Xinyan Velocity Yu

Department of Computer Science
University of Southern California
941 Bloom Walk, Los Angeles, CA 90089

Email: xinyany@usc.edu
Website: velocitycavalry.github.io
Linkedin: velocityyu

Education

† *Indicates expected*

2023 - 2028 †	Ph.D., Computer Science, University of Southern California M.S., expected 12/2025 GPA: 4.0, Supervisor: Jesse Thomason
2021–2023	M.S., Computer Science, University of Washington GPA: 4.0, Supervisors: Hannaneh Hajishirzi, Luke Zettlemoyer
2017–2021	B.S., Computer Science & Applied and Computational Mathematics, University of Washington GPA: 3.91, with Departmental Honor and Cum Laude

Employments

05/2025 – 08/2025	Research Intern, **Microsoft Semantic Machines** Retrieval for Reasoning (Multi-GPU distributed training involved) Host: Patrick Xia, Benjamin Van Durme
07/2022–07/2023	Part-time AI Resident, **Meta AI (Facebook AI Research)** Inference time decoding algorithms Host: Asli Celikyilmaz, Mike Lewis
12/2021–06/2022	Research assistant, **University of Washington** Supervisors: Sewon Min, Hannaneh Hajishirzi, Luke Zettlemoyer
08/2021–11/2021	Software Engineering Intern, Facebook Ads Ranking, **Facebook** Implement a versioning system for counter-based features in Facebook Ads to reduce the bias in internal experimentation using C++ and Python. Implement and test different multi-armed bandits selection strategies for ads options in C++.
05/2021–08/2021	Software Engineering Intern, Search Features, **Pinterest** Implement topic-based user search using Java, Python, Bazel and Apache Hadoop for profile's tab based on query intent detection for Pinterest's profile search to showcase individual creators in key verticals, increases creators profile views by 11%: Related press release
04/2021–06/2021	Undergraduate research assistant, **University of Washington** Supervisors: Akari Asai, Jungo Kasai, Hannaneh Hajishirzi
06/2020–09/2020	Software Engineering Intern, Azure Networking, **Microsoft**

Implement a pipeline using C++ and Python that runs in production nightly for every data center for efficient anomalies detection in running router configurations for Azure Networking.

Honors and Awards

2024	Outstanding Reviewer (**EMNLP 2024**)
2023–2024	Viterbi Graduate Fellowship, **University of Southern California**
2022	Selected Scholar, **The Cornell, Maryland, Max Planck Pre-doctoral Research School**
2021	Computer Science Departmental Honor, **University of Washington**
2021	Infospace Endowed Scholarship, **University of Washington**
2019	Grace Hopper Celebration Scholarship, **University of Washington**
2017–2021	Annual Dean's List, **University of Washington**

Publications

Google Scholar: https://scholar.google.com/citations?user=PoZv5KkAAAAJ
Semantic Scholar: https://www.semanticscholar.org/author/Xinyan-Velocity-Yu/2118211280
* *denotes equal contributions*

Conference Papers

[1] Wang Bill Zhu, Tianqi Chen, **Xinyan Velocity Yu**, Ching Ying Lin, Jade Law, Mazen Jizzini, Jorge J. Nieva, Ruishan Liu, Robin Jia
"Cancer-Myth: Evaluating Large Language Models on Patient Questions with False Presuppositions." 2025.
In: *International Conference on Learning Representations (**ICLR**)*. 2026. [paper]

[2] Zora Zhiruo Wang*, Akari Asai*, **Xinyan Velocity Yu**, Frank F. Xu, Yiqing Xie, Graham Neubig, Daniel Fried.
"CodeRAG-Bench: Can Retrieval Augment Code Generation?"
In: *Findings of the North American Chapter of the Association for Computational Linguistics (**NAACL**)*. 2025. [paper]

[3] Zhaofeng Wu, **Xinyan Velocity Yu**, Dani Yogatama, Jiasen Lu, Yoon Kim.
"The Semantic Hub Hypothesis: Language Models Share Semantic Representations Across Languages and Modalities."
In: *International Conference on Learning Representations (**ICLR**)*. 2025. [paper]

[4] Li Du, Afra Amini, Lucas Torroba Hennigen, **Xinyan Velocity Yu**, Jason Eisner, Holden Lee, Ryan Cotterell.
"Principled Gradient-Based MCMC for Conditional Sampling of Text."
In: *International Conference on Machine Learning (**ICML**)*. 2024. [paper]

[5] Akari Asai, Sneha Kudugunta, **Xinyan Velocity Yu**, Terra Blevins, Hila Gonen, Machel Reid, Yulia Tsvetkov, Sebastian Ruder, Hannaneh Hajishirzi.

"BUFFET: Benchmarking Large Language Models for Few-shot Cross-lingual Transfer."
In: *North American Chapter of the Association for Computational Linguistics (**NAACL**)*. 2024. [paper] **Oral Presentation**

[6] Ting-Rui Chiang, **Xinyan Velocity Yu**, Joshua Robinson, Ollie Liu, Isabelle Lee, Dani Yogatama.
"On Retrieval Augmentation and the Limitations of Language Model Training"
In: *North American Chapter of the Association for Computational Linguistics (**NAACL**)*. 2024. [paper]

[7] Jungo Kasai, Keisuke Sakaguchi, Yoichi Takahashi, Ronan Le Bras, Akari Asai, **Xinyan Yu**, Dragomir Radev, Noah A. Smith, Yejin Choi, Kentaro Inui.
"RealTime QA: What's the Answer Right Now?"
In: *Neural Information Processing Systems (**NeurIPS**) Datasets and Benchmarks Track*. 2023. [paper]

[8] **Xinyan Velocity Yu**, Sewon Min, Luke Zettlemoyer, Hannaneh Hajishirzi.
"Crepe: Open-Domain Question Answering with False Presuppositions."
In: *Annual Meeting of the Association for Computational Linguistics (**ACL**)*. 2023. [paper] **Oral Presentation**

[9] Swarnadeep Saha, **Xinyan Velocity Yu**, Mohit Bansal, Ramakanth Pasunuru, Asli Celikyilmaz.
"MURMUR: Modular Multi-Step Reasoning for Semi-Structured Data-to-Text Generation."
In: *Findings of the Annual Meeting of the Association for Computational Linguistics (**ACL**)*. 2023. [paper]

[10] **Xinyan Velocity Yu***, Akari Asai*, Trina Chatterjee, Junjie Hu, Eunsol Choi.
"Beyond Counting Datasets: A Survey of Multilingual Dataset Construction and Necessary Resources."
In: *Findings of the Empirical Methods in Natural Language Processing (**EMNLP**)*. 2022. [paper]

[11] Akari Asai, **Xinyan Yu**, Jungo Kasai, Hannaneh Hajishirzi.
"One Question Answering Model for Many Languages with Cross-lingual Dense Passage Retrieval."
In: *Neural Information Processing Systems (**NeurIPS**)*. 2021. [paper]

Preprints

[P1] Zefan Cai*, Haoyi Qiu*, Tianyi Ma*, Haozhe Zhao*, Gengze Zhou, Tingting Liao, **Xinyan Velocity Yu**, Kung-Hsiang Huang, Shawn Lin, Parisa Kordjamshidi, Minjia Zhang, Wen Xiao, Jiuxiang Gu, Nanyun Peng, Junjie Hu
"MMGR: Multi-Modal Generative Reasoning"
Submitted to ICML.

[P2] Ting-Rui Chiang, Joshua Robinson, **Xinyan Velocity Yu**, Dani Yogatama.
"LocateBench: Evaluating the Locating Ability of Vision Language Models." *arXiv Preprint*. 2024. [paper]

[P3] Yuehan Qin, Shawn Li, Yi Nian, **Xinyan Velocity Yu**, Yue Zhao, Xuezhe Ma.
"Don't Let It Hallucinate: Premise Verification via Retrieval-Augmented Logical Reasoning."
arXiv Preprint. 2025. [paper]

Professional Services

Reviewer / Program Committee

Conferences:

- ICLR (2024, 2025, 2026)
- ACL Rolling Review (2023, 2024, 2025)
- NeurIPS (2023, 2025)
- EACL (2023)
- AKBC (2022)
- CoNLL (secondary; 2022)

Workshops:

- ML Reproducibility Challenge (2023)
- Multilingual Representation Learning Workshop at EMNLP (2023)
- Multilingual Information Access Workshop at NAACL (2022)

Departmental Services

2019–2020	UW ACM: Treasurer: Prepare budget, tax and legal documents
2018–2019	UW ACM: Associate Officer: Research nights and events outreach

Open-source Contributions

Anserini, Pyserini	Implement feature of indexing using Apache Lucene and retrieving multilingual user documents using BM25 in Java and Python. Integrate Apache Lucene analyzers in Anserini for 15 languages using Java and Apache Maven (Related PR: Anserini#1548) Enable the corresponding Python wrapper in Pyserini using Python and Pyjnius (Related PR: Pyserini#591)
VisualSketchpad	Increase the robustness of multi-agent systems during prompt feeding and path reading (Related commits: #7f311f1, #d504794, #d4b6a86)

Teaching Assistantships

Fall 2024	CSCI 567: Advanced Machine Learning Instructor: Dani Yogatama	University of Southern California
Spring 2023	CSE 444: Database Systems Internal Instructor: Ryan Mass	University of Washington
Winter 2023	CSE 447: Natural Language Processing Instructor: Sofia Serrano	University of Washington
Spring 2022	CSE 415: Artificial Intelligence (non-major) Instructor: Megan Hazen	University of Washington
Winter 2022	CSE 447: Natural Language Processing	University of Washington

	Instructor: Noah Smith	
Spring 2021	CSE 473: Artificial Intelligence	University of Washington
	Instructor: Hannaneh Hajishirzi	
Winter 2021	CSE 473: Artificial Intelligence	University of Washington
	Instructor: Eric Hsu	
Fall 2020	CSE 333: Systems Programming	University of Washington
	Instructor: Hal Perkins	
Spring 2020	CSE 333: Systems Programming	University of Washington
	Instructor: Hal Perkins	
Winter 2020	CSE 332: Data Structures and Parallelisms	University of Washington
	Instructor: Kevin Lin	
Fall 2019	CSE 373: Data Structures and Algorithms	University of Washington
	Instructor: Kevin Lin	
Summer 2019	CSE 373: Data Structures and Algorithms	University of Washington
	Instructor: Robbie Weber	
Spring 2019	CSE 373: Data Structures and Algorithms	University of Washington
	Instructor: Kasey Champion	

Selected Coursework

ML Related	Learning Theory, Machine Learning, Deep Learning, Natural Language Processing, Artificial Intelligence
Math / Stats	Stochastic Processes, Discrete Math Modeling, Continuous Math Modeling, Probabilities I & II
General CS	Operating Systems, Systems Programming, Hardware/Software Interface Programming, Programming Languages, Functional Programming, Advanced Algorithms, Advanced Database, Database Management

Skills

Programming Languages	Python, Java, C, C++, C#, Julia, Scala, Rust, SQL, SQL++, NoSQL, MATLAB, R, Swift, SML, Racket, Ruby, Javascript, Bash
Frameworks	PyTorch, TensorFlow, Apache Hadoop, Spark, Azure, GCP, React, Vue