JUSTIN LOVELACE

jl3353@cornell.edu
https://justinlovelace.github.io/
https://scholar.google.com/citations?user=k2CBs3gAAAAJ

EDUCATION

Cornell University August 2022 - Present
– PhD in Computer Science
– Advisor: Prof. Kilian Q Weinberger
– Google PhD Fellow in Machine Learning & ML Foundations

Carnegie Mellon University August 2020 - August 2022
– M.S. in Language Technologies, School of Computer Science GPA: 4.10
– Advisor: Prof. Carolyn Rosé

Texas A&M University August 2016 - May 2020
– B.S. in Computer Science, Minor in Mathematics GPA: 4.0
– Advisor: Prof. Bobak Mortazavi
– Honors Fellow, Undergraduate Research Scholar
– USA Powerlifting: All-American, 2019 Collegiate National Runner-Up

HONORS AND AWARDS

Google PhD Fellowship—Machine Learning & ML Foundations 2025-2027

NSF National Artificial Intelligence Research Resource (NAIRR) Pilot Award 2024
Principal Investigator

· Project: "Enhancing Large Language Model Alignment through Diffusion Generated Soft-Prompts"
· Award: 10,000 GPU Hours on TACC Vista Supercomputer

Keri Family PhD Fellowship 2022-2023
Cornell University

Carnegie Mellon University Research Fellowship 2020-2022

CRA's Outstanding Undergraduate Researcher Award (Honorable Mention) 2020

PUBLICATIONS

1. "Pre-training Limited Memory Language Models with Internal and External Knowledge"
   Linxi Zhao, Sofian Zalouk, Christian Belardi, **Justin Lovelace**, Jin Peng Zhou, Kilian Q Weinberger, Yoav Artzi, Jennifer Sun
   arXiv Preprint, 2025

2. "SpeechOp: Inference-Time Task Composition for Generative Speech Processing"
   Justin Lovelace, Rithesh Kumar, Jiaqi Su, Ke Chen, Kilian Q Weinberger, Zeyu Jin
   arXiv Preprint, 2025

3. "Stop-Think-AutoRegress: Language Modeling with Latent Diffusion Planning"
   Justin Lovelace, Christian Belardi, Sofian Zalouk, Adhitya Polavaram, Srivatsa Kundurthy, Kilian Q Weinberger
   COLM, 2025

4. "Improving Multislice Electron Ptychography with a Generative Prior"
   Christian Belardi, Chia-Hao Lee, Yingheng Wang, **Justin Lovelace**, Kilian Q Weinberger, David Muller, Carla Gomes
   Computer Vision for Materials Science Workshop (CV4MS) at ICCV, 2025

5. "Sample-Efficient Diffusion for Text-To-Speech Synthesis"
   Justin Lovelace, Soham Ray, Kwangyoun Kim, Kilian Q Weinberger, Felix Wu
   Interspeech, 2024

6. "Diffusion Guided Language Modeling"
   Justin Lovelace, Varsha Kishore, Yiwei Chen, Kilian Q Weinberger
   Findings of the Annual Meeting of the Association for Computational Linguistics (Findings of ACL), 2024

7. "Latent Diffusion for Language Generation"
   Justin Lovelace, Varsha Kishore, Chao Wan, Eliot Shekhtman, Kilian Q Weinberger
   Conference on Neural Information Processing Systems (NeurIPS), 2023

8. "IncDSI: Incrementally Updatable Document Retrieval"
   Varsha Kishore, Chao Wan, **Justin Lovelace**, Yoav Artzi, Kilian Q Weinberger
   International Conference on Machine Learning (ICML), 2023

9. "A Framework For Adapting Pre-Trained Language Models to Knowledge Graph Completion"
   Justin Lovelace and Carolyn Rosé
   Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022

10. "Robust Knowledge Graph Completion with Stacked Convolutions and a Student Re-Ranking Network"
    Justin Lovelace, Denis Newman-Griffis, Shikhar Vashishth, Jill Fain Lehman, and Carolyn Rosé
    Annual Meeting of the Association for Computational Linguistics and the International Joint Conference on Natural Language Processing (ACL-IJCNLP), 2021

11. "Learning to Generate Clinically Coherent Chest X-Ray Reports"
    Justin Lovelace, Bobak Mortazavi
    Findings of the Conference on Empirical Methods in Natural Language Processing (Findings of EMNLP), 2020

12. "Dynamically Extracting Outcome-Specific Problem Lists from Clinical Notes with Guided Multi-Headed Attention"
    Justin Lovelace, Nathan C. Hurley, Adrian Haimovich, Bobak Mortazavi
    Machine Learning for Healthcare Conference (MLHC), 2020

ACADEMIC EXPERIENCE

ML Core Lab at Cornell University — August 2022 - Present
Graduate Research Assistant — *Advisor: Dr. Kilian Q. Weinberger*

· Developing latent diffusion language models.

Teledia Lab at Carnegie Mellon University — August 2020 - August 2022
Graduate Research Assistant — *Advisor: Dr. Carolyn Rosé*

· Adapted pre-trained language models to improve the coverage of knowledge graphs.

STMI Lab at Texas A&M University — April 2018 - May 2020
Undergraduate Researcher — *Advisor: Dr. Bobak Mortazavi*

· Developed convolutional attention models to extract information from clinical notes and predict adverse outcomes for ICU patients.

· Developed an abstractive radiology report generation framework that improved the clinical correctness of generated reports.

INDUSTRY EXPERIENCE

Meta FAIR May 2025 - Present
Research Scientist Intern - Language & Multimodal Foundations *New York City, NY*

· Developing generative models for speech generation.

Adobe Research May 2024 - August 2024
Research Scientist Intern- Speech AI Group *San Francisco, CA*

· Developed SpeechOp, a multi-task latent diffusion model transforming pre-trained TTS into a universal speech processor (TTS, enhancement, separation).
· Introduced novel inference-time Implicit Task Composition (ITC), achieving state-of-the-art content preservation in speech enhancement using Whisper guidance.

ASAPP May 2023 - August 2023
Research Intern *New York City, NY*

· Developed the first diffusion model for text-to-speech synthesis that does not require explicit phoneme alignment for generation.
· Outperformed the state-of-the-art autoregressive TTS system using less than 2% the training data.

PATENTS

Patent Applications

· "Text-to-Audio Conversion with Byte-Encoding Vectors," **Justin Lovelace**, Soham Ray, Felix Wu, Kilian Q. Weinberger, Kwangyoun Kim. U.S. Patent Application No. 20250104692, Filed January 8, 2024, Patent Pending.

TEACHING EXPERIENCE

Cornell University January 2024 - May 2024
Instructor of Record and Course Developer *CS 4782: Introduction to Deep Learning*

· Served as instructor of record (with one co-instructor) for the pilot offering of Cornell's undergraduate deep learning course; course is now part of the permanent curriculum
· Designed complete curriculum including lectures, assignments, and projects for 4000-level CS course
· Covered fundamentals of neural networks, computer vision, NLP, generative models, and reinforcement learning

SERVICE

Reviewer

· NeurIPS (2025)
· ICML (2025)
· ACL Rolling Review (2024, 2025)
· ACM Conference on Health, Inference, and Learning (2020, 2021)
· NeurIPS Machine Learning for Health (ML4H) Workshop (2019, 2020, 2021)

TECHNICAL SKILLS

Python, C++, SQL, Pytorch, Git