Wenyan Li she/her
https://wenyanli.org
Email:weli@di.ku.dk
Mobile : +4531685111

Education

- **University of Copenhagen** — Copenhagen, Denmark
  Ph.D. Fellow in Computer Science — *July 2022 – Present*
- **University of Maryland, College Park** — College Park, MD
  Master of Science in Electrical Engineering (Thesis track) — *Aug 2018*
  - Relevant Coursework: Computational Linguistics, Machine Learning, Database Design, Convex Optimization, Computer Processing of Pictorial Information
- **Northwestern Polytechnical University** — Xian, China
  Bachelor of Engineering in Electrical Engineering (Outstanding Graudate) — *June 2018*

Research Experience

- **COASTAL NLP & LAMP group, University of Copenhagen** — Copenhagen, Denmark
  Multi-modal Learning & Language Modeling; Supervised by Anders Søgaard — *July 2022 – Present*
  - *Retrieval-robust Image Captioning*
    Analyze the robustness of existing **retrieval-augmented image captioning models**, identify their failure modes and build a retrieval-robust image captioning model. (ACL 2024 main conference, oral)
  - *Dynamic Data Curation*
    Improve state-of-the-art image captioning models such as BEiT3 and BLIP with three **dynamic data curation** approaches, including the complete removal of a sample, caption replacement, or image replacement via a text-to-image generation model (EACL 2024 main conference, oral)
  - *Culture awareness multimodal dataset and probing*
    Build **culture-aware multimodal dataset** and probing large language models and vision language models for their bias and understanding towards culture-related concepts (under review)
  - *PIXEL-based Language Modeling*
    Design and pretrain **tokenization-free pixel-based language models** (ViT-based) for generative language modeling and experiment with different masking strategies for masked language modeling (work in progress)
- **CLIP Lab, University of Maryland, College Park** — *Master's Thesis Research* — College Park, MD
  NLP; Supervised by Jordan Boyd-Graber — *Sep 2017 – Aug 2018*
  - *Sentence-final Verb Prediction*
    Processed news corpora of more than 20 million sentences in German and Japanese and trained an **end-to-end and incremental verb prediction** model for reducing translation latency in simultaneous machine translation, improving verb prediction in both German and Japanese; implemented **synonym-aware verb prediction** for German and provided interpretable prediction visualization

Employment

- **Sensetime** — ***Senior NLP researcher, Artificial General Intelligence*** — Shanghai, China
  NLP & Multimodality — *Sept 2021 – Jun 2022*
  - Multi-domain knowledge-enhanced question answering and dialogue system based on dense passage retrieval
  - Multi-modal learning with **pretraining, adapter-based finetuning and prompting**
- **Comcast Applied AI** — ***Senior Research Engineer, Machine Learning*** — Washington, D.C.
  NLP & Data Science — *Jan 2019 – July 2021*

- *User-feedback Automatic Learning Platform*
  Designed and productionized an **unsupervised auto-annotation and active learning pipeline** which models user behaviours to automatically produce reliable training data, identify errors in speech recognition and NLP systems, and suggest corrections. A reduction of 0.907 in word error rate is achieved after applying the auto-suggested fixes. This platform creates a **human-in-the-loop** feedback system for our speech and NLP models to perform more robustly by learning from user behaviors. The platform greatly reduces and prioritizes human annotation effort and improves user experiences, e.g. click through rate. *(project lead and engineering owner)*
- *User behaviour data processing*
  Built a Spark-based automatic data processing pipeline, which cleans and preprocesses 40M raw user logs daily and creates user voice search sessions
- *Multilinguality and Fairness*
  Analyzed fairness of NLP and ASR modules in voice search including user preferences, click through rate and system response time for weak-labeled user groups, taking consideration of multilingual and demographic factors

Publications & Patents

- **Li, W.**, Zhang, X., Li, J., Peng, Q., Tang, R., Zhou, L., ... & Elliott, D. (2024). FoodieQA: A Multimodal Dataset for Fine-Grained Understanding of Chinese Food Culture. *arXiv preprint arXiv:2406.11030.*
- **Li, W.**, Li, J., Ramos, R., Tang, R., & Elliott, D. (2024). Understanding Retrieval Robustness for Retrieval-Augmented Image Captioning. *In ACL 2024.*
- Cao, Y., **Li, W.**, Li, J., Yuan, Y., & Hershcovich, D. (2024). Exploring Visual Culture Awareness in GPT-4V: A Comprehensive Probing. *arXiv preprint arXiv:2402.06015.*
- Li, W., Lotz, J. F., Qiu, C., & Elliott, D. (2024). The Role of Data Curation in Image Captioning. *In EACL 2024.*
- **Li, W.**, Li, D., Li, W., Wang, Y., Jie, H., & Zhong, Y. (2023). MAP: Low-data Regime Multimodal Learning with Adapter-based Pre-training and Prompting. *In Proceedings of the 2023 CLASP Conference on Learning with Small Data (LSD).*
- Tang, R., Kumar, K., Xin, J., Vyas, P., **Li, W.**, Yang, G., ... & Lin, J. (2022). Temporal early exiting for streaming speech commands recognition. *In ICASSP 2022.*
- Tang, R., Kumar, K., Chalkley, K., Xin, J., Zhang, L., textbfLi, W., ... & Lin, J. (2021, November). Voice query auto completion. *In EMNLP, 2021.*
- **Li, W.**, Grissom II, A., & Boyd-Graber, J. (2020). An attentive recurrent model for incremental prediction of sentence-final verbs. *In Findings of EMNLP, 2020.*
- **Li, W.** & Ture, F. (2020). Auto-annotation for voice-enabled entertainment systems. *In SIGIR, 2020.*
- **Li, W.**, Ture, F., Casillas, J., & Des Jardins, G. T. (2022). Systems and Methods for Training Voice Query Models". U.S. Patent Application 17/383,236, filed January 27, 2022.

Service & Professional Activities

- **Reviewer:** ACL, EACL, EMNLP, NeurIPS
- **Teaching Assistant:** Introduction to Data Science, Advanced Topics in Natural Language Processing

Programming Skills

- **Languages**: Python, SQL, C++, MATLAB
- **Frameworks and Tools**: PyTorch, Tensorflow, AWS, PySpark, Git, Docker, Snorkel, Scikit-Learn, Latex