

Jerry Wei

jerrywei@alumni.stanford.edu
<https://www.jerrywei.net>

Experience

Jul. 2023 - *Google DeepMind*
Research Engineer.
LLM alignment and reasoning.

Nov. 2022 - Jul. 2023 *Google Brain*
Student Researcher.
Hosts: Da Huang, Hanxiao Liu, Weizhe Hua.

Jun. - Sep. 2022 *Meta*
Software Engineering Intern.
Messenger for Android team.

Education

Sep. 2021 - *Stanford University*
Bachelor of Science, Computer Science.
GPA: 3.89/4.00.

Publications

Long-form factuality in large language models.

J. Wei, C. Yang, X. Song, Y. Lu, N. Hu, J. Huang, D. Tran, D. Peng, R. Liu, D. Huang, C. Du, and Q. V. Le.
arXiv preprint, 2024.

Simple synthetic data reduces sycophancy in large language models.

J. Wei, D. Huang, Y. Lu, D. Zhou, and Q. V. Le.
arXiv preprint, 2023.

Symbol tuning improves in-context learning in language models.

J. Wei, L. Hou, A. Lampinen, X. Chen, D. Huang, Y. Tay, X. Chen, Y. Lu, D. Zhou, T. Ma, and Q. V. Le.
Conference on Empirical Methods in Natural Language Processing, 2023.

Larger language models do in-context learning differently.

J. Wei, J. Wei, Y. Tay, D. Tran, A. Webson, Y. Lu, X. Chen, H. Liu, D. Huang, D. Zhou, and T. Ma.
arXiv preprint, 2023.

Generative image translation for data augmentation in colorectal histopathology images.

J. Wei, A. Suriawinata, L. Vaickus, B. Ren, X. Liu, J. Wei, and S. Hassanpour.
ML4H Workshop at NeurIPS, 2019 (spotlight talk).

Honors

Regeneron Science Talent Search (STS) Semifinalist (top 300 projects), 2021.
National Merit Scholarship Semifinalist & Leidos Corporation Scholarship Winner, 2021.
Spotlight presentation at the NeurIPS ML4H Workshop, 2019.

Other

Reviewing: ACL '23, EACL'22, WACV '22, EMNLP '21, EACL '21, NeurIPS ML4H '20.