Litao Yan

ltyan@seas.upenn.edu • yanlitao.github.io • litao-yan

Research Interests

Human-Computer Interaction, Large Language Models, Programming Assistants, Code Comprehension, Interactive Data Visualization, Software Engineering

Education

Sep. 2022 – Present: **University of Pennsylvania**, Philadelphia, PA, United States
Ph.D. in Computer and Information Science, Advisor: *Andrew Head*

Sep. 2019 – May 2021: **Harvard University**, Cambridge, MA, United States
M.Eng. in Computational Science and Engineering, Advisor: *Elena L. Glassman*

Sep. 2015 – Jun. 2019: **Xiamen University**, Xiamen, Fujian, China
B.Eng. in Material Science and Engineering, Advisors: *Yixi Zhuang & Ye Luo*

Selected Publications

CHI 2024: **Ivie: Lightweight Anchored Explanations of Just-Generated Code**
Litao Yan, Alyssa Hwang, Zhiyuan Wu, Andrew Head. 2024.
ACM Conference on Human Factors in Computing Systems.
PDF *Video*

UIST 2022: **Concept-Annotated Examples for Library Comparison**
Litao Yan, Miryung Kim, Björn Hartmann, Tianyi Zhang, Elena Glassman. 2022.
ACM Symposium on User Interface Software and Technology.
PDF *Video*

CHI 2021 (Honorable Mention): **Visualizing Examples of Deep Neural Networks at Scale**
Litao Yan, Elena L. Glassman, Tianyi Zhang. 2021.
ACM Conference on Human Factors in Computing Systems.
PDF *Video*

Selected Honors and Awards

2022: John Grist Brainerd Doctoral Fellow, University of Pennsylvania
2021: Honorable Mention at CHI'21 (top 5% of submitted papers)
2019: Outstanding Graduates, Xiamen University (top 3%)
2016 - 2018: Outstanding Scholarship for Undergraduates (top 5%), Xiamen University
2016 & 2017: Triple-A Student, Xiamen University (top 3%)
2016 & 2017: "An An" Scholarship First Prize (top 3%), College of Materials, Xiamen University

Research Experience

University of Pennsylvania **Sep. 2022 – Jun. 2026 (expected)**
Ph.D. Student, PA, United States *Advisor: Prof. Andrew Head*

Current project: Developing software engineering agents that integrate IDE APIs to mitigate model hallucination in answering reachability and control flow questions.

Ivie: Created Ivie, a programming assistant providing real-time, in-situ explanations for code generated by GitHub Copilot. Built as a VSCode extension using TypeScript and OpenAI APIs, Ivie improved code comprehension accuracy by 25% and reduced task completion time compared to GPT-based chatbots, as demonstrated in a user study with 32 participants.

Harvard University **Sep. 2019 – Aug. 2022**
Master's Student & Research Assistant, MA, United States *Advisor: Prof. Elena Glassman*

ParaLib: Developed ParaLib to show 150 concept-annotated code examples in parallel, facilitating library comparison and offering insights into library selections. I conducted a user study with 20 participants to evaluate ParaLib spanning Visualization and NLP domains. ParaLib improved developer confidence in library selection by 35% and reduced cognitive load.

ExampleNet: I created ExampleNet, an interface that visualizes 100 neural network architectures, extracted from 203 highly-starred GitHub projects using call graph analysis in Python AST, identifying 27 layer types and 9 key hyperparameters. The tool helped participants inspect three times more examples and make more diverse design choices in neural network design, as shown in a study with 16 participants.

Massachusetts Institute of Technology **Jun. 2020 – Apr. 2021**
Research Assistant, MA, United States *Advisors: Prof. Michael Cafarella & Prof. Tim Kraska*

VisMeet: Developed VisMeet, an interactive video conferencing system designed to enhance meeting participation and content comprehension through dynamic, real-time visualizations. VisMeet integrates transcript-based topic hierarchies, speaker contributions, and interactive summaries to provide users with better contextual understanding during meetings. I conducted a study with 39 MTurk participants, showing a 140% improvement in users' ability to recall meeting content.

Xiamen University **Sep. 2015 – Jun. 2019**
Undergrad Student, Fujian, China *Advisors: Prof. Yixi Zhuang & Prof. Ye Luo*

Thesis: I worked on predicting the thermodynamic stability of Co-V-Ti-Ta quaternary superalloys using machine learning models. By combining first-principles calculations and density functional theory (DFT), I automated the extraction of 6,219 possible atomic structures. These were used to train machine learning models to predict stability with high accuracy for previously uncalculated configurations. My research reduced the need for extensive first-principles computations while maintaining prediction accuracy, offering a more efficient approach to alloy design in high-temperature applications like aerospace engines and turbines.

Chip-Firing: I simulated the Abelian Sandpile Model, visualizing patterns of 10 million grains on an infinite graph grid. Innovatively, I integrated this model into a new stream cipher by replacing the LFSR and further applied it to study social networks and simulate dynamic systems.

Professional Experience

Intern for Front-End and Data Visualization **Mar. 2018 – May 2019**
PZCNET (Xiamen) Ecommerce Co., Ltd., Fujian, China *Advisor: Prof. Defu Zhang*
Developed a user interface for a real-time agricultural big data platform, optimizing agricultural production for over 500 products. I implemented advanced features like production forecasting, precision marketing, and food safety traceability using TensorFlow for predictive modeling.

Intern for Sports Data and Interface Design **Oct. 2016 – Dec. 2018**
GenGee Sport Co., Ltd., Fujian, China *Advisor: Yelei Zhang (CTO)*
I developed a real-time soccer performance analysis interface for the INSAIT K1 system, visualizing 16 performance metrics, including player fitness and tactical movements. The system reduced post-game analysis time by 27% for over 4,000 coaches and athletes, improving decision-making through data visualization.

Skills

	Level	Skill	Years
Languages	■■■■■	Python	7
	■■■■■	JavaScript	5
	■■■■■	TypeScript	2
	■■■■■	Java, C	3, 8
	■■■■■	MATLAB, R, Go	4, 2, 2
Front-End	■■■■■	HTML & CSS, Sass/SCSS, React, AJAX, jQuery	5
Back-End	■■■■■	Node.js, MongoDB, Prisma, WebSocket	5
	■■■■■	Docker, GitHub Actions, Webpack	3
Data Analysis	■■■■■	scikit-learn, Pandas, Numpy, TensorFlow, PyTorch, Keras	7
Visualization	■■■■■	Matplotlib, Plotly, Seaborn, D3.js, Three.js	6
UX Research	■■■■■	User Interviews, Usability Testing, A/B Testing, Heuristic Evaluation	5
	■■■■■	Contextual Inquiry, Diary Studies	5
Techniques	■■■■■	Data Flow / Call Graph Analysis, Eye-Tracking Analysis	5
	■■■■■	Unit Testing & TDD, Version Control (e.g., Git)	4
	■■■■■	CI/CD Pipelines, API Design & Development	3
Statistics	■■■■■	Statistical Methods (e.g., T-Test, ANOVA, Chi-Square)	5
Design Tools	■■■■■	Figma, Sketch, Adobe Photoshop, Adobe XD, AutoCAD	5

Academic Service

2024: **Reviewer**, The ACM Symposium on User Interface Software and Technology (UIST)
2023: **Reviewer**, ACM Conference on Designing Interactive Systems (DIS)
2023, 2024: **Reviewer**, ACM Conference on Human Factors in Computing Systems (CHI)
Fall 2022: **Teaching Assistant**, CIS 3990: Introduction to Human-Computer Interaction

Invited Talks

Microsoft Research Lab - Cambridge, HCI community talks **Oct. 2, 2024**
Ivie: Lightweight Anchored Explanations of Just-Generated Code

CHI 2024 **May 14, 2024**
Ivie: Lightweight Anchored Explanations of Just-Generated Code

GitHub Copilot teams **Jan. 12, 2024**
Ivie: Lightweight Anchored Explanations of Just-Generated Code

PhD Special Topics: Large Language Models & Programming Languages **Nov. 10, 2023**
Ivie: Lightweight Anchored Explanations of Just-Generated Code

UIST 2021 **Nov. 1, 2022**
Concept-Labeled Examples for Library Comparison

CHI 2021 **May 11, 2021**
Visualizing Examples of Deep Neural Networks at Scale