Kellen Cheng

✉ kellentan@princeton.edu ☐ 4074094268 ⌂ Personal Website ♗ kellentan ♟ U.S. Citizen

Education

Princeton University — **Present**
Ph.D. Electrical and Computer Engineering, GPA: 4.000 — Princeton, NJ

Princeton University — **Aug. 2022- Nov. 2024**
M.A. Electrical and Computer Engineering, GPA: 4.000 — Princeton, NJ

University of California, Los Angeles (UCLA) — **Sept. 2018 - Jun. 2022**
B.S. Electrical Engineering, Technical Breadth Computer Science, GPA: 3.926 — Los Angeles, CA

Publications

- **Kellen Tan Cheng**, Anna Lisa Gentile, Chad DeLuca, Guang-Jie Ren, "STAR: Self-Automated Back-Querying for Production Data Generation", NAACL 2025 Submission.
- Chad DeLuca, Anna Lisa Gentile, Shubhi Asthana, Bing Zhang, Pawan Chowdhury, **Kellen Tan Cheng**, Basel Shbita, Pengyuan Li, Guang-Jie Ren, Sandeep Gopisetty, "OneShield - The Next Generation of LLM Guardrails", WWW 2025 Industry Track Submission.
- **Kellen Tan Cheng**, Anna Lisa Gentile, Pengyuan Li, Chad DeLuca, Guang-Jie Ren, "Don't Be My Doctor! Recognizing Healthcare Advice in Large Language Models", EMNLP Industry Track 2024.
- **Kellen Tan Cheng**, Suma Bhat, "No Context Needed: Contextual Quandary In Idiomatic Reasoning With Pre-Trained Language Models", NAACL Main 2024.
- Ziheng Zeng, **Kellen Tan Cheng**, Srihari Venkat Nanniyur, Jianing Zhou, and Suma Bhat, "IEKG: A Commonsense Knowledge Graph for Idiomatic Expressions", EMNLP Main 2023.
- Jean-Luc Margot, Megan G. Li, Pavlo Pinchuk, Nathan Myhrvold, Larry Lesyna, Lea E. Alcantara, Megan T. Andrakin, Jeth Arunseangroj, Damien S. Baclet, Madison H. Belk, Zerxes R. Bhadha, Nicholas W. Brandis, Robert E. Carey, Harrison P. Cassar, Sai S. Chava, Calvin Chen, James Chen, **Kellen T. Cheng**, Alessia Cimbri, Benjamin Cloutier, Jordan A. Combitsis, Kelly L. Couvrette, Brandon P. Coy, Kyle W. Davis, Antoine F. Delcayre, Michelle R. Du, Sarah E. Feil, Danning Fu, Travis J. Gilmore, Emergy Grahill-Bland, Laura M. Iglesias, Zoe Juneau, Anthony G. Karapetian, George Karfakis, Christopher T. Lambert, Eric A. Lazbin, Jian H. Li, Zhuofu (Chester) Li, Nicholas M. Liskij, Anthony V. Lopilato, Darren J. Lu, Detao Ma, Vedant Mathur, Mary H. Minasyan, Maxwell K. Muller, Mark T. Nasielski, Janice T. Nguyen, Lorraine M. Nicholson, Samantha Niemoeller, Divij Ohri, Atharva U. Padhye, Supreethi V. Penmetcha, Yugantar Prakash, Xinyi (Cindy) Qi, Vedant Sahu, Joshua A. Scally, Zefyr Scott, Trevor J. Seddon, Lara-Lynn V. Shohet, Anchal Sinha, Anthony E. Sinigiani, Jiuxu Song, Spencer M. Stice, Andria Uplisashvili, Krishna Vanga, Amaury G. Vazquez, George Vetushko, Valeria Villa, Maria Vincent, Ian J. Waasdorp, Ian B. Wagaman, Amanda Wang, Jade C. Wight, Ella Wong, Natsuko Yamaguchi, Zijin Zhang, Junyang Zhao, and Ryan S. Lynch, "A Search for Technosignatures Around 11,680 Stars with the Green Bank Telescope at 1.15-1.73 GHz", The Astronomical Journal 2023.
- **Kellen Cheng**, Kunakorn Atchaneeyasakul, Zeid Barakat, David Liebeskind, and Fabien Scalzo, "CT Perfusion Imaging of the Brain with Machine Learning", ISVC 2021.

Experience

Princeton: Graduate Research Assistant — **Nov. 2022 - Present**
Advisor: Dr. Suma Bhat — Princeton, NJ

- Advanced the performance and understanding of figurative language reasoning capabilities of pre-trained language models ranging from 110M to 7B parameters.
- Designed and implemented an end-to-end two-mask infilling fine-tuning framework for idiomatic knowledge injection with compact language models (e.g. BART) using the IEKG dataset.
- Coupled my fine-tuning framework with transfer learning from the MNLI dataset, achieving a new state-of-the-art performance of 83.75% on the IMPLI benchmark, an improvement of 12% compared to previous state-of-the-art.

- Validated my framework through an ablation study, demonstrating gains of 23.08% compared with off-the-shelf models and 5.79% compared with vanilla transfer learning.
- Formulated and conducted a comprehensive context perturbation study for off-the-shelf language models (ranging from BART to Mistral-7B), uncovering that they are negatively affected by the context (up to 3.89% better *without* context).
- Published my works as a first-author NAACL 2024 main conference paper and a second-author EMNLP 2023 main conference paper.

Princeton: Graduate Teaching Assistant
Jan. 2024 - Dec. 2024

SML/PHI 354: Prof. Sarah-Jane Leslie
Princeton, NJ

- Sole assistant instructor (TA) responsible for 56 (Spring 2024) and 47 students (Fall 2024) enrolled in SML/PHI 354 (Artificial Intelligence: A Hands-On Introduction from Basics to ChatGPT).
- Lead 4 (Fall 2024) and 6 (Spring 2024) weekly precept sections (50 min. each) to review fundamental course material, spanning artificial neural networks, convolutional neural networks, and transformer models.
- Held weekly office hours and was responsible for handling assignments, final projects, and exams.

IBM Research: Research Scientist Intern
Jun. 2024 - Sept. 2024

Advisor: Dr. Anna Lisa Gentile
San Jose, CA

- Developed a lightweight BART-Large detector model for health advice as a summer intern in the OneShield AI safety guardrails team at IBM Research-Almaden.
- Manually synthesized a new, gold standard, evaluation benchmark (HeAL) for health advice guardrails, containing 402 samples from a mix of professional, medical, and conversational sources on Common Crawl.
- Designed an on-demand human-in-the-loop system based on grouped clustering for sparse human annotation of synthetic production data at scale.
- Created a new framework to generate synthetic production data for health advice guardrailing, resulting in state-of-the-art performance of 85.07% from a BART-Large detector model, an increase of 3.48% compared to GPT-4o.
- Validated my findings by running additional experiments on proprietary IBM data, demonstrating real world performance gains with a false positive rate reduction of up to 3.18%.
- Programmed and deployed an internal Rest API module to automate the loading and evaluation process of all internal use datasets for the OneShield team.
- Published my work as a first-author EMNLP 2024 industry track paper, and submitted works to both the NAACL 2025 main conference (first-author) and the WWW 2025 industry track.
- Defended and filed a patent detailing a continual learning framework with model version-control and knowledge distillation for health advice detector development (bar date of April 2025).
- Will be returning to IBM Research as a summer 2025 intern!

UCLA: Undergraduate Research Assistant
Jan. 2021 - Jun. 2022

Advisor: Prof. Fabien Scalzo
Los Angeles, CA

- Developed a proof-of-concept end-to-end framework utilizing machine learning models (random forests, kernel ridge regression, etc.) to automate the prediction of CT perfusion parameter maps.
- Incorporated training data from 50 ischemic stroke patients at UCLA's Ronald Reagan hospital to develop our machine learning CT perfusion framework, achieving structural similarities (SSIM) of 84% after motion correction and filtering.
- Published my findings as a first-author paper at ISVC 2021.

UCLA: ECE Course Grader
Sept. 2021 - Jun. 2022

ECE 102, ECE C143A: Prof. Jonathan Kao
Los Angeles, CA

- Graded weekly assignments for over 100 students enrolled in ECE 102 systems and signals (Fall 2021) and ECE C143A neural signal processing (Spring 2022).
- Provided detailed feedback and additional notes for students on assignments.

PerfuseAI: Machine Learning Intern
Jun. 2021 - Sept. 2021

Advisor: Prof. Fabien Scalzo
Remote

- Optimized random forest and kernel ridge regression models to predict CT perfusion blood flow parameters such as rCBF, rCBV, MTT, and TMax.
- Incorporated the end-to-end model pipeline from my ISVC 2021 paper into the PerfuseAI product, with a final objective of deployment for Thailand's hospital network.

- Empirically tested the prototype arterial input function (AIF) image annotation tool designed to be used by physicians on ischemic stroke patients.

Coursework

- UCLA: Probability & Statistics (ECE 131A), Machine Learning (CS M146), Neural Networks & Deep Learning (ECE C147A), Graph Theory (ECE 134), Numerical Computing (ECE 133A).
- Princeton: Natural Language Processing (COS 484), Deep Learning Networks (ECE 571), Graph Signal Processing (ECE 538A), Cryptography (COS 433), Information Theory (COS 585), Complexity Theory (COS 522)

Awards

Toby and Jack Wolf Travel Grant $\qquad$ Oct. 2024
RECIPIENT $\qquad$ PRINCETON, NJ

- Selected to receive the Princeton ECE department travel grant from the Toby and Jack Wolf fund ($500).

Bede Liu Travel Grant $\qquad$ Sept. 2023
RECIPIENT $\qquad$ PRINCETON, NJ

- Selected to receive the Princeton ECE department travel grant from the Bede Liu Fund for Excellence ($500).

Princeton Departmental Fellowship $\qquad$ Aug. 2022 - Jun. 2023
RECIPIENT $\qquad$ PRINCETON, NJ

- Selected to receive the first-year Princeton ECE department fellowship.

UCLA Tau Beta Pi $\qquad$ Mar. 2020 - Jun. 2022
MEMBER $\qquad$ LOS ANGELES, CA

- Inducted into UCLA's Tau Beta Pi, California-Epsilon chapter, awarded to students in the top 12.5% of the junior class.
- Tutored students via weekly help sessions for any undergraduate engineering course.

UCLA Eta Kappa Nu (HKN) $\qquad$ Sept. 2019 - Jun. 2022
MEMBER $\qquad$ LOS ANGELES, CA

- Inducted into UCLA's IEEE-HKN chapter, awarded to students in the top 25% of the junior class.
- Tutored students via weekly help sessions for any undergraduate course in electrical engineering or computer science.

Dean's Honor List $\qquad$ Mar. 2019 - Jun. 2022
RECIPIENT $\qquad$ LOS ANGELES, CA

- Applicable to all quarters at UCLA with at least 15 units taken from Spring 2019 onwards.

Skills

- Python, C++, PyTorch, Tensorflow, Huggingface, Transformers, Jupyter