2313 Devon Circle
Ann Arbor, MI 48105
+1 (650) 815 1197
yukw777@gmail.com
yukw777.github.io
k-peter-yu
yukw777
wPlzAvEAAAAJ

Keunwoo Peter Yu

Curriculum Vitae

Research Interests

Multi-modal learning, vision-language models (videos), embodied AI.

Education

2020–2025 **Ph.D.**, *University of Michigan*, Ann Arbor, MI
Computer Science and Engineering
Advised by Joyce Chai in the Situated Language & Embodied Dialogue (SLED) lab
Dissertation title: Towards Generalist Vision-Language Models for Videos in Embodied AI

2020–2022 **M.S.E.**, *University of Michigan*, Ann Arbor, MI
Computer Science and Engineering
Artificial Intelligence

2009–2013 **B.S.E.**, *Princeton University*, Princeton, NJ
Computer Science
Cum Laude, Member of the Society of Sigma Xi

Publications

Under Review

[1] **Keunwoo Peter Yu** and Joyce Chai. Temporally-grounded language generation: A benchmark for real-time vision-language models. *arXiv preprint arXiv:2505.11326*, 2025.

[2] **Keunwoo Peter Yu**, Achal Dave, Rares Ambrus, and Jean Mercat. Espresso: High compression for rich extraction from videos for your vision-language model. *arXiv preprint arXiv:2412.04729*, 2024.

Published/Accepted Peer-reviewed

[1] **Keunwoo Peter Yu**, Zheyuan Zhang, Fengyuan Hu, Shane Storks, and Joyce Chai. Eliciting in-context learning in vision-language models for videos through curated data distributional properties. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 20416–20431, Miami, Florida, USA, November 2024. Association for Computational Linguistics.

[2] Yuwei Bao, **Keunwoo Peter Yu**, Yichi Zhang, Shane Storks, Itamar Bar-Yossef, Alex de la Iglesia, Megan Su, Xiao Zheng, and Joyce Chai. Can foundation models watch, talk and guide you step by step to make a cake? In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12325–12341, Singapore, December 2023. Association for Computational Linguistics.

[3] Shane Storks, **Keunwoo Peter Yu**, Ziqiao Ma, and Joyce Chai. NLP reproducibility for all: Understanding experiences of beginners. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 10199–10219, Toronto, Canada, July 2023. Association for Computational Linguistics.

[4] Yichi Zhang, Jianing Yang, Jiayi Pan, Shane Storks, Nikhil Devraj, Ziqiao Ma, **Keunwoo Peter**

Yu, Yuwei Bao, and Joyce Chai. DANLI: Deliberative agent for following natural language instructions. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 1280–1298, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics.

Other Manuscripts

[1] **Keunwoo Peter Yu**. Constructing temporal dynamic knowledge graphs from interactive text-based games. *arXiv preprint arXiv:2311.01928*, 2023.

[2] Yichi Zhang, Jianing Yang, **Keunwoo Peter Yu**, Yinpei Dai, Shane Storks, Yuwei Bao, Jiayi Pan, Nikhil Devraj, Ziqiao Ma, and Joyce Chai. Seagull: An embodied agent for instruction following through situated dialog. In *Alexa Prize SimBot Challenge Proceedings*, 2023. **Winner of the Alexa Prize SimBot Challenge**.

[3] **Keunwoo Peter Yu** and Yi Yang. One model to recognize them all: Marginal distillation from ner models with different tag sets. *arXiv preprint arXiv:2004.05140*, 2020.

Patents

[1] Yi Yang and **Keunwoo Peter Yu**. System, method, and computer program for obtaining a unified named entity recognition model with the collective predictive capabilities of teacher models with different tag sets using marginal distillation, November 1 2022. US Patent 11,487,944.

Industry Experience

2025–present **Applied Scientist**, *Wayve*, Sunnyvale, CA
- Researched foundation models for driving.

2024–2024 **Summer Research Intern**, *Toyota Research Institute*, Los Altos, CA
- Researched vision-language models for long-form videos.
- Conducted large-scale, distributed training of vision-language models on video-text data using SageMaker.
- Ran extensive evaluations on the current SOTA vision-language models.

2019–2020 **NLP Researcher**, *ASAPP*, New York, NY
- Developed a novel algorithm to train a unified named entity recognition model from multiple resources with different tag sets.
- Architected an internal natural language processing library that streamlined NLP model deployment to production.

2017–2019 **Lead Software Engineer**, *ASAPP*, New York, NY
- Refactored a monolithic server into microservices communicating over protobuf-based RPC calls and an AMQP message queue to scale real-time chat message delivery and support more resilient websocket connection handling.
- Improved the deployment system to deliver true continuous integration and deployment.
- Developed and deployed a set of microservices to support the "omni-channel" chat experience where customers can chat using their preferred communication channel, e.g. Apple Business Chat, push notifications, SMS, etc.
- Conducted research in proprietary, multi-lingual customer intent classifier.
- Mentored junior developers through the onboarding process.

2016–2017 **Software Engineer**, *Etsy*, New York, NY
- Developed and launched Guest Checkout, a critical project that added millions of dollars to Etsy's global merchandise sales.

2016–2016 **Software Engineer**, *DWNLD*, New York, NY

- Developed and maintained an API written in Scala (Finch, Finagle) that allowed mobile apps to make purchases using Google In-app Billing.
- Extended a Rails API to add new features such as in-app authentication.
- Wrote and deployed a number of big data analytics job using Celery, RabbitMQ, Spark and HBase.
- Developed and maintained an internal web app (Scala Finch, Postgres, Slick) that allowed employees to tag images with predefined characteristics to construct a data set that would be used to train a machine learning model.

2013–2015 **Software Engineer**, *AppNexus*, New York, NY

- Implemented a Kubernetes-like system to integrate Docker to AppNexus' cloud system.
- Developed, automated and maintained a system that rapidly generates clones of the production system that can be used as sand boxes by developers and external clients.
- Developed a proprietary configuration system similar to Puppet's Hiera.
- Managed a summer intern to successfully complete his project.
- Developed and maintained tools to manage AppNexus' cloud system, including a job scheduler and a server management web application.

Technical Skills

- Programming Languages: Python, Golang, Typescript, Javascript, C#, C/C++, Scala, Java, PHP, SQL
- Libraries: PyTorch, Hugging Face Transformers, PyTorch Lightning, TensorFlow, \psi, spaCy, NLTK, pytest, NumPy, scikit-learn, pandas, matplotlib, seaborn
- Developer Tools: Git, LaTeX, Slurm, SageMaker, Docker, GitHub Actions, GitLab Pipelines, Jupyter

Honors and Awards

2023 **First-place Winner, Alexa Prize SimBot Challenge**, *Amazon*

Academic Appointments

2021–2022 **Graduate Student Instructor**, *University of Michigan*, Ann Arbor, MI

- Assisted the professor in running a graduate-level course in natural language processing with 100+ students for two semesters.
- Supported students through office hours.
- Designed assignments and supervised a team of graders.
- Gave guest lectures on specialized topics such as graph neural networks and the SLURM Workload Manager.

Service

2022-present **Peer Reviewer**

COLING 2022; ACL-Industry 2023; EMNLP 2024; COLING 2025; NeurIPS 2025; EMNLP-Industry 2025