JUSTIN LOVELACE

(512) 964-9498 ⋄ jl3353@cornell.edu
https://justinlovelace.github.io/
https://scholar.google.com/citations?user=k2CBs3gAAAAJ

EDUCATION

Cornell University August 2022 - Present
– Ph.D. in Computer Science
– Advisor: Prof. Kilian Q Weinberger

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

PUBLICATIONS

1. "Pre-training Large 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

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

3. "Improving Multislice Electron Ptychography with a Generative Prior"
Travis Zhang, Christian Belardi, **Justin Lovelace**, Jin Peng Zhou, Saebyeol Shin, Carla P Gomes, Kilian Q Weinberger
Computer Vision for Materials Science Workshop (CV4MS) at ICCV, 2025

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

5. "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

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

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

8. "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

9. "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

10. "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

11. "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

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.

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*

· Leveraged 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.

TEACHING EXPERIENCE

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

· Developed and co-taught the pilot offering of Cornell's undergraduate deep learning course, covering fundamentals of neural networks, computer vision, NLP, generative models, and reinforcement learning
· Created the curriculum including lectures, assignments, and projects for 4000-level CS course

INDUSTRY EXPERIENCE

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

· Developing generative audio models.

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.

Facebook May 2020 - August 2020
Software Engineering Intern- Search *Menlo Park, CA*

· Developed a service (C++, Python, SQL) to onboard keywords to the search index in real time from a variety of sources (e.g. news articles).
· Conducted experiment with live traffic and found that my framework improved Facebook's total search volume, search value, and keyword retrieval latency.
· Launched service to production.

Facebook May 2019 - August 2019
Software Engineering Intern-Notification Ranking *Menlo Park, CA*

· Implemented rate limiting service (C++, Python) to protect Facebook's internal reachability service and its dependencies from unstable traffic.
· Developed ML pipeline (Python, SQL) to extend reachability service from email to include SMS.
· Demonstrated that the ML model can reduce the amount of undeliverable SMS sent by over half.

TECHNICAL SKILLS

Python, C++, SQL, Pytorch, Git

HONORS AND AWARDS

NSF National Artificial Intelligence Research Resource (NAIRR) Pilot Award 2024
Principal Investigator *Estimated value: $50,000*

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

Cornell University Ph.D. Fellowship 2022-2023

Carnegie Mellon University Research Fellowship 2020-2022

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

Outstanding Undergraduate Honors Thesis Award 2019

SERVICE

Reviewer

- 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)