

Xiao Chen

Redmond, WA | andrewchenxiao@gmail.com | 315-395-5248
Google Scholar | ORCID: 0009-0000-8828-7330 | github.com/Xiao-Chen-usc

RESEARCH PROFILE

Applied AI and statistical-systems practitioner with 3.5+ years at Amazon Health building and evaluating production multimodal and multi-agent decision-support systems for regulated healthcare workflows. Experience spans image and text understanding, specialist-agent orchestration, expert-grounded evaluation, human-in-the-loop safeguards, and large-scale statistical forecasting. Coauthor of a preprint and five manuscripts under review that use multilevel models and interpretable machine learning to study psychosocial, educational, and behavioral outcomes. Research interests center on reliable AI evaluation, human-AI collaboration, quantitative methods, and real-world decision systems.

RESEARCH INTERESTS

Multimodal and multi-agent AI; reliable AI evaluation; statistical modeling; human-AI collaboration; measurement and assessment; clinical decision support; interpretable machine learning; multilevel and observational methods; model calibration and uncertainty.

INDUSTRY RESEARCH AND ENGINEERING EXPERIENCE

Software Development Engineer, Amazon Health | Seattle, WA Sep 2022 - Mar 2026

Multimodal, Multi-Agent Pharmacy Verification Copilot | AWS Bedrock / Claude, Lambda, Python

- Designed and shipped a real-time multimodal, multi-agent decision-support pipeline processing 3,000+ prescription images per day at under 800 ms average latency, integrating image, extracted text, order, and database evidence in a human-in-the-loop pharmacist workflow.
- Evolved the architecture from a single model to a single agent and then coordinated specialist agents for prescription-rationality, image-to-database, and label-to-database checks; structured cross-checks reduced hallucination risk before pharmacist review.
- Worked with applied scientists and pharmacists to define agent- and system-level evaluation criteria, validate against pharmacist-reviewed cases, and design false-positive handling, gating, escalation, and traceable controls; achieved 92% pre-screening accuracy while preserving pharmacists' final authority.
- Reduced manual verification workload by 95% and increased pharmacist throughput 4x.
- Won first place in an internal hackathon among 50+ teams.

Pharmacy Verification Platform | React, Java / Spring Boot, AWS ECS, DynamoDB

- Architected a full-stack verification platform serving 200+ internal users at 99.9% uptime and handling 10,000+ daily prescription review requests.
- Designed APIs, review interfaces, and backend workflow orchestration for structured human decision-making, traceable verification, and auditable review states.
- Structured review outcomes and system events to support operational monitoring and investigation of exceptions.
- Conducted iterative workflow research with pharmacists and operations staff, reducing manual verification steps by about 30%.

Fulfillment Cost Projection Service | Java, Spring Boot, DynamoDB, statistical modeling

- Built a high-throughput forecasting API using historical facility- and medication-level data, handling 2,000+ TPS across 20+ fulfillment facilities and producing real-time cost estimates for operational routing decisions.
- Built facility- and medication-level data flows connecting statistical estimates to downstream production services.
- Implemented low-latency model-backed decisioning and reduced DynamoDB read latency from about 400 ms to about 10 ms.
- Translated statistical forecasts into production routing decisions that supported \$9M+ in annual savings.

EDUCATION

University of Southern California | M.S. in Applied Data Science May 2022

Boston University | B.A. in Economics and Mathematics, Cum Laude Jun 2019

QUANTITATIVE RESEARCH EXPERIENCE

Machine Learning for STEM Career Interest and Learning Experiences | Research collaborator with Dr. Rongxiu Wu | Manuscript under review

- Contributed to a study of 15,847 students predicting end-of-high-school STEM career interest from in-school and out-of-school experiences.
- Implemented and compared L1-regularized logistic regression, random forest, XGBoost, and SVM pipelines on large-scale student survey data.
- Supported preprocessing, model selection, accuracy and ROC-AUC analysis, and SHAP-based interpretation of feature importance, nonlinear effects, and interactions.
- Compared predictive performance and interpretability across linear and nonlinear models, with XGBoost providing the strongest performance in the study.
- Translated nonlinear feature effects and model comparisons into interpretable evidence for downstream empirical analysis.

Large-Scale Survey Research and Multilevel Modeling | Research collaborator with Dr. Rongxiu Wu and coauthors | Preprint and manuscripts under review

- Credited with methodology, software, formal analysis, and investigation on a PISA 2022 multilevel study of digital self-regulation and emotional control across 19,779 students in 996 schools in Canada and Hong Kong (preprint).
- Contributed to PISA 2022 studies of teacher support, reading performance, immigrant status, self-efficacy, gender differences, school context, student belonging, and cross-cultural learning environments.
- Supported preparation of student- and school-level variables, survey weights, and modeling datasets across country contexts.
- Contributed to interpretable analyses using ANCOVA, school fixed effects, multilevel models, effect sizes, and nested model comparison.
- Supported model diagnostics and comparison using ICC, AIC, BIC, log likelihood, and marginal and conditional R-squared.

PREPRINTS

1. Chen, X., Qin, L., **Chen, X.**, Wu, R., & Sonnert, G. (2026). Digital Self-Regulation Before Sleep and Emotional Control Among Adolescents: A Multilevel Cross-National Study of Canada and Hong Kong, Using PISA 2022. *Preprints*. preprints.org/manuscript/202609.0400/v1

MANUSCRIPTS UNDER REVIEW

1. Wu, R., Song, Z., **Chen, X.**, Su, R., & Shen, Z. (Under Review). Multi-Dimensional Teacher Support and Reading Performance: The Moderating Role of Immigrant Status in the U.S. *Journal of Youth and Adolescence*.
2. Wu, R., **Chen, X.**, & Su, R. (Under Review). Problems With Self-Efficacy Across Educational Systems: Gender Differences and School Contexts in PISA 2022. *Current Psychology*.
3. Su, Y., Shao, Y., Wu, R., & **Chen, X.** (Under Review). Cultural Context and Students' Sense of Belonging: A Multilevel Analysis of PISA 2022 Across Confucian-Heritage and Western Societies. *Learning, Culture and Social Interaction*.
4. Wu, R., Su, Q., **Chen, X.**, & Su, R. (Under Review). "Problem-Driven" and "Student-Led": A Comparative Case Study of Transdisciplinary Talent Cultivation in Four Hong Kong Universities. *Active Learning in Higher Education*.
5. Wu, R., & **Chen, X.** (Under Review). Predicting STEM Career Interest at the End of High School: A Machine Learning Approach to In-School and Out-of-School Time Experiences.

TECHNICAL SKILLS

Programming: Python, R, Java, JavaScript / TypeScript, SQL, Go

AI / ML: multimodal document understanding; multi-agent orchestration; LLM integration and prompting; human-in-the-loop system design; agent- and system-level evaluation; statistical modeling; supervised machine learning; model interpretability; ML pipeline engineering

Quantitative Methods: multilevel modeling; fixed effects; ANCOVA; survey-weighted analysis; model comparison; ROC-AUC; effect sizes; SHAP-based interpretation

Research Tools: lme4; lmerTest; XGBoost; scikit-learn; SHAP; large-scale survey data processing

Systems: Spring Boot; REST APIs; microservices; React; AWS Lambda; Bedrock; ECS; DynamoDB; S3; Docker; Kubernetes; CI/CD; Linux

ADDITIONAL EXPERIENCE

Software Development Intern, ByteDance (TikTok) | Beijing, China Jun 2021 - Aug 2021