in abstract common-sense reasoning. *The Web Conference (TheWebConf)*.

Sun, Y., & Kejriwal, M. (2025). Backlash or reinforcement? Donald Trump's 2017 inauguration and shifting climate beliefs in the United States. *NPJ Climate Action*.

Sun, Y., & Kejriwal, M. (2024). A study of firm-switching of inventors in Big Tech using public patent data. Chapter 12 in *Springer Social Network Analysis and Mining Applications in Healthcare and Anomaly Detection*.

Sun, Y., & Kejriwal, M. (2023). A structural study of Big Tech firm-switching of inventors in the post-recession era. *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 670–677.

Sun, Y., & Kejriwal, M. (2023). DeepGraph: Multi-cluster interactive visualization of complex networks in a learned representation space. *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*, 427–430.

Under Review

Patil, V., Bacha, S. V., Yamani, R., Sun, Y., & Kejriwal, M. (2026). Beyond the star rating: LLM-assisted aspect analysis and large-scale text classification of 4.7 million restaurant reviews. Revised manuscript under review at *PLOS ONE*.

Sun, Y., & Kejriwal, M. (2026). Online sex-work advertising on Backpage in Southeast Asia, 2013–2018: A retrospective analysis of posting, place, and price. Revised manuscript under review at *SN Social Sciences*.

Kejriwal, M., Tang, Z., Nananukul, N., Aravindan, A. V., & Sun, Y. (2026). A hybrid-AI architecture for rapid autonomous scientific discovery. Submitted to *AGI-26*.

Kejriwal, M., & Sun, Y. (2026). Building discoverative AI ecosystems for supporting high-quality agentic science. Submitted to *AIAS+ 2026*.

Sun, Y., & Kejriwal, M. (2026). Hallucinated references in AI-generated publications: A cross-corpus audit. Submitted to *AIAS+ 2026*.

Kejriwal, M., Sun, Y., et al. (2026). Bootstrapping scientific discovery: A case study on AI-driven research workflows in a research group setting. Submitted to *CHI 2026 Posters*.

Sun, Y., et al. (2026). Computational social science for harm reduction: Privacy-preserving analysis of market structure in counter-trafficking work. Submitted to the *AAAI 2027 Artificial Intelligence for Social Impact Track*.

Under Preparation

Sun, Y., & Kejriwal, M. (2026). Interactive exploration of tie-formation hypotheses in complex networks. Manuscript in preparation (demonstration paper).

Sun, Y., & Kejriwal, M. (2026). Coordination and fragmentation in Latin American and Caribbean online sex work markets. Manuscript in preparation.

Sun, Y., & Kejriwal, M. (2026). A framework for organized control at the account level in online sex trafficking detection. Manuscript in preparation.

Presentations

Poster Presentation, 'Building reliable PhD-level agents for computational research', Agentic AI Summit, University of California, Berkeley, Berkeley, CA, August 1–2, 2026.

Poster Presentation, 'Mapping collaboration in AI-driven firms: A co-inventor dataset (2020–2025)', International Conference of the Computational Social Science Society of the Americas (CSSSA), Washington, D.C., November 6–9, 2025.

Poster Presentation, 'Reviewer rating variability, reviewer confidence, and language model-based sentiment prediction of machine learning conference submissions', 10th International Congress on Peer Review and Scientific Publication, Chicago, September 3–5, 2025.

Oral Presentation, 'Collaboration diversity and paper success: A network analysis of co-authorship in a major machine learning conference', INSNA Sunbelt Conference, Paris, June 23–29, 2025.

Poster Presentation, 'Scalable estimation of exponential random graph models on large networks', The Web Conference (TheWebConf), May 13, 2025.

Oral Presentation, 'Navigating semantic relations: Challenges for large language models in abstract common-sense reasoning', The Web Conference (TheWebConf), May 13, 2025.

Oral Presentation, 'AI reveals systemic financial inequities: A study of CFPB complaints for older adults and veterans', USC ShowCAIS Symposium, April 2025.

Oral Presentation, 'Understanding the drivers of human trafficking using massive multi-partite network modeling and analytics', USC Cloudwalkers Premiere, March 2025.

Poster Presentation, 'Modelling campaign finance data using network science to map structures and flows of influence', Southern California Applied Mathematics Symposium (SOCAMS), April 22, 2023.

Oral Presentation, 'A structural study of Big Tech firm-switching of inventors in the post-recession era', Advances in Social Networks Analysis and Mining (ASONAM), November 7, 2023.

Oral Presentation, 'DeepGraph: Multi-cluster interactive visualization of complex networks in a learned representation space', Advances in Social Networks Analysis and Mining (ASONAM), November 6, 2023.

Media Coverage Public Writing:

'5 lessons from 20 years of climate surveys: Political party matters, but it doesn't explain everything', *The Conversation*, July 21, 2026. Syndicated by *The Washington Post*, *NationofChange*, and *ArcaMax*.

Research Coverage:

‘A majority of Americans don’t actually fit either climate camp’, *Anthropocene Magazine*, August 4, 2026.

‘Financial complaint delays hit seniors and veterans hardest, with gaps widening over time’, *Phys.org*, April 16, 2026.

‘Older Americans and veterans wait longer for financial justice’, *USC Information Sciences Institute*, April 16, 2026.

‘Big Tech inventor mobility leads to higher productivity’, *USC Information Sciences Institute*, November 7, 2023.

Teaching Experience
University of Southern California (Teaching Assistant)

ISE 547: Applied Generative Artificial Intelligence for Enterprises Fall 2025

ISE 529: Predictive Analytics Fall 2025

ISE 561: Economic Analysis of Engineering Projects Summer 2025

ISE 540: Text Analytics Fall 2023, Spring 2025, Summer 2025

Mentoring

John Lim (REU Summer program; Northwestern University) Summer 2023

Yining Ma (Directed Research; USC) Spring 2024

Revanth Yamani (Directed Research; USC) Spring 2024

Shree Vaishnavi Bacha (Directed Research; USC) Spring 2024 - Present

Vishal Patil (Directed Research; USC) Spring 2024 - Present

Daiqi He (Directed Research; USC) Summer 2024 - Present

Herry Wang (Directed Research; USC) Summer 2024

Erica Okeh (Viterbi SURE Program; Howard University) Summer 2024

Richard Joungyoon Kim (CURVE Program; USC) Fall 2024 - Present

Angel Otters (Directed Research; USC) Fall 2024

Professional Service**Journal Reviewer for:**

Social Network Analysis and Mining (Springer)

Frontiers in Political Science

Frontiers in Sociology

Conference Reviewer for:

Conference on Empirical Methods in Natural Language Processing (EMNLP), 2026

Industry Experience

GRAIL Los Angeles, CA

Research Scientist May 2026 – Aug. 2026

Developed LLM-based agentic systems for research workflows. Built an autonomous scientist pipeline that decomposes research questions into executable steps, orchestrates tool calls across retrieval, analysis, and code execution, and iterates on intermediate results with human-in-the-loop checkpoints. Designed a retrieval-augmented grant recommender that matches investigator profiles and project abstracts against funding-opportunity corpora, combining dense retrieval with LLM-based relevance ranking and structured eligibility filtering. Implemented a hallucination checker that audits generated scientific text against source documents and bibliographic databases, verifying citation existence, attribution accuracy, and claim–evidence grounding; developed the accompanying evaluation harness and error taxonomy.

USC Information Sciences Institute

Research Engineer

Marina del Rey, CA

Jan. 2023 – Aug. 2023

Conducted applied research in the Artificial Intelligence and Complex Systems group; designed and ran controlled experiments, implemented model variants and evaluation pipelines, and contributed to academic publications through manuscript drafting and preparation of experimental results.

Datago Technology Limited

Quantitative Researcher Intern

Hong Kong

May 2021 – Sep 2021

Built data pipelines to systematically extract and standardize the text features required by sentiment-based factors developed by senior researchers; implemented factor construction logic from raw forum data; ran signal diagnostics and walk-forward backtests to evaluate predictive strength, decay, and out-of-sample performance.

Ping An Insurance

Data Analyst Intern

Shenzhen, China

Jun. 2019 – Sep. 2019

Assisted in designing and executing A/B tests for a new version of the online banking portal; analyzed key performance indicators, checked sample balance, and prepared short diagnostic reports to support product decisions.

Guosen Securities

Algorithmic Trader Intern

Shenzhen, China

Jun. 2018–Sep. 2018

Assisted in automating and optimizing intraday trading strategies, including signal preprocessing, rule refinement, and basic execution tuning; supported backtesting and parameter checks to improve stability and reduce latency-related errors.

Technical Skills **Machine Learning & AI:** Supervised and Unsupervised Learning, Deep Learning, Large Language Models, Graph-Based Learning
LLM & Agentic Systems: Prompt Engineering, Retrieval-Augmented Generation (RAG), Multi-Agent Coordination, LLM Orchestration, Evaluation Pipelines
Network Science & Graph Analytics: Graph Theory, Community Detection, Centrality Analysis, Graph Embeddings, Structural Modeling
Natural Language Processing: Embedding Models, Semantic Search, Text Classification, Sequence Labeling, Topic Modeling
Programming & Tools: Python (PyTorch, TensorFlow, Scikit-Learn, Transformers), R, Julia, SQL, JavaScript (React.js, D3.js)
Data Engineering & Storage: PostgreSQL, Neo4j, Spark, Snowflake
Data Visualization: Matplotlib, Plotly, Seaborn, Altair, Tableau, Gephi
Cloud & Systems: AWS, Linux, Distributed Computing, Workflow Automation