

# DI WU

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

## EDUCATION

---

University of California, Los Angeles	Los Angeles, CA
<i>Ph.D. in Computer Science</i>	09/2022-Present
<i>UCLANLP Group (with Prof. Kai-Wei Chang)</i>	

- Research Interests: keyphrase generation, text generation evaluation, human-in-the-loop NLP

---

University of California, Los Angeles	Los Angeles, CA
<i>B.S. in Computer Science, Summa Cum Laude</i>	08/2018-06/2022

- Overall GPA: 3.973/4.000
- HSSEAS Dean's Honor List (8 times)

## RESEARCH EXPERIENCE

---

UCLA NLP Lab	Los Angeles, CA
<i>Student Researcher, Supervisor: Prof. Kai-Wei Chang</i>	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	Los Angeles, CA
<i>Student Research Assistant, Supervisor: Prof. Aydogan Ozcan</i>	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

---

Wu, D., Yin, D, and Chang, K. W. (2023). KPEval: Towards Fine-grained Semantic-based Evaluation of Keyphrase Extraction and Generation Systems. *ArXiv 2023*.

Wu, D., Ahmad, W. U., and Chang, K. W. (2022). Pre-trained Language Models for Keyphrase Generation: A Thorough Empirical Study. *ArXiv 2022*.

Wu, D., Ahmad, W. U., Dev, S., and Chang, K. W. (2022). 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. (2021). Biopsy-free in vivo virtual histology of skin using deep learning. *Light: Science & Applications*, 10(1), 1-22.

## INTERNSHIP EXPERIENCE

---

Microsoft Research Asia	Beijing, China
<i>Research Intern, Mentor: Ning Shang</i>	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

*Natural Language Processing Intern*

Beijing, China

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.

## LANGUAGE AND SKILLS

---

Programming	C, C++, Python, shell, OCaml, Scheme, Verilog, MATLAB
Machine Learning	NumPy, TensorFlow, PyTorch, pandas, scikit-learn, Matplotlib
Language	Mandarin (Native), English (Professional), Japanese (Beginner)
GRE	337 (V: 167/98%; Q: 170/96%; AW: 4.5/81%)
Hobbies	Air Traffic Control, Flight Simulation