

ELINOR POOLE-DAYAN

✉ elinorpd@mit.edu |  [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 (MIT)

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

Montréal, QC, Canada

Bachelor's in Honours Computer Science and Mathematics

Sept. 2019 – May 2023

- First Class Honours with Distinction. GPA 3.9/4.0

RESEARCH & PROFESSIONAL EXPERIENCE

Predoctoral Researcher

July 2025 – Present

Massachusetts Institute of Technology (MIT)

Cambridge, MA, USA

- Research on pluralistic alignment in LLMs [[W.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

- Conducted research related to fairness and bias in LLMs [[C.1](#), [C.3](#)] and computational social science [[C.2](#), [W.3](#)].
- Played a key role in the design and execution of the [MIT Student Deliberative Assembly on Sustainability](#) [[R.2](#)].
- 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

Montréal, 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

Montréal, 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

Montréal, 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.5](#)] and supervised by [Siva Reddy](#).

Academic Conference Publications

- C.1. **Poole-Dayana**, E., Roy, D. & Kabbara, J. *LLM Targeted Underperformance Disproportionately Impacts Vulnerable Users in Proceedings of the AAAI Conference on Artificial Intelligence* (Jan. 2026). <https://arxiv.org/abs/2406.17737>.
- C.2. Hughes, M., Roy, B., **Poole-Dayana**, E., Roy, D. & Kabbara, J. *Computational Analysis of Conversation Dynamics through Participant Responsivity in Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Nov. 2025). <https://aclanthology.org/2025.emnlp-main.1798/>. (*Oral Presentation*).
- C.3. Fulay, S., Brannon, W., Mohanty, S., Overney, C., **Poole-Dayana**, E., Roy, D. & Kabbara, J. *On the Relationship between Truth and Political Bias in Language Models in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Nov. 2024). <https://aclanthology.org/2024.emnlp-main.508>.
- C.4. Krojer, B., **Poole-Dayana**, E., Voleti, V., Pal, C. & Reddy, S. *Are Diffusion Models Vision-And-Language Reasoners? in Advances in Neural Information Processing Systems (NeurIPS)* (2023). https://proceedings.neurips.cc/paper_files/paper/2023/file/1a675d804f50509b8e21d0d3ca709d03-Paper-Conference.pdf.
- C.5. Meade, N., **Poole-Dayana**, E. & Reddy, S. *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 (ACL)* (May 2022). <https://aclanthology.org/2022.acl-long.132>.

Workshop Papers & Extended Abstracts

- W.1. **Poole-Dayana**, E., Wu, J., Pei, J. & Bakker, M. A. *Benchmarking Overton Pluralism in LLMs* [NeurIPS 2025 LLM Evaluation Workshop](#) (San Diego, CA, USA). Dec. 2025. (*Poster Presentation*).
- W.2. **Poole-Dayana**, E. & Roy, D. *Tracing Idea Evolution in Democratic Deliberation with LLMs* [COLM 2025 NLP4Democracy Workshop](#) (Montréal, Canada). Oct. 2025. (*Oral Presentation*).
- W.3. **Poole-Dayana**, E., Kessler, D. T., Hughes, M., Lin, E. S., Ganz, M. & Roy, D. *Applying Large-Language Models to Characterize Public Narratives* [NAACL 2025 Workshop on Narrative Understanding](#) (Albuquerque, NM, USA). May 2025. (*Poster Presentation*).
- W.4. **Poole-Dayana**, E. & Roy, D. *LLM Targeted Underperformance Disproportionately Impacts Vulnerable Users* [NeurIPS 2024 Safe Generative AI Workshop](#) (Vancouver, Canada). Dec. 2024. (*Poster Presentation*).
- W.5. **Poole-Dayana**, E. & Roy, D. *Sandbagging in LLMs: Systematic Underperformance for Uneducated and ESL Users* [International Conference on Computational Social Science \(IC2S2\)](#) (Philadelphia, PA, USA). July 2024. (*Poster Presentation*).

Pre-prints & Under Review

- R.1. **Poole-Dayana**, E., Wu, J., Sorensen, T., Pei, J. & Bakker, M. A. *Benchmarking Overton Pluralism in LLMs* Manuscript under review at [ICLR 2026](#). Sept. 2025.
- R.2. **Poole-Dayana**, E., Roy, D. & Kabbara, J. *An AI-Powered Framework for Analyzing Collective Idea Evolution in Deliberative Assemblies* ArXiv[Preprint]. Aug. 2025. <https://arxiv.org/abs/2509.12577>.

INVITED TALKS AND RESEARCH PRESENTATIONS

Nov 2025 – Bar Ilan University hosted by Prof. Yanai Elazar. Invited talk: *Measuring Overton Pluralism in LLMs*

Oct 2025 – COLM NLP4Democracy Workshop. Oral presentation: *Analyzing Collective Idea Evolution in Democratic Deliberation with LLMs* [\[W.2\]](#)

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

Montréal, QC, Canada

AWARDS AND HONOURS

2023 – Mila Excellence Scholarship - EDI in Research (\$5,000)

2023 – Canadian Graduate Scholarship - Master's (NSERC) (\$17,500)

2020 – Lev Bukhman Internship Award (\$1,070)

2019–2021 – Dean's Honour List (top 10% of class by GPA)

2019–2021 – J W McConnell Scholarship (\$9,000)

2019 – NY Cardiac Center Scholarship (\$20,000)

2019 – National AP Scholar

2019 – National Merit Scholarship Award (\$2,500)

2018 – Positive Coaching Alliance Triple-Impact Competitor Scholarship (\$1,000)

MENTORSHIP

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

Hannah Chiou, Wellesley College (Spring 2025): Computational Analysis of Public Narratives with LLMs [[W.3](#)]

ACADEMIC SERVICE

Reviewer for ACL Rolling Review: October 2025; 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