KOYENA PAL

Northeastern University ◇ Boston, MA 02115
(401) · 226 · 7477 ◇ pal.k@northeastern.edu ◇ koyenapal.github.io

EDUCATION

Northeastern University 2022 – 2027
Doctor of Philosophy, Computer Science
- GPA: 4.0/4.0
- Thesis Proposal: Model Lakes (website)
- Co-Advisor: Dr. David Bau, Assistant Professor of CS @ Northeastern
- Co-Advisor: Dr. Renée Miller, Canada Excellence Research Chair in Data Intelligence @ UWaterloo. (former University Distinguished Professor @ Northeastern)
- Research Areas: Interpretable AI, NLP, Data Science

Brown University 2021 – 2022
Master of Science, Computer Science
- GPA: 4.0/4.0
- Thesis: Summarization and Generation of Discharge Summary Medical Reports
- Thesis Advisor: Dr. Carsten Eickhoff, Director of Health NLP Lab @ University of Tübingen

Brown University 2017 – 2021
Bachelor of Science, Computer Science, Honors
- GPA: 3.8/4.0
- Thesis: The Effect of Multi-Document Summarizations on User SERP Experience
- Thesis Advisor: Dr. Carsten Eickhoff, Director of Health NLP Lab @ University of Tübingen

PEER-REVIEWED PUBLICATIONS

1. Dmitrii Troitskii*, Koyena Pal*, Chris Wendler, Callum Stuart McDougall. Internal States before Wait Modulate Reasoning Patterns. *Conference on Empirical Methods in Natural Language Processing (EMNLP) Findings. 2025.*

2. Aaron Mueller, Jannik Brinkmann, Millicent Li, Samuel Marks, Koyena Pal, Nikhil Prakash, Can Rager, Aruna Sankaranarayanan, Arnab Sen Sharma, Jiuding Sun, Eric Todd, David Bau, Yonatan Belinkov. The Quest for the Right Mediator: A History, Survey, and Theoretical Grounding of Causal Interpretability. *Computational Linguistics. 2025.*

3. Koyena Pal, David Bau, and Renée J. Miller. 2024. Model Lakes. *28th International Conference on Extending Database Technology (EDBT). 2025.*

4. Jaden Fiotto-Kaufman, Alexander R Loftus, Eric Todd, Jannik Brinkmann, Caden Juang, Koyena Pal, Can Rager, Aaron Mueller, Samuel Marks, Arnab Sen Sharma, Francesca Lucchetti, Michael Ripa, Adam Belfki, Nikhil Prakash, Sumeet Multani, Carla Brodley, Arjun Guha, Jonathan Bell, Byron Wallace, David Bau NNsight and NDIF: Democratizing Access to Foundation Model Internals. *The Thirteenth International Conference on Learning Representations (ICLR). 2025.*

5. Koyena Pal, Aamod Khatiwada, Roee Shraga and Renée J. Miller (2024). ALT-GEN: Benchmarking Table Union Search using Large Language Models. *VLDB 2024 Workshop: Tabular Data Analysis Workshop (TaDA). 2024.* [Best Long Paper]

6. Koyena Pal, Jiuding Sun, Andrew Yuan, Byron C. Wallace, & David Bau (2023). Future Lens: Anticipating Subsequent Tokens from a Single Hidden State. *SIGNLL Conference on Computational Natural Language Learning (CoNLL). 2023.*

7. C. Meyer, D. Adkins, K. Pal, R. Galici, A. Garcia-Agundez, and C. Eickhoff , "Neural text generation in regulatory medical writing," *Frontiers in Pharmacology. 2023*

8. K. Pal, S. Adepu and J. Goh, "Effectiveness of Association Rules Mining for Invariants Generation in Cyber-Physical Systems," *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE). 2017.* [Top 3 Best Student Paper Presentation]

PRE-PRINT PUBLICATIONS

1. Koyena Pal, Seyed Ali Bahrainian, Laura Mercurio, and Carsten Eickhoff. 2023. Neural Summarization of Electronic Health Records.

RESEARCH EXPERIENCE

Cambridge-Alignment Boston Initiative (CBAI) June 2025 - Aug 2025
Fellow *Cambridge, MA*

- Mentor: Chandan Singh, Senior Researcher @ Microsoft
- Interpretable AI: Evaluated the generalizability of Chain-of-Thought reasoning across Large Reasoning Models to determine if one model's CoT can guide another to the same answer. (In submission)

Northeastern University Sept 2022 - Present
PhD Student Researcher *Boston, MA*

- Interpretable AI + NLP: Explored ways to predict distant tokens of a Large Language Model (LLM) through single hidden token representations. (Published at CoNLL 2023)
- Database + AI: Explored ways to create harder benchmarks and simpler solutions for database problems using LLMs. (Best Long Paper at TaDA @ VLDB 2024)
- Database + AI: A vision for future research on model lakes, as an extension to a popular concept in database known as data lakes, that enables model search and understanding in model sharing platforms. (Published at EDBT 2025)

Brown University Sept 2019 - Sept 2022
Student Researcher *Providence, RI*

- AI + Healthcare: Developed a patient-centric literature summarization mechanism by implementing NLP models on biomedical texts.
- AI + HCI: Conducted user study on the effect of multi-document summarizations on User Search Results Page (SERP) experience.

iTrust Centre at Singapore University of Technology and Design April 2016 - Jan 2017
Research Intern *Singapore*

- AI + Cyber Security: Proposed a novel method to detect cyber-attacks on a Cyber-Physical System.

INDUSTRY EXPERIENCE

Akamai Technologies Jan 2022 - May 2022
Information Security Intern *Cambridge, MA*

- Created text summarization and recommender systems for a threat-intelligence dashboard

Akamai Technologies May 2020 - August 2020
Information Security Intern *Cambridge, MA*

- Designed and deployed UI features to achieve consistent language interpretation in Technical Risk Illuminator, a tool built to support executive decisions.
- Modeled a multiple-tag generator to identify key terms and incidents types from cyber-incident reports.

Akamai Technologies May 2019 - August 2019
Information Security Intern *Cambridge, MA*

· Optimized the yearly audit process duration by building a systems register, which generates answers that business units require to show that they are PCI Compliant.
· Revamped the PCI Gap Analysis Template to streamline the process of products entering the PCI Compliance program for the first time.

Brown University May 2018 - July 2018
Information Security Analyst Intern *Providence, RI*

· Substantially improved copyright ticketing process automation through network expansion in support of internal IP address, MAC address, and public IP username identification.

TEACHING AND RESEARCH MENTORSHIP

Northeastern University Sept 2022 - Dec 2022
Graduate TA *Boston, MA*

· CSCI 7150: Deep Learning, with Dr. David Bau (Fall 2022)

Brown University Dec 2018 - Dec 2020
*Teaching Assistant (TA) and Head TA** *Providence, RI*

· CSCI 1470* and 2470*: Deep Learning, with Dr. Daniel Ritchie (Fall 2020)
· CSCI 1010: Theory of Computation, with Dr. Lorenzo De Stefani (Fall 2019)
· CSCI 0220: Intro to Discrete Structures and Probability, with Dr. Caroline J Klivans (Spring 2019)
· CSCI 0170: Computer Science – An Integrated Introduction, with Dr. Philip N. Klein (Fall 2018)
Organized 35 TAs teaching 350+ students. Developed course materials, managed course logistics, led weekly labs, graded student work and held office hours.

Inspirit AI July 2020 - August 2020, Dec 2021
Instructor *Online*

· Conducted online lectures and coding assignments covering concepts such as machine learning fundamentals, NLP, computer vision, and ethics in AI.
· Guided global group of students through AI projects such as Anti-Refugee Semantic Analysis.

INVITED TALKS AND MEDIA COVERAGE

- "All the AI risks we cannot see", Jan 2025
- DS4440: Practical Neural Networks @ Northeastern University, Boston, MA, Nov 2024
- "Opening AI's Black Box with Prof. David Bau, Koyena Pal, and Eric Todd of Northeastern University". The Cognitive Revolution Podcast, Boston, MA, April 2024.
- Mechanistic Interpretability Reading Seminar, Discord, Online, Jan 2024

HONORS AND AWARDS

Northeastern Graduate Fellowship 2022-2027
Cambridge-Boston Alignment Initiative Fellow 2025
Microsoft EDBT Fellowship 2025
Best Long Paper @ VLDB TaDA Workshop 2024
Northeastern Startup Fund 2022
Brown CS Scholarship for Richard Tapia Conference 2021
Brown University Undergraduate Honors Award 2021
Brown Undergraduate Teaching Assistant-ship 2019

Top 3 Best Student Paper Award @ IEEE 18th HASE Conference 2017

ORGANIZED WORKSHOPS AND CONFERENCE

2nd New England Mechanistic Interpretability Workshop 2025
1st New England Mechanistic Interpretability Workshop 2024
Conference on Health, Inference, and Learning (CHIL) 2024

SERVICE

Reviewer @ ICLR 2026
Reviewer @ AAAI 2025
Reviewer @ NeurIPS 2024,2025
Reviewer @ ICML Actionable Interpretability Workshop 2025
Reviewer @ ICML 2025
Reviewer @ NeurIPS Interpretable AI Workshop 2024
Reviewer @ CoNLL 2024
Mentor @ Brown University's Women Launch Pad 2022-Present
Booth Host @ SOURCE 2023
Reviewer @ JAMIA 2023
Reviewer @ Northeastern University's CS PhD Admissions Committee 2022
Mentor @ Brown University's Women in Computer Science 2018-2022