

# ELINOR POOLE-DAYAN

✉ [elinorpd@mit.edu](mailto:elinorpd@mit.edu) | [in elinor-poole-dayan](https://www.linkedin.com/in/elinor-poole-dayan) | [Google Scholar](https://scholar.google.com/citations?user=...) | [elinorpd.github.io](https://github.com/elinorpd)

## EDUCATION

---

### Massachusetts Institute of Technology

Cambridge, MA, USA

*Master's of Science*

*Sept. 2023 – May 2025*

- Advised by [Deb Roy](#) in the [MIT Center for Constructive Communication](#). GPA 5.0/5.0
- Thesis: [From Dialogue to Decision: An LLM-Powered Framework for Analyzing Collective Idea Evolution and Voting Dynamics in Deliberative Assemblies](#) (Grade: A+)

### McGill University

Montreal, QC, Canada

*Bachelor's in Honours Computer Science and Mathematics*

*Sept. 2019 – May 2023*

- GPA: 3.9/4.0, Awards: Dean's Honour List, J W McConnell Scholarship, Canadian Graduate Scholarship - Master's (NSERC) \$17,500, [Mila Excellence Scholarship - EDI in Research](#) \$5,000

## RESEARCH & PROFESSIONAL EXPERIENCE

---

### Predoctoral Researcher

July 2025 – Present

*Massachusetts Institute of Technology*

*Cambridge, MA, USA*

- Research on pluralistic alignment in LLMs [\[P.1\]](#), [\[R.1\]](#) and evaluating the fairness of AI's societal impacts; supervised by [Michiel Bakker](#).

### Graduate Research Assistant

Sept. 2023 – June 2025

*Center for Constructive Communication, MIT Media Lab*

*Cambridge, MA, USA*

- Played a key role in the design and execution of the [MIT Student Deliberative Assembly on Sustainability](#).
- Supported engineering and design efforts for tech-enhanced deliberative assembly pilots in Ireland and Oregon.
- Led lab efforts to organize and write documentation and developed custom API wrappers for Fora data handling.

### Data Science Intern

May 2022 – Aug. 2022

*Unity Technologies*

*Montreal, QC, Canada*

- Optimized deep learning algorithms to throttle bid requests in Unity's Ad Exchange using Tensorflow.
- Decreased model training time by 25% and reduced model size and number of parameters by 50%.
- Created a text data preprocessing pipeline on Google Cloud Platform Dataflow using Apache beam.

### NLP Research Intern

Jan. 2022 – May 2022

*McGill University*

*Montreal, QC, Canada*

- Identified the geo-indicativeness of text using BERT applied to geosocial datasets to build a safety tool for social media. Supervised by [Grant McKenzie](#).

### NLP Research Intern

June 2021 – Aug. 2021

*Shamoon College of Engineering*

*Be'er Sheva, Israel*

- Classified author gender of books to perform a case study on female authors who wrote under male pseudonyms.
- Preprocessed data using CoreNLP and scikit-learn. Designed and implemented baseline experiments using SVMs.
- Supervised by [Irina Rabaev](#) and [Marina Litvak](#).

### NLP Research Intern

Jan. 2021 – May 2021

*Mila - Quebec AI Institute*

*Montreal, QC, Canada*

- Investigated the effect of gender debiasing on fine-tuned language models such as BERT using PyTorch.
- Explored debiasing methods and reformulated bias metrics for racial and religious biases.
- Published paper in ACL with Nick Meade [\[C.4\]](#) and supervised by [Siva Reddy](#).

## Conference Proceedings

- [C.1] M. Hughes, B. Roy, E. **Poole-Dayana**, D. Roy, and J. Kabbara, “Computational Analysis of Conversation Dynamics through Participant Responsivity,” in *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, Suzhou, China, Nov. 2025. [Online]. Available: <https://arxiv.org/abs/2509.16464>.
- [C.2] S. Fulay, W. Brannon, S. Mohanty, C. Overney, E. **Poole-Dayana**, D. Roy, and J. Kabbara, “On the Relationship between Truth and Political Bias in Language Models,” in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Miami, Florida, USA, Nov. 2024, pp. 9004–9018. [Online]. Available: <https://aclanthology.org/2024.emnlp-main.508>.
- [C.3] B. Krojer, E. **Poole-Dayana**, V. Voleti, C. Pal, and S. Reddy, “Are Diffusion Models Vision-And-Language Reasoners?” In *Advances in Neural Information Processing Systems*, 2023. [Online]. Available: [https://proceedings.neurips.cc/paper\\_files/paper/2023/file/1a675d804f50509b8e21d0d3ca709d03-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/1a675d804f50509b8e21d0d3ca709d03-Paper-Conference.pdf).
- [C.4] N. Meade, E. **Poole-Dayana**, and S. Reddy, “An Empirical Survey of the Effectiveness of Debiasing Techniques for Pre-trained Language Models,” in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Dublin, Ireland, May 2022, pp. 1878–1898. [Online]. Available: <https://aclanthology.org/2022.acl-long.132>.

## Pre-prints & Under Review

- [R.1] E. **Poole-Dayana**, J. Wu, T. Sorensen, J. Pei, and M. A. Bakker, *Benchmarking Overton Pluralism in LLMs*, Manuscript under review, Sep. 2025,  
– Under Review at ICLR 2026.
- [R.2] E. **Poole-Dayana**, D. Roy, and J. Kabbara, *An AI-Powered Framework for Analyzing Collective Idea Evolution in Deliberative Assemblies*, ArXiv[Preprint], Aug. 2025. [Online]. Available: <https://arxiv.org/abs/2509.12577>,  
– Under Review at AAAI Social Impact Track 2026.
- [R.3] E. **Poole-Dayana**, D. Roy, and J. Kabbara, *LLM Targeted Underperformance Disproportionately Impacts Vulnerable Users*, ArXiv[Preprint], Jun. 2024. [Online]. Available: <https://arxiv.org/abs/2406.17737>,  
– Under Review at AAAI Social Impact Track 2026.

## Talks

- [T.1] E. **Poole-Dayana** and D. Roy, *Tracing Idea Evolution in Democratic Deliberation with LLMs*, [COLM 2025 NLP4Democracy Workshop](#) (Montréal, Canada), Oct. 2025,  
– *Oral Presentation*.

## Posters

- [P.1] E. **Poole-Dayana**, J. Wu, J. Pei, and M. A. Bakker, *Benchmarking Overton Pluralism in LLMs*, [NeurIPS 2025 LLM Evaluation Workshop](#) (San Diego, CA, USA), Dec. 2025.
- [P.2] E. **Poole-Dayana**, D. T. Kessler, M. Hughes, E. S. Lin, M. Ganz, and D. Roy, *Applying Large-Language Models to Characterize Public Narratives*, [NAACL 2025 Workshop on Narrative Understanding](#) (Albuquerque, NM, USA), May 2025.
- [P.3] E. **Poole-Dayana** and D. Roy, *LLM Targeted Underperformance Disproportionately Impacts Vulnerable Users*, [NeurIPS 2024 Safe Generative AI Workshop](#) (Vancouver, Canada), Dec. 2024.
- [P.4] E. **Poole-Dayana** and D. Roy, *Sandbagging in LLMs: Systematic Underperformance for Uneducated and ESL Users*, [International Conference on Computational Social Science \(IC2S2\)](#) (Philadelphia, PA, USA), Jul. 2024.

## TEACHING EXPERIENCE

---

### **Kaufman Teaching Certificate**

Feb. 2025 – May 2025

*MIT Teaching + Learning Lab*

*Cambridge, MA, USA*

- Participated in eight practice-based workshops, evaluated on my teaching skills through 2 microteaching sessions, received individual feedback from peers and teaching experts, and implemented evidence-based teaching techniques grounded in the scholarship of teaching and learning.
- Developed a syllabus for a course titled *Ethics, Fairness, and Bias in Generative Language Models*.

### **Teaching Assistant: Intro to Media Arts & Sciences**

Sept. 2024 – Dec. 2024

*MIT Media Lab*

*Cambridge, MA, USA*

### **Teaching Assistant: Honours Algorithms & Data Structures**

Jan. 2022 – May 2022

*McGill University*

*Montreal, QC, Canada*

## MENTORSHIP

---

**Jiayi Wu, Brown University** (Summer 2025): Benchmarking Overton Pluralism with LLM-as-a-Judge [\[P.1\]](#), [\[R.1\]](#)

**Hannah Chiou, Wellesley College** (Spring 2025): Computational Analysis of Public Narratives with LLMs [\[P.2\]](#)

## ACADEMIC SERVICE

---

**Reviewer for ACL Rolling Review:** May 2025; December 2024; October 2024 (Emergency Reviewer)

**Reviewer for AAAI:** Social Impact Track 2026

**Reviewer for ICLR:** 2026

## TECHNICAL SKILLS

---

**Programming Languages:** Python, Java, Javascript, C, Unix/Linux, OCaml, SQL

**Machine Learning & Data Science:** TensorFlow, PyTorch, Keras, scikit-learn, pandas, NumPy, matplotlib, seaborn, plotly

**Cloud Computing:** Google Cloud Platform, Amazon Web Services, Docker