

Daniel Li

[Identifying Information Redacted]

Email: [write2daniel.li\[at\]gmail.com](mailto:write2daniel.li[at]gmail.com)

Citizenship: United States of America

Current Position(s)

- 2022-present *09S Officer Candidate*, United States Army
- Active Duty, aspiring 11A Infantry
- 2018-Present *Limited Partner*, Dekrypt Capital
- Blockchain venture capital firm, **9 figure** (USD) AUM.
 - Co-investors include a16z, Coinbase Ventures, Arrington XRP.
- 2018-Present *Ph.D Candidate*, Columbia University
- Expected graduation in **Summer 2022**.
 - Natural language processing and applications.
 - Supervised 2 M.Sc. and 2 B.Sc. students and projects.

Areas of Specialization

Artificial Intelligence • Natural Language Processing • Financial Markets

Industry Positions

- 2017-2022 *Co-Founder (2)*, Unnamed Capital Group
- **7 figure** (USD) accounts with a **400% ROI** since inception.
 - Various strategies, instruments, and markets.
- 2021 *Research Intern*, Google Research @ *Cloud AI Translation*
- Confidence score calibration in multilingual quality estimation
- 2020 *Research Intern*, Google Research @ *Google Translate*
- Neural machine translation robustness for automatic speech recognition
- 2019S *Researcher*, Memorial Sloan Kettering Cancer Center
- Research in cancer tumor biopsy images
- 2017S *Research Intern*, NEC Laboratories
- Deep learning on memory recurrent networks.
- 2016S *Software Engineering Intern*, Factual Inc (Foursquare)
- Fuzzy entity resolution of large scale databases

Academic Positions

- 2017-2020 *Graduate Student Instructor*, UC Berkeley and Columbia University
• User Interface Design & Machine Learning for Functional Genomics
- 2015-2018 *Research Assistant*, Pachter Lab at UC Berkeley
• Research in approaches to RNA-sequencing with features in abundance estimation, transcript annotation difficulties, differential expression
- 2016-2018 *Research Assistant*, Rao Group at UC Berkeley
• Investigation on gene feature identification and accurate dimensionality reduction through memory autoencoders

Education

- 2018-2022 PH.D. *Computer Science*, Columbia University.
- 2018-2021 M.PHIL. *Computer Science*, Columbia University.
- 2017-2018 M.Sc. *Electrical Engineering and Computer Science*, University of California, Berkeley.
- 2014-2017 B.Sc. *Electrical Engineering and Computer Science*, University of California, Berkeley.
- 2011-2014 DIPLOMA. La Cañada High School.

Publications

- 2022 **Daniel Li***, Thomas Chen*, Thomas Chen, Albert Tung, Lydia Chilton. *A Hierarchical Dialog Corpus: Designing, Collecting, and Annotating Intermediate Text for Abstract Summarization*. Columbia University, Microsoft, Stanford University. *In Preparation*.
- Daniel Li***, Thomas Chen*, Albert Tung, Lydia Chilton. *Improving Usability of Automatic Summarization for Longform Spoken Dialog*. Columbia University, Microsoft, Stanford University. *In Preparation*.
- 2021 **Daniel Li**, Junpei Zhou, Music Li Lydia Chilton. *Investigating Translation Quality Estimation in Bilingual Models*. Google, Columbia University. *Under Review*.
- Daniel Li***, Thomas Chen*, Albert Tung, Lydia Chilton. *Hierarchical Summarization for Longform Spoken Dialog*. Columbia University, Microsoft, Stanford University. **UIST 2021**.

Jennifer Chen, Faisal Ladhak, **Daniel Li** and Noémie Elhadad. *Incorporating Human Explanations for Robust Hate Speech Detection*. Columbia University. **ACL-IJCNLP 2021 Workshop: Understanding Implicit and Underspecified Language**.

2020 **Daniel Li**, Te I, Naveen Arivazhagan, Colin Cherry, Dirk Padfield. *Sentence Boundary Augmentation For Neural Machine Translation Robustness*. Google Research. **ICASSP 2021**.

Daniel Li*, Jennifer Chen*, Qiang Ma, Andrew Liu, Justin Cheung, Yubin Xie, Herman Gudjonson, Dana Pe'er, Itsik Pe'er. *Counterfactual Hypothesis Testing of Tumor Microenvironment Scenarios Through Semantic Image Synthesis*. Columbia University, Google, Memorial Sloan Kettering Cancer Center. **BioArXiv**.

2019 Justin M. Cheung, Hanan Baker, **Daniel Li**, Daniel Stor, Daniel A. Heller. *Abstract 1719: Drug-loaded porphyrin nanoparticles as a platform for targeted and photodynamic combination therapy*. Memorial Sloan Kettering Cancer Center. **American Association for Cancer Research**.

2017 **Daniel Li**, Asim Kadav. *Adaptive Memory Networks*. University of California, Berkeley, NEC Laboratories America. **ICLR 2018**.

Patents

2020 *System and method for faster interfaces on text-based tasks using adaptive memory networks*. Patent Number: 10853575. Asim Kadav, **Daniel Li**. NEC Laboratories America.

Skills

Programming	Python, PyTorch, TensorFlow, Scikit-Learn, NumPy
Finance	Trading (equities, commodities, digital assets), Technical Analysis, Fundamental Analysis, Algorithmic Trading, Venture Capital (valuation, go to market, etc.), Risk & Compliance
Design/UI/UX	Graphic Design (Figma, Omnigraffle), Ideation, Interviewing, Iterating and Prototyping, Pitch Deck and Presentation Design, Typography
Miscellaneous	Public Speaking, <i>Actual</i> Attention to Detail, Teaching, Open Ended Research, Technical Writing

Awards

2020	Google Cloud Grant
2019	Ph.D. Service Award
2018	ICLR Travel Grant
2017	NVIDIA Grant

Selected Coursework

Theory	Algorithms & Uncertainty, Beyond Worst Case Analysis, Combinatorial Algorithms, Computational Geometry
AI/ML	Artificial Intelligence, Computer Vision, Graphical Models, Deep Learning
Applications	Machine Learning for Functional Genomics, Statistical Genomics, Computational Photography, User Interface Design and Development
Miscellaneous	Entrepreneurship, Cybersecurity Venture Capital, Data and Ethics, Public Speaking, U.S. Education Institutions, Critical Studies in Education, The Southern Border