

# KOYENA PAL

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

## EDUCATION

---

- |                                                                                                                                                                                                                              |             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>Northeastern University</b><br><i>Doctor of Philosophy, Computer Science</i><br>- GPA: 4.0/4.0<br>- Advisor: Dr. David Bau and Dr. Renée Miller<br>- Research Areas: Interpretable AI, NLP, Data Science                  | 2022 – 2027 |
| <b>Brown University</b><br><i>Master of Science, Computer Science</i><br>- GPA: 4.0/4.0<br>- Thesis: Summarization and Generation of Discharge Summary Medical Reports<br>- Thesis Advisor: Dr. Carsten Eickhoff             | 2021 – 2022 |
| <b>Brown University</b><br><i>Bachelor of Science, Computer Science, Honors</i><br>- GPA: 3.8/4.0<br>- Thesis: The Effect of Multi-Document Summarizations on User SERP Experience<br>- Thesis Advisor: Dr. Carsten Eickhoff | 2017 – 2021 |

## PEER-REVIEWED PUBLICATIONS

---

1. Koyena Pal, David Bau, and Renée J. Miller. 2024. Model Lakes. *28th International Conference on Extending Database Technology (EDBT)*. 2025.
2. 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 2024. NNsight and NDIF: Democratizing Access to Foundation Model Internals. *The Thirteenth International Conference on Learning Representations (ICLR)*. 2025.
3. Koyena Pal, Aamod Khatiwada, Roe Shraga and Renée J. Miller (2024). UGEN: Benchmarking Table Union Search using Large Language Models. *VLDB 2024 Workshop: Tabular Data Analysis Workshop (TaDA)*. [Best Long Paper]
4. 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)*.
5. C. Meyer, D. Adkins, K. Pal, R. Galici, A. Garcia-Agundez, and C. Eickhoff , “Neural text generation in regulatory medical writing,” *Frontiers in Pharmacology*, vol. 14, 2023, doi: 10.3389/fphar.2023.1086913.
6. 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, pp. 124-127, doi: 10.1109/HASE.2017.21. [Top 3 Best Student Paper Presentation]

## PRE-PRINT PUBLICATIONS

---

1. 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. 2024. The Quest for the Right Mediator: A History, Survey, and Theoretical Grounding of Causal Interpretability.
2. Koyena Pal, Seyed Ali Bahrainian, Laura Mercurio, and Carsten Eickhoff. 2023. Neural Summarization of Electronic Health Records.

## RESEARCH EXPERIENCE

---

### Northeastern University

*PhD Student Researcher*

Sept 2022 - Present

*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: Creating a vision for future research on model lakes, as an extension to a popular concept in database known as data lakes, to better understand and organize AI models in model sharing platforms. (Initial draft on arXiv: Model Lakes)

### Brown University

*Student Researcher*

Sept 2019 - Sept 2022

*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

*Research Intern*

April 2016 - Jan 2017

*Singapore*

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

## INDUSTRY EXPERIENCE

---

### Akamai Technologies

*Information Security Intern*

Jan 2022 - May 2022

*Cambridge, MA*

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

### Akamai Technologies

*Information Security Intern*

May 2020 - August 2020

*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

*Information Security Intern*

May 2019 - August 2019

*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***Information Security Analyst Intern*

May 2018 - July 2018

*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***Graduate TA*

Sept 2022 - Dec 2022

*Boston, MA*

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

**Brown University***Teaching Assistant (TA) and Head TA\**

Dec 2018 - Dec 2020

*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)

**Inspirit AI***Instructor*

July 2020 - August 2020, Dec 2021

*Online*

- Conducted online lectures and coding assignments covering concepts such as machine learning fundamentals, NLP, computer vision, and ethics in AI.

**INVITED TALKS**

---

- 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

**SERVICE**

---

<b>Lead Student Organizer @ New England Mechanistic Interpretability Meetup</b>	2024
<b>Reviewer @ NeurIPS 2024</b>	2024
<b>Reviewer @ NeurIPS Interpretable AI Workshop 2024</b>	2024
<b>Reviewer @ CoNLL 2024</b>	2024
<b>Comms Sub-Chair @ CHIL 2024</b>	2023-Present
<b>Mentor @ Brown University’s Women Launch Pad</b>	2022-Present
<b>Booth Host @ SOURCE</b>	2023
<b>Reviewer @ JAMIA</b>	2023
<b>Reviewer @ Northeastern University’s CS PhD Admissions Committee</b>	2022
<b>Mentor @ Brown University’s Women in Computer Science</b>	2018-2022

**HONORS AND AWARDS**

---

<b>Northeastern Graduate Fellowship</b>	2022-2027
<b>Best Long Paper @ VLDB TaDA Workshop</b>	2024
<b>Northeastern Startup Fund</b>	2022
<b>Brown CS Scholarship for Richard Tapia Conference</b>	2021
<b>Brown University Undergraduate Honors Award</b>	2021
<b>Brown Undergraduate Teaching Assistant-ship</b>	2019
<b>Top 3 Best Student Paper Award @ IEEE 18th HASE Conference</b>	2017