XINGDA LYU

Seattle, WA xlyu18@uw.edu Website Google Scholar

Education

University of Washington, Seattle, WA *Expected Jun 2027*
B.S. in Statistics (Data Science Track) and B.S. in Informatics (Departmental Honors)
Minor in Philosophy
Cumulative GPA: 3.90/4.00
Selected Coursework: Machine Learning, Data Structures and Algorithms, Matrix Algebra, Probability and Mathematical Statistics I–III, Real Analysis, Generalized Linear Models, Statistical Computing.

Research Experience

Data Mining Group Jun 2026 – Present
Research Assistant University of Illinois Urbana-Champaign
Advisor: Prof. Jiawei Han; Research Mentor: Joey Chan.

- Lead development of RetroRAG for training-free, single-step LLM retrosynthesis using reaction-class, reaction-center, and structure-guided multi-stage retrieval with template-supported precursor selection.
- Analyze the predictive utility of reaction precedents and develop retrieval strategies to select the most informative reaction evidence for inference-time precursor prediction across LLMs. Targeting submission to KDD 2027.

Computational Minds and Machines Lab Aug 2026 – Present
Research Assistant University of Washington
Advisor: Prof. Max Kleiman-Weiner; Research Mentor: Kirandeep Kaur.

- Develop a benchmark for meta-ignorance in proactive AI assistants, evaluating how models reason about users' epistemic states and adapt assistance strategies accordingly.
- Co-lead experimental design and evaluation across distinct user epistemic states using controlled multi-turn scenarios with varied metacognitive cues; analyze assistance-strategy selection and failure modes. Targeting submission to ACL 2027.

UW NLP Group Apr 2026 – Present
Research Collaborator University of Washington
Research Mentor: Shuyue Stella Li.

- Lead development of VVRBench v2, improving constraint-based task sampling and programmatic verification for difficulty-controlled, compositional spatial-reasoning tasks.
- Extend VVRBench to 2D and 3D spatial reasoning in LLMs and VLMs, identifying model capability gaps and enabling reinforcement learning with verifiable rewards.

InfoSeeking Lab Nov 2025 – Aug 2026
Research Assistant University of Washington
Advisor: Prof. Chirag Shah; Research Mentor: Kirandeep Kaur.

- Co-developed an epistemic–behavioral coupling framework drawing on the philosophy of ignorance and grounding behavioral commitment in epistemic legitimacy; proposed epistemic partnership for joint human–AI inquiry into unknown unknowns and evolving understanding. Co-authored and presented a position paper at **ICML 2026** in Seoul.

Proactive AI for Scientific Communication Sep 2025 – Present
Research Project Lead Student-Led Research Project

- Initiated and led development of PROS, a proactive AI system integrating source-grounded problem discovery, user-governed intervention, and validated native-object PPTX editing.
- Developed PROS-Bench (120 source papers, 320 editable posters) and an end-to-end diagnosis–repair–outcome evaluation framework; achieved 87.6% operator-verified resolution of accepted issues and +22.7-point VLM-judged accepted-target uplift. First-author arXiv preprint; presented at **CVPR 2026** HiGen Workshop.

Publications & Presentations

ICML '26 Position: Knowing Isn't Understanding: Re-grounding Generative Proactivity with Epistemic and Behavioral Insight. Kirandeep Kaur, **Xingda Lyu**, and Chirag Shah. *ICML 2026.* [paper] [poster]

arXiv '26 Beyond Instruction-Driven Editing: Source-Grounded Problem Discovery with User-Governed Repair for Scientific Posters. **Xingda Lyu**, Honglin Lu, Xinye Luo, and Shiqi Yang. *arXiv 2026.* [paper]

CVPR WS '26 PROS: A Proactive Multimodal Refinement Agent for Scientific Posters. **Xingda Lyu**, Xinye Luo, Honglin Lu, and Shiqi Yang. *CVPR 2026 HiGen Workshop.* [workshop]

NeurIPS '26 Agents' Last Exam. Yiyou Sun, . . . , **Xingda Lyu**, . . . , and Dawn Song. *Data contributor. NeurIPS 2026.* [paper]

CONF-CDS '25 P-RAG: Prompt-Enhanced Parametric RAG with LoRA and Selective CoT for Biomedical and Multi-Hop QA. **Xingda Lyu**, Gongfu Lyu, Zitai Yan, and Yuxin Jiang. *International Conference on Computing and Data Science (CONF-CDS), 2025.* [paper]

Selected Presentations

NSF AI-SDM '26 Beyond Instruction Following: Toward Proactive Human–AI Collaboration in Scientific Poster Refinement. **Xingda Lyu**, Honglin Lu, Xinye Luo, and Shiqi Yang. *NSF AI-SDM Workshop, Carnegie Mellon University, 2026. Poster.* [abstract] [poster]

CHIIR WS '26 Decision-Making Under Unknown Unknowns in Proactive Agentic Systems. **Xingda Lyu**. *ACM CHIIR Workshop on Human-Centered Proactive & Personalized Agents, 2026. Lightning Talk & Poster.* [workshop]

Amazon RD '26 ProPer: Proactivity Driven Personalized Agents for Knowledge Gap Navigation. Kirandeep Kaur, Vinayak Gupta, Aditya Gupta, **Xingda Lyu**, and Chirag Shah. *Amazon Research Day, AI Co-Scientist, 2026. Poster Presentation.* [poster]

Industry Experience

Amazon Jun 2025 – Aug 2025
Data Analyst Intern Seattle, WA

- Built Python and SQL data-quality pipelines processing 300K+ transportation records; conducted routing and anomaly-detection analyses to identify operational inefficiencies, refund issues, and underperforming carriers.

Baynovation Jul 2024 – Sep 2024
Data Analyst Intern San Jose, CA

- Developed a mortgage-refinancing decision tool and interactive dashboard using real-time rate data; modeled delinquency risk with GLMs, Random Forest, and XGBoost.

Honors & Service

Reviewer: *AI Magazine* (AAAI); ACM IUI 2027 (Posters & Demos).
iSchool Conference Travel Award, University of Washington, 2026.
NSF AI-SDM Workshop Travel Funding, 2026.
Invited Panelist: UW Allen School CSE 394 Undergraduate Researcher Panel.

Skills

Programming & Tools: Python, Java, R, SQL; Git, Linux

ML Frameworks: PyTorch, Hugging Face Transformers, scikit-learn

ML & Agents: LLM Agent Development, Benchmark Design & Evaluation, RAG, Reinforcement Learning, LLM Fine-Tuning