

# Maria Ryskina

Toronto, ON, Canada ◇ [maria.ryskina@vectorinstitute.ai](mailto:maria.ryskina@vectorinstitute.ai) ◇ <https://ryskina.github.io>

---

## EDUCATION

### Carnegie Mellon University

*Ph.D., Language and Information Technologies*

Pittsburgh, PA, USA

Aug 2016–Sep 2022

### Skolkovo Institute of Science and Technology

*M.Sc., Information Technology, with honors*

Moscow, Russia

Aug 2014–Jul 2016

### Moscow Institute of Physics and Technology

*M.Sc., Applied Mathematics and Physics, with honors*

Moscow, Russia

Sep 2014–Jun 2016

*B.Sc., Applied Mathematics and Physics, with honors*

Sep 2010–Jun 2014

## RESEARCH EXPERIENCE

### Vector Institute for Artificial Intelligence

*CIFAR AI Safety Postdoctoral Fellow*

Toronto, ON, Canada

Jun 2025–present

- My current project aims to develop RL training procedures that can enhance normative reasoning in LLMs in multi-agent settings, building on theories of how human societies create social norms.
- Advisor: Dr. Gillian K. Hadfield

### Massachusetts Institute of Technology

*Postdoctoral Associate, McGovern Institute for Brain Research*

Cambridge, MA, USA

Oct 2022–Oct 2024

- Localized brain networks that represent modality-independent conceptual meaning using neuroimaging data and used LLM representations to predict brain responses in these networks [C7]
- Advisors: Dr. Evelina Fedorenko, Dr. Kyle Mahowald

### Carnegie Mellon University

*Graduate Research Assistant, Language Technologies Institute*

Pittsburgh, PA, USA

Aug 2016–Sep 2022

- Developed structured unsupervised probabilistic models for character-level transduction problems [C3, W4] and for applications in digital humanities [C1]
- Modeled how new words have emerged in historical corpora [C2] and on social media, and how morphological inflection models adapt to new words
- Thesis: “Learning Computational Models of Non-Standard Language” [T1]
- Advisors: Dr. Matthew R. Gormley, Dr. Eduard Hovy, Dr. Taylor Berg-Kirkpatrick

### DiDi Labs

*Research Intern*

Remote

May 2020–Aug 2020

- Learned embeddings of number tokens from mathematical sequences and probed them for number-theoretic knowledge [W5, P2]
- Trained task-oriented reinforcement learning agents in a self-play setting [X1, P1]
- Mentor: Dr. Kevin Knight

### Massachusetts Institute of Technology

*Visiting Graduate Student, CSAIL Infolab*

Cambridge, MA, USA

Sep 2015–Jan 2016

- Used visual representations of text (sampled from a video captioning model) to detect entailment.
- Mentors: Dr. Boris Katz, Dr. Andrei Barbu

**Moscow Institute of Physics and Technology & Dorodnicyn Computing Centre of RAS**

Moscow, Russia

*Undergraduate Student, Dept. of Control and Applied Mathematics*

Sep 2012–Jun 2014

- Designed regularizers for topic models towards stable convergence to an interpretable solution.
- Advisor: Dr. Konstantin Vorontsov

## AWARDS AND FELLOWSHIPS

|   |           |
|---|-----------|
| CIFAR AI Safety Postdoctoral Fellowship   | 2025–2026 |
| Outstanding Reviewer Award, Annual Meeting of the ACL   | 2023      |
| Best Paper Award, ACM Conference on Fairness, Accountability, and Transparency<br>Awarded for “Queer In AI: A Case Study in Community-Led Participatory AI” | 2023      |
| Diverse Intelligences Summer Institute Fellowship<br>Templeton World Charity Foundation   | 2021      |
| Skoltech Academic Excellence Award, Russia  | 2016      |
| State Academic Scholarship, Russia  | 2013–2014 |
| Foundation for the Development of Innovative Education Scholarship, Russia  | 2010–2012 |

## PUBLICATIONS

\* denotes equal contribution

### JOURNAL ARTICLES

- [J2] [Elements of World Knowledge \(EWoK\): A Cognition-Inspired Framework for Evaluating Basic World Knowledge in Language Models](#)  
A. A. Ivanova\*, A. Sathe\*, B. Lipkin\*, ... V. Paulun, M. Ryskina, E. Akyürek, ... J. Tenenbaum, J. Andreas  
*Transactions of the Association for Computational Linguistics (TACL)* 13, 2025.
- [J1] [A taxonomy and review of generalization research in NLP](#)  
D. Hupkes, M. Giulianelli, V. Dankers, ... L. Khalatbari, M. Ryskina, R. Frieske, R. Cotterell, Z. Jin  
*Nature Machine Intelligence* 5(10), 2023.

### CONFERENCE PROCEEDINGS

- [C7] [Language models align with brain regions that represent concepts across modalities](#)  
M. Ryskina, G. Tuckute, A. Fung, A. Malkin, E. Fedorenko  
**[Spotlight]** *Conference on Language Modeling (COLM)*, 2025.
- [C6] [Queer In AI: A Case Study in Community-Led Participatory AI](#)  
Organizers of Queer in AI: A. Ovalle, A. Subramonian, ... M. L. Pacheco, M. Ryskina, M. Mundt, ... L. Stark  
**[Best Paper]** *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2023.
- [C5] [UniMorph 4.0: Universal Morphology](#)  
K. Batsuren\*, O. Goldman\*, ... Y. Ghanggo Ate, M. Ryskina, S. Mielke, ... R. Tsarfaty, E. Vylomova  
*Language Resources and Evaluation Conference (LREC)*, 2022.

- [C4] [NoiseQA: Challenge Set Evaluation for User-Centric Question Answering](#)  
A. Ravichander, S. Dalmia, **M. Ryskina**, F. Metze, E. Hovy, A. W. Black  
*Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021.
- [C3] [Phonetic and Visual Priors for Decipherment of Informal Romanization](#)  
**M. Ryskina**, M. R. Gormley, T. Berg-Kirkpatrick  
*Annual Meeting of the Association for Computational Linguistics (ACL)*, 2020.
- [C2] [Where New Words Are Born: Distributional Semantic Analysis of Neologisms and Their Semantic Neighborhoods](#)  
**M. Ryskina**, E. Rabinovich, T. Berg-Kirkpatrick, D. R. Mortensen, Y. Tsvetkov  
*Annual Meeting of the Society for Computation in Linguistics (SCiL)*, 2020.
- [C1] [Automatic Compositor Attribution in the First Folio of Shakespeare](#)  
**M. Ryskina**, H. Alpert-Abrams, D. Garrette, T. Berg-Kirkpatrick  
*Annual Meeting of the Association for Computational Linguistics (ACL)*, 2017.

## WORKSHOP PROCEEDINGS

- [W5] [Learning Mathematical Properties of Integers](#)  
**M. Ryskina**, K. Knight  
*BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, 2021.
- [W4] [Comparative Error Analysis in Neural and Finite-state Models for Unsupervised Character-level Transduction](#)  
**M. Ryskina**, E. Hovy, T. Berg-Kirkpatrick, M. R. Gormley  
*SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 2021.
- [W3] [SIGMORPHON 2021 Shared Task on Morphological Reinflection: Generalization Across Languages](#)  
T. Pimentel\*, **M. Ryskina\***, S. Mielke, ... R. Cotterell, B. Ambridge, E. Vylomova  
*SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 2021.
- [W2] [OPERA: Operations-oriented Probabilistic Extraction, Reasoning, and Analysis \(2019\)](#)  
E. Hovy, J. Carbonell, ... X. Ma, **M. Ryskina**, R. Sanabria, V. Gangal  
*Text Analysis Conference (TAC)*, 2019.
- [W1] [OPERA: Operations-oriented Probabilistic Extraction, Reasoning, and Analysis \(2018\)](#)  
E. Hovy, T. Berg-Kirkpatrick, ... D. Rajagopal, **M. Ryskina**, R. Sanabria  
*Text Analysis Conference (TAC)*, 2018.

## THESES

- [T1] [Learning Computational Models of Non-Standard Language](#)  
**M. Ryskina**  
*Ph.D. dissertation*, Carnegie Mellon University, 2022.

## PATENTS

- [P2] US Patent 11460982, [Number embedding application system](#)  
K. Knight, **M. Ryskina**  
Oct 4, 2022

[P1] US Patent 12026544, [Self-play to improve task-oriented dialog systems and methods](#)  
K. Knight, M. Ryskina, A. Arkhangorodsky, A. Nagesh, S. Fang  
Jul 2, 2024

## MAGAZINE ARTICLES

[M1] [Queer in AI](#)  
Organizers of Queer in AI: H. Jethwani, A. Subramonian, ... M. Ryskina, J. Xiong  
*XRDS: Crossroads, The ACM Magazine for Students*, 28(4), 2022.

## PREPRINTS

[X1] [Two Approaches to Building Collaborative, Task-Oriented Dialog Agents through Self-Play](#)  
A. Arkhangorodsky, S. Fang, V. Knight, A. Nagesh, M. Ryskina, K. Knight  
*Preprint*, 2021.

## INVITED TALKS

|   |              |
|---|--------------|
| <b>Normative justification for scalable LLM oversight</b><br>Canadian AI Safety Institute (CAISI) Research Program Annual Meeting, CIFAR<br>Toronto AI Safety Reading Group, Vector Institute/University of Toronto/Schwartz Reisman Institute  | Oct 2025     |
| <b>Language models as models of the brain and as normative social agents</b><br>NLP Seminar Series, National Research Council Canada  | Oct 2025     |
| <b>Untangling Language: Representation Levels and the Long Tail Challenge in NLP</b><br>Visitor Talk Series, Vector Institute for Artificial Intelligence   | Apr 2025     |
| <b>Learning Computational Models of Non-Standard Language</b><br>LCC Lab, Department of Computer Science, University of Toronto<br>EvLab, Department of Brain and Cognitive Sciences, MIT<br>CLunch Seminar Series, Department of Computer and Information Science, University of Pennsylvania<br>MCQLL Lab, Department of Linguistics, McGill University | Feb–Mar 2022 |
| <b><a href="#">Romanization with Friends: Deciphering Informally Romanized Text</a></b><br>NLP With Friends Seminar Series  | Nov 2021     |
| <b>Unsupervised Decipherment of Informal Romanization</b><br>NLPhD Speaker Series, Department of Language Science and Technology, Saarland University   | Jun 2021     |
| <b><a href="#">Informal Romanization across Languages and Scripts</a></b><br>SIGTYP Lecture Series  | Jun 2021     |

## TEACHING & MENTORING

### Courses:

- TA and Guest lecturer for [CMU 11-711: Algorithms for NLP](#). Instructor: Prof. Yulia Tsvetkov Fall 2018
- TA for [CMU 11-711: Algorithms for NLP](#). Instructor: Prof. Taylor Berg-Kirkpatrick. Fall 2017

### Students mentored:

- Ashley Malkin, high school intern, MIT Jul–Aug 2023 & Jun–Jul 2024
- Sahal Ahmed, undergraduate student, MIT Sep 2023–May 2024

## SERVICE

### Professional service:

- Treasurer, Toronto Professional ACM-W chapter, 2025–present
- Diversity & Inclusion Chair, Annual Meeting of the Association for Computational Linguistics (ACL) 2025
- Core organizer at Queer in AI, 2022–present
- Mentor for MIT Research Science Institute, Summer 2024
- Co-organizer for the Queer in AI workshops at NAACL 2022, ACL 2023
- Co-organizer for SIGMORPHON Shared Task 0: Generalization in Morphological Inflection Generation, 2021
- Co-host for SIGTYP Lecture Series, 2021

### Area chairing:

- Action Editor for ACL Rolling Review / Area Chair, Interpretability and Analysis of Models for NLP: ACL 2024, NAACL 2024, EACL 2024, NAACL 2025
- Area Chair, Conference on Language Modeling (COLM 2025)

### Reviewing (journal articles):

- Cognitive Science (2021)
- Review of General Psychology (2021)

### Reviewing (conference proceedings):

- Conference on Language Modeling (COLM 2024)
- Annual Meeting of the Cognitive Science Society (CogSci 2024, 2025)
- Annual Meeting of the Association for Computational Linguistics (ACL 2021, 2023)
- Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021, 2023)
- SIGNLL Conference on Computational Natural Language Learning (CoNLL 2020, 2021, 2023, 2025)
- Conference on Empirical Methods in Natural Language Processing (EMNLP 2018, 2020, 2021, 2022)
- ACL Rolling Review (ARR 2021, 2022)

### Reviewing (workshop proceedings):

- GenBench Workshop on Generalisation (Benchmarking) in NLP (EMNLP 2023)
- SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology (ACL 2023)
- CAWL Workshop on Computation and Written Language (ACL 2023)
- Queer in AI (ACL 2023, NAACL 2022)
- Student Research Workshop (ACL 2020, ACL 2021, NAACL 2021, AACL 2020)
- Human And Model in the Loop Evaluation and Training Strategies (NeurIPS 2020)

### Volunteering and community service:

- Co-organizer for oSTEM/Queer in AI Graduate School Application Review and Fee Aid Programs, 2021–2025
- Student reviewer for CMU LTI Graduate Admissions, 2021
- Mentor for CMU SCS Graduate Application Support Program, 2020–2021
- Mentor for CMU LTI Peer Mentoring Program for new graduate students, 2021
- Mentor for CMU AI Mentoring Program for undergraduates from underrepresented groups, 2020