John Thickstun

+1 (315) 601-5709

jthickstun@stanford.edu

https://johnthickstun.com/

Current Position

- **Stanford University** — Palo Alto, CA
  Postdoctoral Scholar - Stanford Artificial Intelligence Laboratory — *September 2021 - Present*
  - Advised by Percy Liang (joint Computer Science & Statistics).

Education

- **University of Washington** — Seattle, WA
  Ph.D. in Computer Science and Engineering — *August 2021*
  - Advised by Sham M. Kakade (joint Computer Science & Statistics) and Zaid Harchaoui (Statistics).
  - Doctoral Committee: Sham M. Kakade, Zaid Harchaoui, Noah A. Smith, Sewoong Oh, Lalit Jain.
  - Dissertation: Leveraging Generative Models for Music and Signal Processing.
- **University of Washington** — Seattle, WA
  M.Sc. in Computer Science and Engineering — *December 2017*
  - Coursework in Optimization, Algorithms, Learning Theory, Information Theory, Reinforcement Learning.
- **Brown University** — Providence, RI
  Sc.B. Magna cum Laude with Honors in Applied Mathematics — *May 2013*
  - Advised by Björn Sandstede and Eugene Charniak.
  - Coursework in Machine Learning, Computer Vision, Natural Language Processsing, Probability Theory, Stochastic Processes, Real and Complex Analysis, Operator Theory.
- **Hamilton College Bridge Program** — Clinton, NY
  (High School Credit) — *2005-2007*
  - Coursework in Programming Languages, Computer Architecture, Operating Systems, Abstract Algebra.

Major Grants and Awards

- Stanford HAI Google Cloud Credits Grant: $15,000 (2022).
- Neurips Outstanding Paper Award: 6 / 9122 paper submissions (2021).
- Qualcomm Innovation Fellowship: $100,000 (2020).
- NSF Graduate Research Fellowship: $138,000 (2017-2019).
- Brown University Distinguished Senior Thesis Award (university-wide award, 2013).
- Sigma Xi Scientific Honor Society (2013).

Teaching

- Predoctoral Instructor (Instructor of Record). CSE599i: Generative Models (Autumn, 2020).
  - Created a new course offering covering advances in generative modeling from 2010-2020.
  - Developed course materials from scratch including lecture notes, slides, and homework.
  - Top-decile teaching reviews: 4.9/5.0 overall course quality with a 53% response rate (16/30 students).
- Teaching Assistant
  - CSE547: Machine Learning for Big Data (Spring, 2016).
  - CSE546: Machine Learning (Autumn, 2015).
- Guest Lecturer
  - STAT558: Statistical Machine Learning for Data Scientists (Spring 2020).
  - CSE490: An Introduction to Deep Learning (Autumn, 2018).
  - STAT558: Statistical Machine Learning for Data Scientists (Spring 2018).
  - CSE546: Machine Learning (Spring 2018).
  - CSE547: Machine Learning for Big Data (Spring, 2016).

Research Advising and Mentoring

- Xiang Lisa Li, (PhD Student) Stanford University.
  - Lisa and I co-authored a paper appearing at Neurips 2022.
- Vivek Jayaram, (PhD Student) University of Washington.
  - Vivek and I co-authored two papers appearing at ICML 2020, and ICML 2021.
- Harsh Verma, (Undergraduate Student) University of Washington, Class of 2019.
  - Current Masters student at Concordia University.
  - Harsh and I co-authored a paper appearing at ISMIR 2019.

Professional Experience

- **Clear Ventures** - Technology Venture Capital — Palo Alto, CA
  DeepTech Fellow — *September 2022 - Present*
- **Human Exploratorium** - Music Therapy and Recommendation — San Francisco, CA
  Advisor — *May 2021 - Present*
  - Provided introductions, employee mentoring, and advice on hiring and technical strategy.
- **Panjandrum.ai** - Virtual Avatar Music — Seattle, WA
  Advisor — *October 2019 - Present*
  - Provided advice and technical strategy for applications of machine learning to music and visual production.
- **University of Washington** — Seattle, WA
  Graduate Researcher - Computer Science and Engineering — *September 2015 - August 2021*
  - Built a track-record of machine learning research with publications at ICML, Neurips, ICLR, ICASSP, ISMIR.
  - Created the current best music transcription model in the MIREX Multi-F0 Challenge.
  - Designed, built, and administered a GPU computing cluster to support two research groups ($\sim$ 40 GPUs).
- **Amazon** — Seattle, WA
  Applied Science Intern - Amazon Music Machine Learning — *June 2019 – Aug. 2019*
  - Hosts: Ted Sandler, Ben London.
  - Built a recommendation model using deep contextual bandits to sequence tracks on Amazon Music stations.
  - Used counterfactual risk minimization to train models off-policy from logged user interaction data.
- **Bracebridge Capital** - Fixed Income Arbitrage Hedge Fund — Boston, MA
  Quantitative Developer (Lead Developer) - Quantitative Research — *July. 2013 – May 2015*
  - Led and mentored a team of three software engineers developing C++ software infrastructure.
  - Maintained models governing a billion dollar asset backed structured product portfolio.
  - Organized front-office data acquisition, coordinating between research, vendors, the trading floor, and IT.
  - Built in-house models of implied volatilities and sensitivities, with applications to rates products.
- **Bracebridge Capital** — Boston, MA
  Summer Analyst - Quantitative Research — *June 2012 – Aug. 2012*
  - Rebuilt a legacy Excel product model using modern technologies: C#, Postgres, and JavaScript.
  - Completed a quant training course on valuation and risk models, with a focus on fixed income products.
- **Sirius Software** - Database Vendor — Cambridge, MA
  Software Engineer - Systems Software Engineering — *Jan. 2008 – May 2012*
  - Wrote compiler extensions for a database bytecode query language used by 1000+ developers worldwide.
  - Developed core system libraries for reporting, parsing, and web, supporting applications with millions of users.
  - Delivered international on-site programmer training and product demonstrations.
  - Linux systems administration: DNS, backups, software updates, security, documentation wiki.

Academic Service

- Journal Reviewer:
  - Transactions of the International Symposium on Music Information Retrieval 2021, 2022.
  - Journal of Creative Music Systems 2022.
  - IEEE Signal Processing Letters 2022.
- Conference Reviewer:
  - Advances in Neural Information Processing Systems 2016, 2020, 2021, 2022.
  - International Conference on Machine Learning 2018, 2021, 2022.
  - International Conference on Learning Representations 2022, 2023.
- Workshop Reviewer:
  - NeurIPS ML Safety Workshop, 2022.
- UW CSE Application Reader, PhD Admissions: 2018, 2019, 2020.
  - Screened and reviewed 30-50 PhD applications annually for the UW CSE Machine Learning group.
- UW CSE Machine Learning Graduate Student Recruiting Activities Coordinator: 2018, 2019.
  - Planned and organized on-campus recruiting events and activities for 100+ current and prospective students.
- UW CSE Graduate Social Co-Chair, 2017.
  - Organized weekly student social events for the Allen School graduate student community.
  - Worked to create inclusive activities that are accessible and appealing to our diverse community of students.
- Panelist:
  - June 2021: Howard University Karsh STEM Scholars Research Panel for Incoming First-Year Students
  - March 2021: UW CSE PhD Student Experience Panel for Admitted Graduate Students
  - February 2020: CSE 142 Careers in Research Panel for First-Year Computer Science Students
  - May 2018: ACM Research Night for UW Undergraduate Students
- Stanford Student-Applicant Support Program (SASP) 2022.
  - Provided feedback on PhD applications to prospective Stanford students from under-represented groups.
- Stanford CS Undergraduate Mentoring, 2021-2022.
  - This program provides early research mentoring to undergraduate students from underrepresented groups.
  - Met regularly with my mentee during the 2021-2022 academic year.
- UW CSE Pre-Application Mentorship Service (PAMS) 2021.
  - Mentored prospective applicants to the UW PhD program from historically marginalized groups.
- Co-founder and organizer of the UW Machine Learning and Optimization Reading Group.
  - Organized and scheduled speakers for a weekly seminar for 5 years (2015-2020).
  - In 2020, this seminar grew into the regular meeting of ADSI/ IFDS, funded by an NSF Tripods grant.

Publications and Preprints

Theses

- Leveraging Generative Models for Music and Signal Processing.
  University of Washington Dissertation
  Paul G. Allen School of Computer Science & Engineering, 2021.
  John Thickstun.
- Statistical Inference on Music with Applications to the Transcription Problem.
  Brown University Senior Thesis
  Department of Applied Mathematics, 2013.
  Brown University **Distinguished Senior Thesis Award** (university-wide award).
  John Thickstun.

Conference Publications

- Melody Transcription via Generative Pre-Training.
  In International Symposium on Music Information Retrieval (**ISMIR**) 2022. Acceptance rate: 43.3%
  Chris Donahue, John Thickstun, Percy Liang.
- Diffusion-LM Improves Controllable Text Generation.
  In Advances in Neural Information Processing Systems (**Neurips**) 2022. Acceptance rate: 25.6%
  Selected for **Oral Presentation**.
  Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, Tatsunori B. Hashimoto.
- MAUVE: Measuring the Gap Between Neural Text and Human Text using Divergence Frontiers.
  In Advances in Neural Information Processing Systems (**Neurips**) 2021. Acceptance rate: 25.7%
  Outstanding Paper Award: 6 / 9122 paper submissions.
  Krishna Pillutla, Swabha Swayamdipta, Rowan Zellers, John Thickstun, Sean Welleck, Yejin Choi, Zaid Harchaoui.
- Parallel and Flexible Sampling from Autoregressive Models via Langevin Dynamics.
  In International Conference on Machine Learning (**ICML**) 2021. Acceptance rate: 21.5%
  Vivek Jayaram*, John Thickstun* (*equal contribution).
- Faster Policy Learning with Continuous-Time Gradients.
  In Learning for Dynamics & Control (**L4DC**) 2021.
  Samuel Ainsworth, Kendall Lowrey, John Thickstun, Zaid Harchaoui, Siddhartha Srinivasa.
- An Information Bottleneck Approach for Controlling Conciseness in Rationale Extraction.
  In Empirical Methods in Natural Language Processing (**EMNLP**) 2020. Acceptance rate: 24.5%
  Bhargavi Paranjape, Mandar Joshi, John Thickstun, Hannaneh Hajishirzi, Luke Zettlemoyer.
- Source Separation with Deep Generative Priors.
  In International Conference on Machine Learning (**ICML**) 2020. Acceptance rate: 21.8%
  Vivek Jayaram*, John Thickstun* (*equal contribution).
- Convolutional Composer Classification.
  In International Symposium on Music Information Retrieval (**ISMIR**) 2019. Acceptance rate: 45.1%
  Harsh Verma, John Thickstun.
- Coupled Recurrent Models for Polyphonic Music Composition.
  In International Symposium on Music Information Retrieval (**ISMIR**) 2019. Acceptance rate: 45.1%
  John Thickstun, Zaid Harchaoui, Dean P. Foster, Sham M. Kakade.
- Invariances and Data Augmentation for Supervised Music Transcription.
  In International Conference on Acoustics, Speech, and Signal Processing (**ICASSP**) 2018. Acceptance rate: 49.7%
  Selected for **Oral Presentation**.
  John Thickstun, Zaid Harchaoui, Dean P. Foster, Sham M. Kakade.
- Frequency Domain Convolutions for Multiple F0 Estimation.
  MIREX Abstract (Technical Report) 2017.
  John Thickstun, Zaid Harchaoui, Dean P. Foster, Sham M. Kakade.
- MusicNet: Learning Features of Music from Scratch.
  In International Conference on Learning Representations (**ICLR**) 2017. Acceptance rate: 39.1%
  John Thickstun, Zaid Harchaoui, Sham M. Kakade.

Pre-print Reports

- Evaluating Human-Language Model Interaction.
  Under Review, 2022.
  Mina Lee, Megha Srivastava, Amelia Hardy, John Thickstun, Esin Durmus, Ashwin Paranjape, Ines Gerard-Ursin, Xiang Lisa Li, Faisal Ladhak, Frieda Rong, Rose E. Wang, Minae Kwon, Joon Sung Park, Hancheng Cao, Tony Lee, Rishi Bommasani, Michael Bernstein, Percy Liang.

- Reconstruction of Visual Images from Murine Retinal Ganglion Cell Spiking Activity using Convolutional Neural Networks.
  Under Review, 2022.
  Tyler Benster, Darwin Babino, John Thickstun, Matthew Hunt, Xiyang Liu, Zaid Harchaoui, Sewoong Oh, Russell N. Van Gelder.

Invited Talks and Presentations

- Controlling Generative Models for Content Creation - CLEAR Ventures - Palo Alto, CA - 10/24/2022
- Controlling Generative Models in Diverse Media Domains - Meta - New York, NY - 10/18/2022
- Audio Source Separation with Deep Generative Priors - Mila - Quebec, Canada - 8/19/2022
- Classifier-Guided Controllable Text Generation with Diffusion-LM - AI2 - Seattle, WA - 8/3/2022
- Generative Modeling of Classical Western Music - SAIL - Stanford University - 12/4/2020
- Source Separation with Deep Generative Priors - ICML - Vienna, Austria - 7/14/2020
- Convolutional Composer Classification - ISMIR - Delft, Netherlands - 11/6/2019
- Autoregressive Modeling of Musical Scores - ISMIR - Delft, Netherlands - 11/5/2019
- Robust Generative Modeling in Generic Problem Domains - CSE Colloquium - UW Seattle - 10/31/2019
- Neural Music Transcription - ICASSP - Calgary, Canada - 4/18/2018
- MusicNet: Learning Features of Music from Scratch - ICLR - Toulon, France - 4/25/2017
- Automatic Music Transcription - CS Department - Brown University - 5/1/2013
- Introducing the Janus XmlParser - Sirius User Group - St. Louis, MO - 5/2/2010
- Tokenization and Collection Objects - Centrelink - Canberra, Australia - 3/23/2010

Media Coverage

- TechCrunch - MusicNet aims to give machine learning algorithms a taste for Beethoven.
- The Times of London - Bach to the future: computer will finish composer's work.
- A Tempo with Rachel Katz (WWFM Radio) - Computers and music.
- CNET - Bach to the future: AI, meet classical music.
- CIFAR - Learning algorithms find a new music teacher.
- RouteNote - How do you advance machine learning? Teach them Beethoven and Bach.
- UW Today - What makes Bach sound like Bach? New dataset teaches algorithms classical music.
- Allen School News - UW researchers hit the right note with new machine learning tool for music.
- Allen School News - Jayaram and Thickstun win Qualcomm Fellowship for work in source separation.

Open Source Contributions

- Stanford CRFM Mistral
  - https://github.com/stanford-crfm/mistral
  - A framework for replicable training of GPT-2 scale Transformer models.
  - I contributed abstractions for training models over non-text datasets including, for example, music.
- HuggingFace Datasets -
  - https://github.com/huggingface/datasets/
  - A Python library of natural language processing datasets and utilities.
  - I contributed model evaluation metrics that implement the MAUVE algorithm for measuring the quality of machine-generated text.
- FIFE Engine
  - https://github.com/fifengine/fifengine
  - A multi-platform isometric game engine written in C++ with Python bindings for scripting.
  - I was an early core contributor from 2007-2008 with 241 commits touching all aspects of the engine.