

Anthony Chen

✉ anthony.chen@uci.edu
📄 [anthonywchen.github.io](https://github.com/anthonywchen)

Education

- 2018-2023 **Ph.D. in Computer Science**, *University of California, Irvine*,
Factuality and Large Language Models: Evaluation, Characterization, and Correction.
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

- 2023 - Now **Software Engineer**, *Google DeepMind*.
- Summer 2022 **Research Scientist, Intern**, *Google Research, Mountain View, California*.
Hosted by Kelvin Guu and Hongrae Lee
- Summer 2021 **Research Scientist, Intern**, *Verneek, New York, New York*.
Hosted by Sam Anzaroot and Nasrin Mostafazadeh
- Summer 2020 **Research Scientist, Intern**, *Apple, Cupertino, California*.
Hosted by Pallavi Gudipati
- Summer 2018 **Data Scientist, Intern**, *Ancestry, San Francisco, California*.
- Summer 2017 **Data Scientist, Intern**, *Allstate, Menlo Park, California*.
- Summer 2014 **Software Engineer, Intern**, *Intel, Folsom, California*.

Publications

- 2024 **Can Long-Context Language Models Subsume Retrieval, RAG, SQL, and More?**,
Jinhyuk Lee, [Anthony Chen](#), Zhuyun Dai, Dheeru Dua, Devendra Singh Sachan,
Michael Boratko, Yi Luan, Sébastien M. R. Arnold, Vincent Perot, Siddharth Dalmia,
Hexiang Hu, Xudong Lin, Panupong Pasupat, Aida Amini, Jeremy R. Cole, Sebastian
Riedel, Iftekhhar Naim, Ming-Wei Chang, Kelvin Guu.
arXiv
- 2023 **The Data Provenance Initiative**,
Shayne Longpre, Robert Mahari, [Anthony Chen](#), Niklas Muennighoff, Kartik Perisetla,
William Brannon, Jad Kabbara, Emily Reif, Luis Villa, Sara Hooker.
Generative AI and Law Workshop @ International Conference on Machine Learning (ICML)
Spotlight
- 2023 **PURR: Efficiently Editing Language Model Hallucinations by Denoising Language Model Corruptions**,
[Anthony Chen](#), Panupong Pasupat, Sameer Singh, Hongrae Lee, 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, 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, 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, Sameer Singh.
Association for Computational Linguistics (ACL)
- 2020 **MOCHA: A Dataset for Training and Evaluating Generative Reading Comprehension Metrics**,
[Anthony Chen](#), Gabriel Stanovsky, Sameer Singh, Matt Gardner.
Empirical Methods in Natural Language Processing (EMNLP)
- 2019 **Evaluating Question Answering Evaluation**,
[Anthony Chen](#), Gabriel Stanovsky, Sameer Singh, 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 for Question Answering (MRQA).
Awarded to **Anthony Chen**, Gabriel Stanovsky, Sameer Singh, and Matt Gardner.

Invited Talks

- 2024 **USC ISI NLG Seminar**, *The Data Provenance Initiative: A Large Scale Audit of Dataset Licensing & Attribution in AI*, Los Angeles, CA.
- 2023 **UC Irvine AI/ML Seminar**, *RARR: Researching and Revising What Language Models Say, Using Language Models*, Irvine, CA.
- 2020 **MRQA Workshop @ EMNLP**, *Evaluating Question Answering Evaluation*, Hong Kong, SAR China.

Research Services

- 2023 **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)*.