

# DI WU

+1 (310)498-8193 ◊ [diwu@cs.ucla.edu](mailto:diwu@cs.ucla.edu) ◊ <https://xiaowu0162.github.io/>

## EDUCATION

### University of California, Los Angeles

*Ph.D. in Computer Science*

*UCLANLP Group (with Prof. Kai-Wei Chang)*

Los Angeles, CA, USA

09/2022 - present

- Overall GPA: **4.000/4.000**

- Research Interests: retrieval-augmented language models, text generation evaluation, keyphrase generation

### University of California, Los Angeles

*B.S. in Computer Science, Summa Cum Laude*

Los Angeles, CA, USA

08/2018 - 06/2022

- Overall GPA: **3.973/4.000**

- HSSEAS Dean's Honor List (8 times)

## RESEARCH EXPERIENCE

### UCLA NLP Lab

*Student Researcher, Supervisor: Prof. Kai-Wei Chang*

Los Angeles, CA, USA

02/2021 - present

- Research on neural keyphrase generation methods and their evaluation.
- Research on building robust and reliable evaluation metrics for text generation.
- Collaborate with Taboola. Using an internal dataset as the use case, experiment with keyphrase generation and adaptation methods and compare with the company's working pipeline.

### The Ozcan Research Group, UCLA

*Student Research Assistant, Supervisor: Prof. Aydogan Ozcan*

Los Angeles, CA, USA

09/2019 - 06/2022

- Researched on virtual staining of skin tissues with an emphasis on Basal Cell Carcinoma. Designed methods to improve both single-image quality and temporal coherence of the predictions.
- Explored dataset engineering and data augmentation methods to mitigate class imbalance.

## PUBLICATIONS

Li, Y., Gu, J. C., **Wu, D.**, Chang, K. W., Peng, N., 2024. BRIEF: Bridging Retrieval and Inference via Multi-hop Reasoning and Compression. *In preparation*.

**Wu, D.**, Wang, H., Yu, W., Zhang, Y., Chang, K. W., Yu, D., 2024. LongMemEval: Benchmarking Chat Assistants on Long-Term Interactive Memory. *Preprint*.

**Wu, D.**, Gu, J. C., Yin, F., Peng, N., and Chang, K. W., 2024. Synchronous Faithfulness Monitoring for Trustworthy Retrieval-Augmented Generation. *EMNLP 2024*.

Wan, Y., **Wu, D.**, Wang, H., and Chang, K. W., 2024. The Factuality Tax of Diversity-Intervened Text-to-Image Generation: Benchmark and Fact-Augmented Intervention. *EMNLP 2024*.

**Wu, D.**, Shen, X., and Chang, K. W., 2024. MetaKP: On-Demand Keyphrase Generation. *Findings of the ACL: EMNLP 2024*.

**Wu, D.**, Yin, D., and Chang, K. W., 2024. KP Eval: Towards Fine-grained Semantic-based Keyphrase Evaluation. *Findings of the ACL: ACL 2024*.

**Wu, D.**, Ahmad, W. U., Zhang, D., Ramanathan, M. K., and Ma, X., 2023. Repoformer: Selective Retrieval for Repository-level Code Completion. *ICML 2024*.

**Wu, D.**, Ahmad, W. U., and Chang, K. W., 2023. On Leveraging Encoder-only Pre-trained Language Models for Effective Keyphrase Generation. *LREC-COLING 2024*.

Li, Y., Pillar, N., Li, J., Liu, T., **Wu, D.**, Sun, S., Ma, G., de Haan, K., Huang, L., Zhang, Y. and Hamidi, S., 2024. Virtual histological staining of unlabeled autopsy tissue. *Nature Communications*, 15(1), p.1684.

**Wu, D.**, Ahmad, W. U., and Chang, K. W., 2023. Rethinking Model Selection and Decoding for Keyphrase Generation with Pre-trained Sequence-to-Sequence Models. *EMNLP 2023*.

Kung, P., Yin F., **Wu, D.**, Chang, K. W., and Peng N., 2023. Active Instruction Tuning: Improving Cross-Task Generalization by Training on Prompt Sensitive Tasks. *EMNLP 2023*.

**Wu, D.**, Ahmad, W. U., and Chang, K. W., 2023. Pre-trained Language Models for Keyphrase Generation: A Thorough Empirical Study. *Preprint*.

**Wu, D.**, Ahmad, W. U., Dev, S., and Chang, K. W., 2023. Representation Learning for Resource-Constrained Keyphrase Generation. *Findings of the ACL: EMNLP 2022*.

Li, J., Garfinkel, J., Zhang, X., **Wu, D.**, Zhang, Y., de Haan, K., Wang, H., Liu, T., Bai, B., Rivenson, Y., Rubinstein, G., Scumpia, P., and Ozcan, A., 2023. Biopsy-free in vivo virtual histology of skin using deep learning. *Light: Science & Applications*, 10(1), 1-22.

## INTERNSHIP EXPERIENCE

---

### Tencent AI Lab

Bellevue, WA, USA

*Research Intern; Mentors: Hongwei Wang, Wenhao Yu*

06/2024 - 09/2024

- Researched on long-term memory for chat assistants and built a large-scale high-quality evaluation benchmark.
- Designed indexing strategies for improving long-term memory retrieval performance.

### AWS AI Labs

New York, NY, USA

*Applied Scientist Intern; Mentors: Wasi Ahmad, Dejiao Zhang*

06/2023 - 09/2023

- Researched on improving retrieval-augmented code language models for repository-level code completion.
- Formulated the task of selective retrieval-augmented infilling. Designed approaches from the perspective of in-repository code retrievers and the code generator models.
- The designed model achieves no performance loss with only 10% of retrieval budget.

### Microsoft Research Asia

Beijing, China

*Research Intern; Mentor: Ning Shang*

04/2021 - 09/2021

- Participated in the research project "Sparse Analysis". Compared over 10 model compression methods including pruning, quantization, and knowledge distillation on MobileNetV2 and Transformers. For pruning, the analysis involved criterion, sparsity scheduling, and learning scheduling.
- Participated in the open source AutoML project [NNI](#). Contributed over 5000 lines of code including a hyperparameter optimization benchmark, a Transformer pruner, and three examples.

### NewsBreak

Beijing, China

*Natural Language Processing Intern*

07/2020 - 10/2020

- Worked on a hierarchical multi-label classification problem with 268 categories. Improved the f1-score by 49% with multiple statistical and deep learning methods.
- Improved the performance of model pre-training, fine-tuning, and online serving pipelines.

## TEACHING

---

### Teaching Assistant

- UCLA CS 33, Introduction to Computer Organization, Fall 2023, with Prof. Glenn Reinman.
- UCLA CS 33, Introduction to Computer Organization, Spring 2024, with Prof. Glenn Reinman.
- UCLA CS 33, Introduction to Computer Organization, Fall 2024, with Prof. Tony Nowatzki.

## SERVICES

---

- Reviewer: ACL 2023, EMNLP 2023, AAAI 2023-2025, JAIR, ACL Rolling Review.