Anthony Chen

✉ anthony.chen@uci.edu
🖥 anthonywchen.github.io

My current research focuses on large language models and aims to create benchmarks to evaluate their many diverse, emergent capabilities and methodologies to address their weaknesses (*e.g.,* lack of attribution and hallucinations).

Education

2018-2023	**Ph.D. in Computer Science**, *University of California, Irvine*. Advised by Sameer Singh.
2016-2018	**M.S. in Computer Science**, *University of California, Irvine*.
2012-2016	**B.S. in Computer Science**, *University of California, Davis*.

Professional History

Summer 2022	**Research Scientist, Intern**, *Google*, *Mountain View, California*, Hosted by Kelvin Guu and Hongrae Lee. Developed novel models for editing large language model generations for improved facutality.
Summer 2021	**Research Scientist, Intern**, *Verneek*, *New York, New York*, Hosted by Sam Anzaroot and Nasrin Mostafazadeh. Worked on retrieval-augmented methods for automatic construction of knowledge bases.
Summer 2020	**Research Scientist, Intern**, *Apple*, *Remote*, Hosted by Pallavi Gudipati. Developed datasets to measure the robustness of Siri's entity linking systems to bias.
Summer 2018	**Data Scientist, Intern**, *Ancestry*, *San Francisco, California*.
Summer 2017	**Data Scientist, Intern**, *Allstate*, *Menlo Park, California*.
2014	**Software Engineer, Intern**, *Intel*, *Folsom, California*.

Publications

2023	**PURR: Efficiently Editing Language Model Hallucinations by Denoising Language Model Corruptions**, *Anthony Chen*, Panupong Pasupat, Sameer Singh, Hongrae Lee, and Kelvin Guu. arXiv
2023	**RARR: Researching and Revising What Language Models Say, Using Language Models**, Luyu Gao, Zhuyun Dai, Panupong Pasupat, *Anthony Chen*, Arun Tejasvi Chaganty, Yicheng Fan, Vincent Y. Zhao, Ni Lao, Hongrae Lee, Da-Cheng Juan, and Kelvin Guu. Association for Computational Linguistics (ACL)
2021	**Entity-Based Knowledge Conflicts in Question Answering**, Shayne Longpre, Kartik Perisetla, *Anthony Chen*, Nikhil Ramesh, Chris DuBois, and Sameer Singh. Empirical Methods in Natural Language Processing (EMNLP)

2021 **Evaluating Entity Disambiguation & the Role of Popularity in Retrieval-Based NLP**,
Anthony Chen, Pallavi Gudipati, Shayne Longpre, Xiao Ling, and Sameer Singh.
Association for Computational Linguistics and the International Joint Conference on Natural Language Processing (ACL-IJCNLP)

2020 **MOCHA: A Dataset for Training and Evaluating Generative Reading Comprehension Metrics**,
Anthony Chen, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.
Empirical Methods in Natural Language Processing (EMNLP)

2019 **Evaluating Question Answering Evaluation**,
Anthony Chen, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.
Machine Reading for Question Answering Workshop @ Empirical Methods in Natural Language Processing (EMNLP) Best Paper

Awards

2022 **Noyce Fellowship**, *Assessing and Addressing the Societal Implications of Using Neural Retrievers in Real-World NLP.*
Awarded to **Anthony Chen** and Sameer Singh.

2019 **Best Paper**, *Evaluating Question Answering Evaluation*,
EMNLP Workshop on Machine Reading and Question Answering (MRQA).
Awarded to **Anthony Chen**, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.

Invited Talks

2023 **UC Irvine AI/ML Seminar**, *RARR: Researching and Revising What Language Models Say, Using Language Models*, Irvine, CA.

2020 **Apple**, *Developing Learned Metrics for Generative Reading Comprehension*, Virtual.

2020 **MRQA Workshop @ EMNLP**, *Evaluating Question Answering Evaluation*, Hong Kong.

Professional Service

2021 **Reviewer**, *Open-Domain Reasoning Under Multi-Modal Settings Workshop @ Computer Vision and Pattern Recognition (CVPR).*

2021 **Site Developer**, *Automated Knowledge Base Construction (AKBC).*

2021 **Reviewer**, *Empirical Methods in Natural Language Processing (EMNLP).*

2018 **Reviewer**, *Empirical Methods in Natural Language Processing (EMNLP).*