

Stella Li

📍 Seattle, WA ✉ stelli@cs.washington.edu 🌐 stellalisy.com 📧 stellalisy

Research Interests

Reinforcement Learning, Meta-Cognition, Personalization, Safety & Reliability, Social Reasoning

Education

University of Washington

Sep. 2023 – Current

Ph.D. in Computer Science

- Advisor: Yulia Tsvetkov

Johns Hopkins University

Sep. 2022 – May 2023

M.S. in Computer Science

- Advisors: Philipp Koehn & Kenton Murray
- Thesis: Learning from Gibberish: Code-Mixing Data Augmentation for Sentiment Analysis
- Cumulative GPA: 4.0/4.0

Johns Hopkins University

Sep. 2019 – Dec 2022

B.A. in Cognitive Science (linguistics focus)

- Additional Majors: Computer Science, Applied Math (statistics focus); Minor: Mathematics
- Cumulative GPA: 3.99/4.0, Major GPA: 4.0/4.0

Select Publications

Verifiable Visual Rewards Transfer from Synthetic Scenes to Natural Prompts

Shuyue Stella Li, Xiaochuang Han, Yulia Tsvetkov, Luke Zettlemoyer. *Preprint 2026.*

ExploreNet: Predictive Exploration Scaling for Diffusion GRPO

Shuyue Stella Li, Xiaochuang Han, Yulia Tsvetkov, Luke Zettlemoyer. *Preprint 2026.*

IsoVGRBench: Video Reward Modeling Without Human Feedback

Shuyue Stella Li, Jacqueline He, Jiahui Chen, Sreya Dutta Roy, Yulia Tsvetkov, Luke Zettlemoyer, Marjan Ghazvininejad, Xiaochuang Han. *Preprint 2026.*

Fighting Fire with Fire: Assessing Test Set Contamination Via Deliberate Training on Test Data

Shuyue Stella Li*, Melanie Sclar*, Niloofar Mireshghallah, Yejin Choi, Yulia Tsvetkov. *Preprint 2026.*

HorizonBench: Long-Horizon Personalization with Evolving Preferences

Shuyue Stella Li, Bhargavi Paranjape, Kerem Oktar, Zhongyao Ma, Gelin Zhou, Lin Guan, Na Zhang, Sem Park, Lin Chen, Diyi Yang, Yulia Tsvetkov, Asli Celikyilmaz. *Preprint 2026.*

EvoLM: Self-Evolving Language Models through Co-Evolved Discriminative Rubrics

Shuyue Stella Li*, Rui Xin*, Yike Wang, Yike Wang, Teng Xiao, Rulin Shao, Zoey Hao, Melanie Sclar, Faeze Brahman, Pang Wei Koh, Yulia Tsvetkov. *Preprint 2026.*

Cold-Start Personalization via Training-Free Priors from Structured World Models

Avinandan Bose*, Shuyue Stella Li*, Faeze Brahman, Pang Wei Koh, Simon Shaolei Du, Yulia Tsvetkov, Maryam Fazel, Lin Xiao, Asli Celikyilmaz. *In Proceedings of ICML 2026.*

Ads in AI Chatbots? An Analysis of How Large Language Models Navigate Conflicts of Interest

Addison J Wu*, Ryan Liu*, Shuyue Stella Li, Yulia Tsvetkov, Thomas L Griffiths. *COLM 2026.*

Personalized Reasoning: Just-In-Time Personalization and Why LLMs Fail At It

Shuyue Stella Li*, Avinandan Bose*, Faeze Brahman, Simon Shaolei Du, Pang Wei Koh, Maryam Fazel, Yulia Tsvetkov. *In Proceedings of ICLR 2026.*

Cognitive Foundations for Reasoning and Their Manifestation in LLMs

Priyanka Kargupta*, Shuyue Stella Li*, Haocheng Wang, Jinu Lee, Dean Light, Shan Chen, Orevaoghene Ahia, Tom Griffiths, Max Kleiman-Weiner, Jiawei Han, Asli Celikyilmaz, Yulia Tsvetkov. *Preprint 2025.*

OLMo 3

Team Olmo, . . . , [Shuyue Stella Li](#), et al. *Technical Report 2025*.

Spurious Rewards: Rethinking training signals in RLVR

Rulin Shao*, [Shuyue Stella Li](#)*, Rui Xin*, Scott Geng*, Yiping Wang, Sewoong Oh, Simon Shaolei Du, Nathan Lambert, Sewon Min, Ranjay Krishna, Yulia Tsvetkov, Hannaneh Hajishirzi, Pang Wei Koh, Luke Zettlemoyer. *In Proceedings of ICML 2026*.

ALFA: Aligning llms to ask good questions a case study in clinical reasoning

[Shuyue Stella Li](#)*, Jimin Mun*, Pedram Hosseini, Bryceton G. Thomas, Jessica M. Sin, Bing Ren, Jonathan S. Ilgen, Yulia Tsvetkov, Maarten Sap. *In Proceedings of COLM 2025*.

PrefPalette: Personalized preference modeling with latent attributes

[Shuyue Stella Li](#), Melanie Sclar, Hunter Lang, Ansong Ni, Jacqueline He, Puxin Xu, Andrew Cohen, Chan Young Park, Yulia Tsvetkov, Asli Celikyilmaz. *In Proceedings of COLM 2025. Spotlight Paper*.

Medical hallucinations in foundation models and their impact on healthcare

Yubin Kim, Hyewon Jeong, Shan Chen, [Shuyue Stella Li](#), . . . , Cynthia Breazeal. *Preprint 2025*.

MediQ: Question-asking LLMs and a benchmark for reliable interactive clinical reasoning

[Shuyue Stella Li](#), Vidhisha Balachandran, Shangbin Feng, Jonathan S Ilgen, Emma Pierson, Pang Wei Koh, Yulia Tsvetkov. *In Proceedings of Neurips 2024*.

Valuescope: Unveiling implicit norms and values via return potential model of social interactions

Chan Young Park*, [Shuyue Stella Li](#)*, Hayoung Jung*, Svitlana Volkova, Tanushree Mitra, David Jurgens, Yulia Tsvetkov. *In Proceedings of EMNLP 2024*.

A false sense of privacy: Evaluating textual data sanitization beyond surface-level privacy leakage

Rui Xin, Niloofar Mireshghallah, [Shuyue Stella Li](#), Michael Duan, Hyunwoo Kim, Yejin Choi, Yulia Tsvetkov, Sewoong Oh, Pang Wei Koh. *In SaTML 2025*.

Learning from mistakes: Towards robust neural machine translation for disfluent L2 sentences

[Shuyue Stella Li](#), Philipp Koehn. *In Proceedings of MT Summit 2023*.

A Quantitative approach to understand self-supervised models as cross-lingual feature extractors

[Shuyue Stella Li](#)*, Beining Xu*, Xiangyu Zhang*, Hexin Liu, Wenhan Chao, Leibny Paola Garcia. *In Proceedings of IJCNLP 2023*.

A new approach to extract fetal electrocardiogram using affine combination of adaptive filters

Yu Xuan, Xiangyu Zhang, [Shuyue Stella Li](#), Zihan Shen, Xin Xie, Leibny Paola Garcia, Roberto Togneri. *In Proceedings of ICASSP 2023*.

Teaching Experience

Ethics in AI (CSE 582)

Teaching Assistant

University of Washington

Jan. 2025 – March 2025

Introduction to Statistics (EN.553.430)

Teaching Assistant

Johns Hopkins University

Sep. 2021 – May 2023

Human-Computer Interaction (EN.601.490)

Course Assistant

Johns Hopkins University

Sep. 2022 – Dec. 2022

Computer Ethics (EN.601.104)

Head Course Assistant

Johns Hopkins University

May 2022 – Aug. 2022

Intermediate Programming (EN.601.220)

Course Assistant

Johns Hopkins University

Sep. 2021 – May 2022

Mentoring

Mohammadamin Shafiei, M.S., University of Milan

Sep. 2025 – July 2026

- Sole mentor; LLM bias behind logical puzzles; outcome: EMNLP 2026 paper. COLM Workshop Oral.

Jack Zhang, Undergraduate, UW

Feb. 2025 – July 2026

- Mentored end-to-end on mid-training influence. Spotlight paper at COLM Workshop for Efficient Reasoning.

Arnav Mazumder, Undergraduate, UW

Feb. 2025 – July 2026

- Mentored with another PhD student; contextual machine translation; outcome: WMT 2026 paper.

Work Experience

Visiting Researcher

Meta FAIR

Seattle, WA

Sep. 2024 – Current

- Visiting researcher at FAIR working on social reasoning & alignment projects.

Research Intern

MSU Genetic Programming Lab

East Lansing, MI

May 2021 – Aug 2021

- Advised by Dr. Wolfgang Banzhaf; outcome: 2 research publications, 1 [open-source software](#) .
- Designed and implemented novel GP algorithm for LLVM compiler flag optimization (20%).

Research Intern

JHU Language & Cognition Lab

Baltimore, MD

Jan 2019 – May 2021

- Advised by Dr. Barbara Landau, studying developmental spatial cognition through block building.
- Designed experiments, collected eye-tracking data from 50+ participants, developed coding scheme and analysis pipeline for the collected data.

Honors & Awards

Qualcomm Fellowship Finalist

2026

Jane Street Fellowship Honorable Mention

2026

IEEE Rising Stars 2026

2026

COLM Spotlight Paper Award (24/418 papers)

2025

JHU Responsible AI for Health Student Travel Grant

2024

Weil Family Endowed Fellowship in Computer Science & Engineering

2023-24

Johns Hopkins Dean's List

2019-23

Upsilon Pi Epsilon Computer Science Honor Society

2022

GECCO-GI Best Presentation Award

2022

ACM Student Travel Grant

2022

PAJH Greek Scholars Award

2022

Grace Hopper Scholarship Award

2022

Omicron Delta Kappa National Leadership Honors Society

2022

Omega Psi National Cognitive Science Honors Society

2022

Cum Laude Society

2022

Service

Reviewer: ACL, EMNLP, NAACL, NeurIPS, ICLR, AISTATS, COLM

Skills

Languages: Python, C++, C, Java, R, MATLAB, HTML/CSS, JavaScript, ReactJS, Go