

WILLIAM RUDMAN

EDUCATION

Brown University Providence, RI	2020 - 2025
<i>Ph.D. in Computer Science</i>	
Research Areas: Large Language Model & Vision-Language Model Interpretability	
King's College London London, England	2018 - 2019
<i>M.Sc. in Complex Systems Modeling</i> (Pass with Distinction)	
Thesis: Frustration Solitaire with Stochastic Resetting	
University of Aberdeen Aberdeen, Scotland	2014 - 2018
<i>M.A. in Mathematics & M.A. in Sociology</i> (First Class Honours)	
Thesis: Representation Theory of Finite Groups	

PUBLICATIONS

- **William Rudman***, Frauke Andersen, Ruochen Zhang, and Carsten Eickhoff. *Accelerated Path Patching with Task-Specific Pruning*. Under Anonymous Review at EACL 2025.
- Florentin Beck, **William Rudman**, and Carsten Eickhoff. *TRIM: Achieving Extreme Sparsity with Targeted Row-wise Iterative Metric-driven Pruning*. Under Anonymous Review at EACL 2025.
- **William Rudman***, Michal Golovanevsky*, Amir Bar, Michael Lepori, Amir Bar, Ritambhara Singh, and Carsten Eickhoff. *Pixels Versus Priors: Controlling Knowledge Priors in Vision-Language Models through Visual Counterfactuals*. EMNLP 2025.
- **William Rudman***, Michal Golovanevsky*, Amir Bar, Vedant Palit, Ritambhara Singh, Carsten Eickhoff, and Yann LeCun. *Forgotten Polygons: Multimodal Large Language Models are Shape-Blind*. Findings of the Association of Computational Linguistics: ACL 2025.
- **William Rudman***, Michal Golovanevsky*, Vedant Palit, Ritambhara Singh, and Carsten Eickhoff. *What Do VLMs NOTICE? A Mechanistic Interpretability Pipeline for Noise-free Text-Image Corruption and Evaluation*. Nations of the Americas Chapter of the Association for Computational Linguistics: NAACL 2025.
- **William Rudman** and Carsten Eickhoff. *Stable Anisotropic Regularization*. International Conference on Learning Representations: ICLR 2024.
- Aleksa Kaye, **William Rudman**, and Stephanie Ruest. *Evaluation of a Child Abuse Screen Performed by Nurses among Young Children with Fractures*. Rhode Island Medical Journal 2024.
- **William Rudman**, Catherine Chen, and Carsten Eickhoff. *Outlier Dimensions Contain Task-Specific Knowledge*. Empirical Methods in Natural Language Processing: EMNLP 2023.
- Isaac Sears, Augusto Garcia-Agundez, George Zerveas, **William Rudman**, Laura Mercurio, Corey E Venter, Adeel Abbasi, and Carsten Eickhoff *Leveraging Unlabeled Electroencephalographic Data to Predict Neurological Recovery for Comatose Patients Following Cardiac Arrest*. Computing in Cardiology: CinC 2023.
- William Jurayj, **William Rudman** and Carsten Eickhoff. *Garden Path Traversal in GPT-2*. BlackBox NLP Workshop on Analyzing and Interpreting Neural Networks at EMNLP 2022.
- **William Rudman**, Nate Gillman, Taylor Rayne, and Carsten Eickhoff. *IsoScore: Measuring the Uniformity of Embedding Space Utilization*. Findings of the Association of Computational Linguistics: ACL 2022.
- **William Rudman**, Jack Merullo, and Carsten Eickhoff. *ACQuA: Anomaly Classification with Quasi-Attractors*. Computing in Cardiology: CinC 2022.
- **William Rudman**, Peter Corcoran, and Kimberly Elliot. *Prescription Drug Monitoring Programs and Opioid Death Rates*. Journal of the American Medical Association: JAMA 2017.

WORK EXPERIENCE

The University of Texas at Austin Austin, TX	August 2025 - Present
Postdoctoral Research Fellow: Computational Linguistics & CosmicAI	
PayPal San Jose, CA	June 2024 - September 2024
Machine Learning Research Scientist Intern	
Lincoln College of Technology Chicago, IL	January 2020 - September 2020
Adjunct Instructor: Technical Mathematics	
Rosalind Franklin University of Medicine and Science North Chicago, IL	June 2017 - September 2017
Public Health Research Intern	

GRANTS, HONORS, AND AWARDS

- ITSERR Research Fellow: National Research Council of Italy April 2025
- NIH: National Institute of General Medical Sciences April 2022 - June 2025
Title: *Derivation of a Clinical Prediction Rule for Pediatric Abusive Fractures.*
Award Amount per Year: \$431,771.
- T-32 NIH Predoctoral Biological Data Science Fellowship September 2020 - May 2022
- King's College London Complex Systems Modeling Prize: Best Master's Thesis December 2019
- King's College London International Master's Scholar September 2018

PROFESSIONAL ACTIVITIES

Mentorship

- *TRIM: Achieving Extreme Sparsity with Targeted Row-wise Iterative Metric-driven Pruning.* Florentin Beck, **William Rudman** and Carsten Eickhoff. Master's Thesis: University of Tübingen 2025.
- *SPOT: Sparse Parameter-efficient Outlier-guided fine-Tuning.* Selina Uecker, **William Rudman** and Carsten Eickhoff. Master's Thesis: University of Tübingen 2025.
- *A Comprehensive Metric for Circuit Overlap in Large Language Models.* Frauke Andersen, Ruochen Zhang, **William Rudman** and Carsten Eickhoff. Master's Thesis: University of Tübingen 2025.
- *Garden Path Traversal in GPT-2.* William Jurayj, **William Rudman** and Carsten Eickhoff. Master's Thesis: Brown University 2022.
- Google's Explore CSR: Socially Responsible AI for Computational Creativity 2021 - 2022.

Invited Talks

- UT-Austin: Computational Linguistics Department May 2025
- MIT: Geometric Data Processing Group. April 2025
- National Research Council of Italy. March 2025
- University of Hawaii at Manoa: AI Precision Health Institute. December 2023
- New England NLP: UMass Lowell. April 2023

Reviewing: EMNLP (2023-2025), ACL (2024-2025), NAACL (2025), ICLR (2025), NeuRIPS (2025), SIGGRAPH (2025).

SKILLS

Programming Languages: Python, PyTorch, Pandas, Git, Numpy, NLTK, spaCy, DeepSpeed, PySpark, SQL

Technical Knowledge: Natural Language Processing, Large Language Models, Vision-Language Models, Parameter-Efficient Fine-Tuning, Deep Learning, Mechanistic Interpretability, Machine Learning, Statistics