# Siva Reddy

McGill University
Montreal, Quebec.

email: siva.reddy@mcgill.ca
https://www.sivareddy.in
https://mcgill-nlp.github.io/

| | | |
|---|---|---|
| **Academic Position** | **Assistant Professor**<br>Mila – Quebec Artificial Intelligence Institute<br>School of Computer Science, McGill University<br>Department of Linguistics, McGill University | Jan 2020 – current |
| **Education** | **Postdoctoral Researcher**<br>Computer Science Department, Stanford University<br>Advisor: Christopher D. Manning | June 2017 – Sep 2019 |
| | **PhD, Informatics**<br>University of Edinburgh, UK<br>Supervisors: Mirella Lapata, Mark Steedman<br>Committee: Luke Zettlemoyer, Adam Lopez<br>Thesis: Syntax-Mediated Semantic Parsing | Oct 2012 – Nov 2017 |
| | **M.S. by Research, Computer Science**,<br>University of York, UK<br>Thesis: Polysemy in Compositional Distributional Semantics | Oct 2010 – Sep 2011 |
| | **M.S. by Research, Computer Science**,<br>IIIT Hyderabad, India | Jul 2009 – Sep 2010 |
| | **B.Tech with Honours, Computer Science**<br>IIIT Hyderabad, India | Jul 2005 – Jul 2009 |
| **Employment** | **Research Scientist (20%)**<br>ServiceNow Research (previously Element AI) | Jul 2022 – current |
| | **Visiting Scientist**<br>Simons Institute, UC Berkeley. | Spring 2025 |
| | **Scientific Advisor (20%)**<br>IVADO Labs, Montreal | May 2021 – Dec 2021<br>https://ivadolabs.com/en/ |
| | **Research Intern**<br>Google Research, New York | Jul 2014 – Nov 2014 |
| | **Research Engineer**<br>Lexical Computing Ltd., UK | Oct 2011 – Aug 2012<br>http://sketchengine.co.uk |
| | **Google Summer of Code** | Summer 2009 |
| **Awards** | CS-Can Outstanding Early Career Computer Science Researcher Award, 2025<br>Canada CIFAR AI Chair, 2025<br>Senior Area Chair Paper Award, NAACL 2025<br>Best Paper Award, EMNLP 2021<br>VentureBeat AI Innovation of the Year Award in NLP, July 2020 | |

Facebook CIFAR AI chair, 2020
Google PhD Fellowship, 2015-2017
The Principal's Career Development Scholarship, 2012-2014, University of Edinburgh
The Edinburgh Global Research Scholarship, 2012-2014, University of Edinburgh
Fully Funded Overseas Research Scholarship, 2010-2011, University of York
Best Paper Award, IJCNLP 2011

**Grants**

**Google Gemini Academic Program Award, USD $100,000 (GCP credits)**, 2025–2026
People: **Siva Reddy** (PI)

**Canada CIFAR AI Chair Renewal, CAD $1.2 Million**, 2025–2030
People: **Siva Reddy**

**IVADO R3AI NLP Regroupement, CAD $3 Million**, 2024–2030
People: **Siva Reddy** (co-Leader) and Amal Zouaq (co-Leader), and 20 other researchers.

**IVADO R3AI Safety Regroupement, CAD $40,000**, 2024–2030
People: **Siva Reddy**

**CIFAR Catalyst Grant, CAD $100,000**, 2024
Natural Language Processing for Users from Diverse Cultures
People: Vered Shwartz (PI) and **Siva Reddy** (co-PI)

**Google-Mila Grant, CAD $100,000**, 2024
Vision Language Models for Everyone: Building geo-diverse and culturally aware vision-language models
People: Aishwarya Agrawal (PI), **Siva Reddy** (co-PI), Lisa Anne Hendricks, Sjoerd van Steenkist, Fei Sha.

**NSERC-Mitacs Alliance Grant, CAD $1.5 Million**, 2024–2029
Overcoming Data Scarcity in Natural Language Processing for Industrial Applications through Knowledge Engineering and Reasoning
People: Jian-Yun Nie (PI), Yoshua Bengio (co-PI), Philippe Langlais (co-PI), Jackie Cheung (co-PI), Luc Lamontagne (co-PI), **Siva Reddy** (co-PI), Samira Kahou (co-PI), Richard Khoury (co-PI), Bang Liu (co-PI), Jian Tang (co-PI), Amal Zouaq (co-PI)

**Samsung-Mila Grant, CAD $200,000**, 2024–2026
People: Siva Reddy (PI)

**Mitacs Grant with ServiceNow, CAD $30,000**, 2024
People: Siva Reddy (PI)

**Mila Technology Maturation Grant, CAD $165,000**, 2023–2026
People: Siva Reddy (PI)

**MSR-Mila Grant, CAD $120,000**, 2023–2024
People: Siva Reddy (PI)

**Google-Mila Grant, CAD $50,000**, 2023–2024
People: Siva Reddy (PI), Chris Pal

**Intel-Mila Grant, CAD $90,000**, 2022–2025
People: Siva Reddy (PI)

**FRQNT New Academics, CAD $50,800,**                                        2022–2024
People: Siva Reddy (PI)

**MITACS Accelerate, CAD$32,500**                              Apr 2022 – May 2023
People: Devang Kulshreshtha, Nicholas Meade, Iulian Vlad Serban (Korbit AI), Siva Reddy (PI)

**Samsung-Mila Grant, CAD $180,000,**                                        2021–2023
People: Yash Goyal, Aishwarya Agrawal (co-PI), Siva Reddy (co-PI), Aaron Carouville (PI)

**IBM-Mila Grant, CAD $90,000,**                                             2021–2023
People: Siva Reddy (PI)

**CFI JELF Innovation Fund, CAD $450,000,**                                        2021
People: Tim O'Donnell (PI), Siva Reddy (co-PI)

**IVADO – Fundamental Research Grant, CAD$200,000**           Sep 2020 – Aug 2022
People: Gilles Pesant (PI), Sarath Chandar (co-PI), Siva Reddy (co-PI)

**MITACS Accelerate, CAD$32,500**                              Sep 2020 – Aug 2021
People: Devang Kulshreshtha, Iulian Vlad Serban (Korbit AI), Siva Reddy (PI)

**MITACS Accelerate, CAD$30,000**                              May 2020 – Jan 2021
People: Emily Goodwin, Dzmitry Bahdanau (Element AI), Timothy J. O'Donnell, Siva Reddy (PI)

**Scale AI: Access to resources on COVID-19
through a chatbot, CAD $560,000 (30K is my allocation)**.     May 2020 – Sep 2020
Partners: Dialogue, Mila, Nu Echo, Samasource, Google, JHU and Dataperformers

**NSERC Discovery Grant CAD $238,000**                         Apr 2020 – Mar 2027
People: Siva Reddy (PI)

**NSERC Discovery Launch Supplement CAD $12,500**             Apr 2020 – Mar 2027
People: Siva Reddy (PI)

**MSR-Mila Grant, CAD$55,000**                                  Apr 2020-Mar 2021
People: Shehzaad Dhuliawala, Kaheer Suleman, Siva Reddy (PI)

**Facebook CIFAR AI Chair, CAD $1.1 million**                                2020-2024
People: Siva Reddy (PI)

**Amazon Research Awards, US $100,000**                                      Jan 2018
Title: Enabling Multilingual Language Understanding
People: Christopher D. Manning (PI), Siva Reddy

**Stanford AI Lab Seed Grant, US $50,000**                                  Dec 2017
Title: Algorithms for Developing Context-Dependent Understanding of Language
People: Christopher D. Manning (PI), Siva Reddy

**Facebook AI Research Grant, US $20,000**                                  Oct 2017
Title: A Task for Developing Context-Dependent Understanding of Language
People: Christopher D. Manning (PI), Siva Reddy, Drew Arad

**Journal Publications**

[J19] DeepSeek-R1 Thoughtology: Let's think about LLM Reasoning
Sara Vera Marjanović, Arkil Patel, Vaibhav Adlakha, Milad Aghajohari, Parishad BehnamGhader, Mehar Bhatia, Aditi Khandelwal, Austin Kraft, Benno Krojer, Xing Han Lù, Nicholas Meade, Dongchan Shin, Amirhossein Kazemnejad, Gaurav Kamath, Marius Mosbach, Karolina Stańczak, **Siva Reddy**
*TMLR 2026.*

[J18] Reframing linguistic bootstrapping as joint inference using visually-grounded grammar induction models
Eva Portelance, **Siva Reddy**, Timothy O'Donnell
*Journal of Memory and Language 2025.*

[J17] Semantic change in adults is not primarily a generational phenomenon
Gaurav Kamath, Michelle Yang, **Siva Reddy**, Morgan Sonderegger, Dallas Card
*Proceedings of the National Academy of Sciences 2025.*

[J16] The BrowserGym Ecosystem for Web Agent Research
Thibault Chezelles et al. (20 authors)
*TMLR 2025.*

[J15] Investigating Adversarial Trigger Transfer in Large Language Models
Nicholas Meade, Arkil Patel, **Siva Reddy**
*TACL 2025.*

[J14] Large language models deconstruct the clinical intuition behind diagnosing autism
Jack Stanley, Emmett Rabot, **Siva Reddy**, Eugene Belilovsky, Laurent Mottron, Danilo Bzdok
*Cell 2025.*

[J13] Scope Ambiguities in Large Language Models
Gaurav Kamath, Sebastian Schuster, Sowmya Vajjala, **Siva Reddy**
*TACL 2024.*

[J12] Evaluating Correctness and Faithfulness of Instruction-Following Models for Question Answering
Vaibhav Adlakha, Parishad BehnamGhader, Xing Han Lu, Nicholas Meade, **Siva Reddy**
*TACL 2024.*

[J11] Data science opportunities of large-language models for neuroscience and biomedicine
Danilo Bzdok, Andrew Thieme, Oleksiy Levkovskyy, Paul Wren, Thomas Ray, **Siva Reddy**
*Neuron 2024.*

[J10] StarCoder: may the source be with you!
Raymond Li et al. (67 authors)
*TMLR 2023.*

[J9] Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models
Aarohi Srivastava et al. (450 authors)
*TMLR 2023.*

[J8] FaithDial: A Faithful Benchmark for Information-Seeking Dialogue
Nouha Dziri, Ehsan Kamalloo, Sivan Milton, Osmar Zaiane, Mo Yu, Edoardo M Ponti, **Siva Reddy**
*TACL 2022.*

[J7] TopiOCQA: Open-domain Conversational Question Answering with Topic Switching
Vaibhav Adlakha, Shehzaad Dhuliawala, Kaheer Suleman, Harm de Vries, **Siva Reddy**
*TACL 2022*

[J6] Post-hoc interpretability for neural nlp: A survey
Andreas Madsen, **Siva Reddy**, Sarath Chandar
*ACM Computing Surveys 2022*

[J5] Does Entity Abstraction Help Generative Transformers Reason?
Nicolas Gontier, **Siva Reddy**, Christopher Pal
TMLR 2022

[J4] CoQA: A Conversational Question Answering Challenge.
**Siva Reddy**\*, Danqi Chen\*, Christopher D. Manning.
*TACL, 2019.*

[J3] Learning an Executable Neural Semantic Parser.
Jianpeng Cheng, **Siva Reddy**, Vijay Saraswat, Mirella Lapata.
*Computational Linguistics, 2019.*

[J2] Transforming Dependency Structures to Logical Forms for Semantic Parsing.
**Siva Reddy**, Oscar Täckström, Michael Collins, Tom Kwiatkowski, Dipanjan Das,
Mark Steedman, Mirella Lapata.
*TACL, 2016.*

[J1] Large-scale Semantic Parsing without Question-Answer Pairs.
**Siva Reddy**, Mirella Lapata, Mark Steedman.
*TACL, 2014.*

**Conference Publications**

[C67] The Markovian Thinker: Architecture-Agnostic Linear Scaling of Reasoning
Milad Aghajohari, Kamran Chitsaz, Amirhossein Kazemnejad, Sarath Chandar, Alessandro Sordoni, Aaron Courville, **Siva Reddy**
*ICLR 2026.*

[C66] The Promise of RL for Autoregressive Image Editing
Saba Ahmadi, Rabiul Awal, Ankur Sikarwar, Amirhossein Kazemnejad, Ge Ya Luo, Juan A Rodriguez, Sai Rajeswar, **Siva Reddy**, Christopher Pal, Benno Krojer, Aishwarya Agrawal
*NeurIPS 2025.*

[C65] AgentRewardBench: Evaluating Automatic Evaluations of Web Agent Trajectories
Xing Han Lù, Amirhossein Kazemnejad, Nicholas Meade, Arkil Patel, Dongchan Shin, Alejandra Zambrano, Karolina Stanczak, Peter Shaw, Christopher Pal, **Siva Reddy**
*COLM 2025.*

[C64] Not All Data Are Unlearned Equally
Aravind Krishnan, **Siva Reddy**, Marius Mosbach
*COLM 2025.*

[C63] SafeArena: Evaluating the Safety of Autonomous Web Agents
Ada Tur, Nicholas Meade, Xing Han Lù, Alejandra Zambrano, Arkil Patel, Esin Durmus, Spandana Gella, Karolina Stanczak, **Siva Reddy**
*ICML 2025.*

[C62] VinePPO: Unlocking RL Potential For LLM Reasoning Through Refined Credit Assignment
Amirhossein Kazemnejad, Milad Aghajohari, Eva Portelance, Alessandro Sordoni, **Siva Reddy**, Aaron Courville, Nicolas Le Roux
*ICML 2025.*

[C61] REARANK: Reasoning Re-ranking Agent via Reinforcement Learning
Le Zhang, Bo Wang, Xipeng Qiu, **Siva Reddy**, Aishwarya Agrawal
*EMNLP 2025.*

[C60] WebMMU: A Benchmark for Multimodal Multilingual Website Understanding and Code Generation
Rabiul Awal, Mahsa Massoud, Aarash Feizi, Zichao Li, Suyuchen Wang, Christopher Pal, Aishwarya Agrawal, David Vazquez, **Siva Reddy**, Juan A. Rodriguez, Perouz Taslakian, Spandana Gella, Sai Rajeswar
*EMNLP 2025.*

[C59] Exploiting Instruction-Following Retrievers for Malicious Information Retrieval
Parishad BehnamGhader, Nicholas Meade, **Siva Reddy**
*ACL Findings 2025.*

[C58] Warmup Generations: A Task-Agnostic Approach for Guiding Sequence-to-Sequence Learning with Unsupervised Initial State Generation
Senyu Li, Zipeng Sun, Jiayi Wang, Xue Liu, Pontus Stenetorp, **Siva Reddy**, David Ifeoluwa Adelani
*ACL 2025.*

[C57] Understanding the Influence of Synthetic Data for Text Embedders
Jacob Mitchell Springer, Vaibhav Adlakha, **Siva Reddy**, Aditi Raghunathan, Marius Mosbach
*ACL Findings 2025.*

[C56] Language Models Largely Exhibit Human-like Constituent Ordering Preferences
Ada Tur, Gaurav Kamath, **Siva Reddy**
*NAACL 2025.* SAC Paper Award

[C55] MMTEB: Massive Multilingual Text Embedding Benchmark
Kenneth Enevoldsen, ... (70+ authors), **Siva Reddy**, Niklas Muennighoff
*ICLR 2025.*

[C54] BigDocs: An Open and Permissively-Licensed Dataset for Training Multimodal Models on Document and Code Tasks
Juan Rodriguez et al. (39 authors)
*ICLR 2025.*

[C53] Benchmarking vision language models for cultural understanding
Shravan Nayak, Kanishk Jain, Rabiul Awal, **Siva Reddy**, Sjoerd van Steenkiste, Lisa Anne Hendricks, Karolina Stańczak, Aishwarya Agrawal
*EMNLP 2024.*

[C52] Learning Action and Reasoning-Centric Image Editing from Videos and Simulations
Benno Krojer, Dheeraj Vattikonda, Luis Lara, Varun Jampani, Eva Portelance, Christopher Pal, **Siva Reddy**
*NeurIPS (Spotlight) Datasets & Benchmarks 2024.*

[C51] LLM2Vec: Large Language Models Are Secretly Powerful Text Encoders
Parishad BehnamGhader, Vaibhav Adlakha, Marius Mosbach, Dzmitry Bahdanau, Nicolas Chapados, **Siva Reddy**
*COLM 2024.* Three million+ downloads on HuggingFace

[C50] Are self-explanations from Large Language Models faithful?
Andreas Madsen, Sarath Chandar, **Siva Reddy**
*ACL Findings 2024.*

[C49] WebLINX: Real-World Website Navigation with Multi-Turn Dialogue
Xing Han Lu, Zdeněk Kasner, **Siva Reddy**
*ICML 2024 (Spotlight).* 1000+ Github stars

[C48] Faithfulness Measurable Masked Language Models
Andreas Madsen, **Siva Reddy**, Sarath Chandar
*ICML 2024 (Spotlight).*

[C47] Evaluating In-Context Learning of Libraries for Code Generation
Arkil Patel, **Siva Reddy**, Dzmitry Bahdanau, Pradeep Dasigi
*NAACL 2024.*

[C46] The Impact of Positional Encoding on Length Generalization in Transformers
Amirhossein Kazemnejad, Inkit Padhi, K. Ramamurthy, Payel Das, **Siva Reddy**
*NeurIPS 2023.*

[C45] Are Diffusion Models Vision-And-Language Reasoners?
Benno Krojer, Elinor Poole-Dayan, Vikram S. Voleti, C. Pal, **Siva Reddy**
*NeurIPS 2023.*

[C44] Syntactic Substitutability as Unsupervised Dependency Syntax
Jasper Jian, **Siva Reddy**
*EMNLP 2023.*

[C43] MAGNIFICo: Evaluating the In-Context Learning Ability of Large Language Models
to Generalize to Novel Interpretations
Arkil Patel, Satwik Bhattamishra, **Siva Reddy**, Dzmitry Bahdanau
*EMNLP 2023.*

[C42] Can Retriever-Augmented Language Models Reason? The Blame Game Between the
Retriever and the Language Model
Parishad BehnamGhader, Santiago Miret, **Siva Reddy**
*EMNLP Findings 2023.*

[C41] Evaluating Dependencies in Fact Editing for Language Models: Specificity and Impli-
cation Awareness
Zichao Li, Ines Arous, **Siva Reddy**, Jackie Chi Kit Cheung
*EMNLP Findings 2023.*

[C40] Using In-Context Learning to Improve Dialogue Safety
Nicholas Meade, Spandana Gella, Devamanyu Hazarika, Prakhar Gupta, Di Jin, **Siva
Reddy**, Yang Liu, Dilek Z. Hakkani-Tür
*EMNLP Findings 2023.*

[C39] Combining Parameter-efficient Modules for Task-level Generalisation
Edoardo M Ponti, Alessandro Sordoni, Yoshua Bengio, **Siva Reddy**
*EACL 2023.*

[C38] The StatCan Dialogue Dataset: Retrieving Data Tables through Conversations with
Genuine Intents
Xing Han Lu, **Siva Reddy**, Harm de Vries
*EACL 2023.*

[C37] On the Origin of Hallucinations in Conversational Models: Is it the Datasets or the
Models?
Nouha Dziri, Sivan Milton, Mo Yu, Osmar Zaiane, **Siva Reddy**
*NAACL 2022*

[C36] An Empirical Survey of the Effectiveness of Debiasing Techniques for Pre-Trained
Language Models
Nicholas Meade, Elinor Poole-Dayan, **Siva Reddy**
*ACL 2022*

[C35] Image Retrieval from Contextual Descriptions
Benno Krojer, Vaibhav Adlakha, Vibhav Vineet, Yash Goyal, Edoardo Ponti, **Siva
Reddy**
*ACL 2022*

[C34] Using Interactive Feedback to Improve the Accuracy and Explainability of Question
Answering Systems Post-Deployment

Zichao Li, Prakhar Sharma, Xing Han Lu, Jackie CK Cheung, **Siva Reddy**
*ACL Findings 2022*

[C33] Evaluating the faithfulness of importance measures in nlp by recursively masking allegedly important tokens and retraining
Andreas Madsen, Nicholas Meade, Vaibhav Adlakha, **Siva Reddy**
*EMNLP Findings 2022.*

[C32] The Power of Prompt Tuning for Low-Resource Semantic Parsing
Nathan Schucher, **Siva Reddy**, Harm de Vries
*ACL 2022*

[C31] Compositional generalization in dependency parsing
Emily Goodwin, **Siva Reddy**, Timothy J O'Donnell, Dzmitry Bahdanau
*ACL 2022*

[C30] IGLUE: A Benchmark for Transfer Learning across Modalities, Tasks, and Languages
Emanuele Bugliarello, Fangyu Liu, Jonas Pfeiffer, **Siva Reddy**, Desmond Elliott, Edoardo Maria Ponti, Ivan Vulic
*ICML 2022*

[C29] End-to-End Training of Multi-Document Reader and Retriever for Open-Domain Question Answering
Devendra Singh Sachan, **Siva Reddy**, William Hamilton, Chris Dyer, Dani Yogatama
*NeurIPS 2021.*

[C28] Few-shot Question Generation for Personalized Feedback in Intelligent Tutoring Systems
Devang Kulshreshtha, Muhammad Shayan, Robert Belfer, **Siva Reddy**, Iulian Vlad Serban, Ekaterina Kochmar
*PAIS'22: Prestigious Applications of Intelligent Systems 2022*

[C27] Back-Training excels Self-Training at Unsupervised Domain Adaptation of Question Generation and Passage Retrieval
Devang Kulshreshtha, Robert Belfer, Iulian Vlad Serban, **Siva Reddy**.
*EMNLP 2021*

[C26] Mind the Context: The Impact of Contextualization in Neural Module Networks for Grounding Visual Referring Expressions.
Arjun Akula, Spandana Gella, Keze Wang, Song-Chun Zhu, **Siva Reddy**.
*EMNLP 2021*

[C25] Visually Grounded Reasoning across Languages and Cultures
Fangyu Liu, Emanuele Bugliarello, Edoardo Maria Ponti, **Siva Reddy**, Nigel Collier, Desmond Elliott.
*EMNLP 2021* Best Paper Award

[C24] StereoSet: Measuring stereotypical bias in pretrained language models
Moin Nadeem, Anna Bethke, **Siva Reddy**.
*ACL 2021.* VentureBeat AI Innovation of the year award

[C23] Minimax and Neyman-Pearson Meta-Learning for Outlier Languages.
Edoardo Maria Ponti, Rahul Aralikatte, Disha Shrivastava, **Siva Reddy** and Anders Søgaard.
*ACL Findings 2021.*

[C22] Explicitly Modeling Syntax in Language Models with Incremental Parsing and a Dynamic Oracle
Yikang Shen, Shawn Tan, Alessandro Sordoni, **Siva Reddy**, Aaron Courville.
*NAACL 2021*

[C21] Understanding by Understanding Not: Modeling Negation in Language Models.
Arian Hosseini, **Siva Reddy**, Dzmitry Bahdanau, R Devon Hjelm, Alessandro Sordoni, Aaron Courville.
*NAACL 2021*

[C20] Learning Improvised Chatbots from Adversarial Modifications of Natural Language Feedback.
Makesh Narsimhan, Ryan Ni, **Siva Reddy**.
*EMNLP Findings, 2020.*

[C19] Words aren't enough, their order matters: On the Robustness of Grounding Visual Referring Expressions.
Arjun Akula, Spandana Gella, Yaser Al-Onaizan, Song-Chun Zhu, **Siva Reddy**.
*ACL, 2020.*

[C18] Measuring Systematic Generalization in Neural Proof Generation with Transformers.
Nicolas Gontier, Koustuv Sinha, **Siva Reddy**, Chris Pal.
*NeurIPS, 2020.*

[C17] Learning Typed Entailment Graphs with Global Soft Constraints.
Mohammad Hosseini, Nathanael Chambers, **Siva Reddy**, Xavier Holt, Shay Cohen, Mark Johnson, Mark Steedman.
*TACL, 2018.*

[C16] Universal Semantic Parsing.
**Siva Reddy**, Oscar Täckström, Slav Petrov, Mark Steedman, Mirella Lapata.
*EMNLP, 2017.*

[C15] Learning to Paraphrase for Question Answering.
Li Dong, Jonathan Mallinson, **Siva Reddy**, Mirella Lapata.
*EMNLP, 2017.*

[C14] Learning Structured Natural Language Representations for Semantic Parsing.
Jianpeng Cheng, **Siva Reddy**, Vijay Saraswat, Mirella Lapata.
*ACL, 2017.*

[C13] Question Answering on Knowledge Bases and Text using Universal Schema and Memory Networks.
Rajarshi Das, Manzil Zaheer, **Siva Reddy**, Andrew McCallum.
*ACL, 2017.*

[C12] Predicting Target Language CCG Supertags Improves Neural Machine Translation.
Maria Nadejde, **Siva Reddy**, Rico Sennrich, Tomasz Dwojak, Marcin Junczys-Dowmunt, Philipp Koehn, Alexandra Birch.
*Conference on Machine Translation, 2017.*

[C11] Question Answering on Freebase via Relation Extraction and Textual Evidence.
Kun Xu, **Siva Reddy**, Yansong Feng, Songfang Huang, Dongyan Zhao.
*ACL, 2016.*

[C10] Evaluating Induced CCG Parsers on Grounded Semantic Parsing.
Yonatan Bisk*, **Siva Reddy***, John Blitzer, Julia Hockenmaier, Mark Steedman.
*EMNLP, 2016.*

[C9] Assessing Relative Sentence Complexity using an Incremental CCG Parser.
Bharat Ram Ambati, **Siva Reddy**, Mark Steedman.
*NAACL, 2016.*

[C8] Paraphrase Generation from Latent-Variable PCFGs for Semantic Parsing.
Shashi Narayan, **Siva Reddy**, Shay B. Cohen.
*International Natural Language Generation Conference, 2016.*

[C7] Semantic Word Sketches.
Diana McCarthy, Adam Kilgarriff, Milos Jakubícek, **Siva Reddy**.
*Corpus Linguistics, 2015.*

[C6] Word Sketches for Turkish.
Bharat Ram Ambati, **Siva Reddy**, Adam Kilgarriff.
*LREC, 2012.*

[C5] Dynamic and Static Prototype Vectors for Semantic Composition.
**Siva Reddy**, Ioannis P. Klapaftis, Diana McCarthy, Suresh Manandhar
*IJCNLP, 2011,* Best Paper Award.

[C4] An Empirical Study on Compositionality in Compound Nouns.
**Siva Reddy**, Diana McCarthy, Suresh Manandhar.
*IJCNLP, 2011.*

[C3] A corpus factory for many languages.
Adam Kilgarriff, **Siva Reddy**, Jan Pomikálek, Avinesh PVS.
*LREC, 2010.*

[C2] Hindi semantic category labeling using semantic relatedness measures.
**Siva Reddy**, Abhilash Inumella, Navjyoti Singh, Rajeev Sangal.
*Global WordNet Conference, 2010.*

[C1] All Words Unsupervised Semantic Category Labeling for Hindi.
**Siva Reddy**, Abhilash Inumella, Rajeev Sangal, Soma Paul.
*RANLP, 2009.*


**Workshop Publications**

[W13] Societal Alignment Frameworks Can Improve LLM Alignment
Karolina Stańczak, Nicholas Meade, Mehar Bhatia, Hao Zhou, Konstantin Böttinger, Jeremy Barnes, Jack Stanley, **Siva Reddy**
Bi-Align Workshop, ICLR 2025

[W12] How to Get Your LLM to Generate Challenging Problems for Evaluation
Arkil Patel, **Siva Reddy**, Dzmitry Bahdanau
LLM-eval Workshop, NeurIPS 2025

[W11] In-Context Learning for Text Classification with Many Labels
Aristides Milios, **Siva Reddy**, Dzmitry Bahdanau
GenBench Workshop, EMNLP 2023

[W10] MeDAL: Medical Abbreviation Disambiguation Dataset for Natural Language Understanding Pretraining
Zhi Wen, Xing Han Lu, **Siva Reddy**
Clinical NLP workshop, EMNLP 2020

[W9] Multilingual Parsing from Raw Text to Universal Dependencies
Daniel Zeman et al.
CoNLL 2017 Shared Task, 2017

[W8] Universal Dependencies to Logical Forms with Negation Scope.
Federico Fancellu, **Siva Reddy**, Adam Lopez and Bonnie Webber.
*EACL — Workshop on Computational Semantics Beyond Events and Roles, 2017.*

[W7] DNN-based Speech Synthesis for Indian Languages from ASCII text.
Srikanth Ronanki, **Siva Reddy**, Bajibabu Bollepalli, Simon King.
*Speech Synthesis Workshop, 2016.*

[W6] Feed Corpus : An Ever Growing Up-To-Date Corpus.
Akshay Minocha, **Siva Reddy**, Adam Kilgarriff.
*Web as Corpus Workshop, 2013.*

[W5] DSS: text similarity using lexical alignments of form, distributional semantics and grammatical relations.
Diana McCarthy, Spandana Gella, **Siva Reddy**.
*SemEVal, 2012.*

[W4] Cross Language POS Taggers (and other Tools) for Indian Languages: An Experiment with Kannada using Telugu Resources.
**Siva Reddy**, Serge Sharoff.
*IJCNLP — Cross Lingual Information Access Workshop, 2011.*

[W3] Exemplar-based Word-Space Model for Compositionality Detection: Shared task system description.
**Siva Reddy**, Diana McCarthy, Suresh Manandhar, Spandana Gella.
*ACL — Distributional Semantics and Compositionality Workshop, 2011.*

[W2] IIITH: Domain specific word sense disambiguation.
**Siva Reddy**, Abhilash Inumella, Diana McCarthy, Mark Stevenson.
*SemEval, 2010.*

[W1] WSD as a distributed constraint optimization problem.
**Siva Reddy**, Abhilash Inumella.
*ACL Student Research Workshop, 2010.*

**Preprints/ Technical Reports**

[T9] Value Drifts: Tracing Value Alignment During LLM Post-Training
Mehar Bhatia, Shravan Nayak, Gaurav Kamath, Marius Mosbach, Karolina Stańczak, Vered Shwartz, **Siva Reddy**
*arXiv.*, 2025

[T8] Position: Build the web for agents, not agents for the web
Xing Han Lù, Gaurav Kamath, Marius Mosbach, **Siva Reddy**
*arXiv.*, 2025

[T7] Uncertainty Quantification of Large Language Models using Approximate Bayesian Computation
Mashrur Sharma, Arkil Patel, Zaeem De Souza, Seyed Amir Rahimi, **Siva Reddy**, Sreenath Madathil
*arXiv.*, 2025

[T6] ReTreever: Tree-based Coarse-to-Fine Representations for Retrieval
Shubham Gupta, Zichao Li, Tianyi Chen, Cem Subakan, **Siva Reddy**, Perouz Taslakian, Valentina Zantedeschi
*arXiv.*, 2025

[T5] A Compositional Typed Semantics for Universal Dependencies
Laurestine Bradford, Timothy O'Donnell, **Siva Reddy**
*arXiv:2403.01187*, 2024

[T4] When Word Order Matters: Unraveling Redundancy and Sequence Sensitivity In Language Models Through Information Theory
Xuanda Chen, **Siva Reddy**, Timothy O'Donnell
*arXiv:2402.18838*, 2024

[T3] Rosa: Random subspace adaptation for efficient fine-tuning
Marawan Hameed, Aristides Milios, **Siva Reddy**, Guillaume Rabusseau
*arXiv:2407.07802*, 2024

[T2] Modelling Latent Translations for Cross-Lingual Transfer
Edoardo Maria Ponti, Julia Kreutzer, Ivan Vulic, **Siva Reddy**.
*arXiv:2107.11353, 2021.*

[T1] A Data Elicitation Framework for Semantic Parsing.
Jianpeng Cheng, **Siva Reddy**, Mirella Lapata.
*arXiv:1812.10037, 2018.*

| | | |
|---|---|---|
| **Teaching** | Instructor, Natural Language Understanding with Deep Learning, McGill | F2025 |
| | Instructor, Graduate School Fundamentals, McGill | W2025 |
| | Instructor, From Natural Language to Data Science, McGill | W2025 |
| | Instructor, Large Language Models, McGill | F2024 |
| | Instructor, From Natural Language to Data Science, McGill | W2024 |
| | Instructor, Natural Language Understanding with Deep Learning, McGill | W2024 |
| | Instructor, From Natural Language to Data Science, McGill | W2023 |
| | Instructor, Natural Language Understanding with Deep Learning, McGill | F2022 |
| | Instructor, Natural Language Understanding with Deep Learning, McGill | W2022 |
| | Instructor, From Natural Language to Data Science, McGill | F2021 |
| | Instructor, Natural Language Understanding with Deep Learning, McGill | W2021 |
| | Instructor, From Natural Language to Data Science, McGill | F2020 |
| | Instructor, Natural Language Understanding with Deep Learning, McGill | W2020 |
| | | |
| | Advisor, Independent Project, COMP 396 | F2024 |
| | Advisor, Independent Project, LING 488 | F2024 |
| | Advisor, Independent Project, COMP 400 | F2022 |
| | Advisor, Independent Project, LING 488 | F2021 |
| | Advisor, Honors Project, LING 480 | 2021–2022 |
| | Advisor, Honors Project, COGS 444, | 2021–2022 |
| | Advisor, Independent Project, COMP 400 | W2021 |
| | | |
| | Non-McGill: | |
| | Retrieval-Augmented Generation, IVADO NLP workshop series | Apr 2024 |
| | Question Answering, IVADO NLP workshop series | Apr 2023 |
| | IVADO NLP course | Oct 2021 |
| | AI4Good Summer School | June 2022 |
| | From document to excerpt, IVADO NLP workshop series | Oct 2021 |
| | RNNs for NLP, AI4Good Summer School | June 2021 |
| | RNNs for NLP, AI4Good Summer School | June 2020 |

| | | |
|---|---|---|
| **Panels** | Panel on AI Safety | July 2025 |
| | AI Safety Social at ICML 2025 | |
| | Panelists: David Duvenaud, Aditi Raghunathan, Siva Reddy | |
| | | |
| | The Future of AI Safety | April 2025 |
| | Simons Workshop: Safety Guaranteed LLMs | |
| | Panelists: Yoshua Bengio, Roger Grosse, Geoffrey Irving, | |
| | Siva Reddy (moderator), Dawn Song | |
| | | |
| | Mila Panel on Safe and aligned AI | Oct 2024 |
| | Panelists: Yoshua Bengio, Aaron Courville, Tegan Maharaj, | |
| | Chris Pal (moderator), Siva Reddy, Dhanya Sridhar. | |
| | | |
| | Mila Lunch and Learn: Challenges faced by new faculty members. | Sep 2024 |
| | Panelists: Golnoosh Farnadi, Simon Lacoste-Julien, Joelle Pineau, | |
| | Siva Reddy, Dhanya Sridhar (moderator) | |

| | | |
|---|---|---|
| | Fireside Chat on LLMs: Opportunities and Risks<br>Schulich Leaders SLXCA event, McGill | Sep 2024 |
| | McGill Homecoming Panel: Language diversity and digital inclusion<br>Panelists: Carola Weil (moderator), María Sierra Córdoba Serrano,<br>Michael David Miller, Cecily Raynor, Siva Reddy<br>School of Continuing Studies | Oct 2024 |
| | EduGenAI Panel: Artificial Intelligence in Research<br>Panelists: Golnoosh Farnadi (moderator), Nikki Lobczowski,<br>Siva Reddy, Gabriel Vigliensoni | May 2024 |
| | McGill Homecoming Panel: AI in Education<br>Panelists: Eric D. Kolaczyk, Joelle Pineau, Doina Precup, Siva Reddy<br>Faculty of Science | Oct 2023 |
| | LingCon Computational Linguistics Conference for High Schoolers<br>Panelists: Margaret Mitchell, Siva Reddy, Jakob Uszkoreit<br>Fremont, California | Nov 2017 |
| **Media** | If Chat-GPT told you that you created a new kind of math<br>Let's Go, CBC Radio. | Sep 2025 |
| | Metacognition and Related Abilities of Large Language Models<br>Simons Institute for the Theory of Computing Blog. | Aug 2025 |
| | What vision AI could be used for in the future<br>Marketplace Podcast. | Aug 2025 |
| | AI generates harsher punishments for people who use Black dialect<br>Science News | Sep 2024 |
| | Trois risques concrets des modèles de langage<br>La Presse | Apr 2024 |
| | AI supercomputer in Canada<br>CBC interview. | Jan 2024 |
| | The Art of (AI) Conversation<br>CIFAR. | Dec 2019 |
| | Training a single AI model can emit as much carbon as five car lifetimes<br>MIT Technology review. | Jun 2019 |
| | The good, the bad, and the ugly of leaderboards and science<br>AI2 NLP Highlights Podcast | Nov 20 2018 |
| **Invited Talks** | Value Drifts: Tracing Value Alignment During LLM Post-Training<br>IVADO Workshop on Deploying Autonomous Agents: Lessons, Risks, and Real-World<br>Impact | Nov 2025 |
| | The state of web agents<br>IVADO Bootcamp on Agents | Aug 2025 |

The Path to Scalable and Safe Web Agents                          July 2025
Research on Agent Language Models (REALM) at ACL 2025

Linguistics from First Principles and LLMs                        July 2025
Dagstuhl Seminar on Linguistics and Language Models

Don't Take LLM Fundamentals for Granted                           July 2025
New In ML at ICML 2025

Data-efficient Training and Reasoning in MLMs and LLMs            June 2025
CIFAR AICan, Ste-Adèle

Jailbreaking Aligned Models, Reasoning Models and Agents
Simons Workshop: Safety Guaranteed LLMs                           April 2025
Singapore Alignment Workshop                                      April 2025

DeepSeek-R1 Thoughtology: <Thinking> about LLM Reasoning
Simons Workshop: The Future of Language Models and Transformers   Mar 2025
Mila Partners Symposium                                           Jun 2025

Learning to Reason with Generative Models
University of Waterloo                                            Dec 2024
UBC Future of NLP Workshop                                        Dec 2024

Paradoxes in Transformer Language Models: Causal Masking, Positional Encoding,
and Routing
CIFAR AICan, Banff                                                June 2024
UNC Chapel Hill                                                   Mar 2024
New York University                                               Feb 2024

ChatGPT, is natural language understanding solved?               Feb 2024
McGill Faculty of Science, Soup and Science

Holistic Evaluation of Language Models for Information Seeking
SemEval workshop, ACL 2023                                        July 2023

Do we still need inductive biases after Transformer language models?
Structured Prediction for NLP, ACL 2022                          May 2022
ENLSP workshop, NeurIPS 2022                                      Dec 2022

How grounded is document-grounded conversational AI?
DialDoc Workshop, ACL 2022                                       May 2022
Mila Partners Symposium                                          June 2022

Robustness of Neural Language Models to Language, Social and Cultural Complexity
Qatar Computing Research Institute                               Dec 2021
University of Dusseldorf                                         2021
George Mason University                                          2021

Unlikelihood-training and Back-training for robust natural language understanding
AKBC 2021                                                        2021
Cambridge University                                             2021
Samsung AI Centre, Toronto                                       July 2021

Challenges in building Natural Language Inferfaces
Mila TechAide, Montreal                                                    Apr 2021

Pathological Behaviours of Large Neural Models of Language
Uppsala University                                                         Mar 2021
Mila Tea Talk                                                              Nov 2020
National Research Council of Canada                                        Oct 2020

Measuring societal biases in pretrained language models                   Nov 2020
Montreal NLP Consortium

Building computers that understand and interact in human language         July 2020
McGill Faculty of Science, Soup and Science

Interacting with machines in natural language
Carnegie Mellon University                                                Mar 2019
University of California, Irvine                                          Mar 2019
University of Pennsylvania, Philadelphia                                  Mar 2019

CoQA: A Conversational Question Answering Challenge
Uber ConvAI Symposium                                                     Nov 28 2018
BayArea NLP & ML Meetup                                                   Oct 24 2018
Facebook AI Research                                                      Oct 19 2018
Massachusetts Institute of Technology                                     Oct 18 2018
University of California Los Angeles                                      Oct 05 2018
ISI, University of Southern California                                    Oct 04 2018
University of California Irvine                                           Oct 03 2018
Google Assistant and Dialog Workshop                                      Jun 26 2018

From Natural Language to Machine Interpretable Representations
University of Cambridge                                                   Mar 16 2018

The State of Natural Language Understanding: Past, Present and Future
AI Assistant Summit San Francisco                                         Jan 25 2018

Linguists-defined vs. Machine-induced Natural Language Structures for Executable
Semantic Parsing
University of California Berkeley                                         Oct 09 2017

Linguists-defined and Machine-induced Representations for Semantic Parsing
University of Cambridge                                                   Jun 15 2017

Towards a Compositional Typed Semantics for Universal Dependencies
Technical University of Darmstadt                                         Nov 23 2016
University of Washington                                                  Sep 09 2016
Microsoft Research Redmond                                                Sep 08 2016
Stanford University                                                       Aug 26 2016

Semantic Parsing with and without Question-Answer Pairs
University of Amsterdam                                                   Nov 21 2016
Google Research NYC                                                       Sep 19 2016
University of Massachusetts Amherst                                       Sep 16 2016
Allen Institute of Artificial Intelligence                               Sep 07 2016

**Current Supervision Activities**

- **Postdoc:**
  - Marius Mosbach (PhD at Saarland University; since Apr 2024)
  - Verna Dankers (PhD at University of Edinburgh; since June 2025)
- **PhD students:**
  - Vaibhav Adlakha (Computer Science, since Sep 2020)
  - Nicholas Meade (Computer Science, since Sep 2020)
  - Zichao Li (Computer Science, since Sep 2020; co-advised with Jackie Cheung)
  - Benno Krojer (Computer Science, since June 2021)
  - Xing Han Lu (Computer Science, since Sep 2021)
  - Parishad BehnamGhader (Computer Science, since Jan 2022)
  - Gaurav Kamath (Linguistics, since June 2022)
  - Arkil Patel (Computer Science, since Sep 2022; co-advised with Dzmitry Bahdanau)
  - Mehar Bhatia (Computer Science, since Sep 2024; co-advised with Vered Shwartz)
  - Aditi Khandelwal (Computer Science, since Sep 2024; co-advised with Golnoosh Farnadi)
  - Austin Kraft (Linguistics Eval 2 project, since May 2025)
  - Tomás Vergara Browne (Computer Science, since Sep 2025)
- **MSc:**
  - Jay Gala (Computer Science, Sep 2025; co-advised with Dzmitry Bahdanau)
  - Fengyuan Liu (Computer Science, Sep 2025; co-advised with Dzmitry Bahdanau)
  - Sreenath Madathil (Computer Science, Sep 2025)
  - Marzia Nouri (Computer Science, Jan 2026)
- **Undergrad:**
  - Ada Tur (Computer Science and Linguistics, since May 2024)
- **Research Assistants:**
  - Amirhossein Kazemnejad (Mila, since Dec 2023)
- **Interns:**
  - Charbel El Feghali (Mila, since Sep 2025)
  - Matteo Boglioni (Mila, since Oct 2025)

**Past Supervision Activities (Alumni)**

- **Postdoc:**
  - Edoardo Ponti (PhD at Cambridge University; Sep 2020 – July 2022) ⟶ Lecturer (Assistant Professor) at School of Informatics, University of Edinburgh
  - Eva Portelance (PhD at Stanford University; Jan 2023 – Aug 2024; co-advised with Timothy O'Donnell) ⟶ Assistant Professor at HEC Montreal
  - Karolina Stańczak (PhD at University of Copenhagen; Apr 2024 – Mar 2025) ⟶ Postdoctoral Researcher at the ETH AI Center.
- **PhD:**
  - Andreas Madsen (Computer Science, Sep 2020 – Dec 2024; co-supervised with Sarath Chandar) ⟶ Founding Research Scientist at Guide Labs, San Francisco.
- **MSc/MA students:**
  - Nathan Schucher (MSc Computer Science, Sep 2020 – Dec 2022) ⟶ Google DeepMind

- Devang Kulshreshtha (MSc Computer Science, Sep 2020 – June 2022) $\longrightarrow$ Amazon AWS AI Applied Scientist
- Emily Goodwin (MA Linguistics, Jan 2020 – May 2021; co-supervised with Timothy O'Donnell) $\longrightarrow$ PhD at Stanford University
- Clemencia Siro (Remote Thesis Supervision for the African Masters of Machine Intelligence, July 2020 – July 2021) $\longrightarrow$ PhD at Informatics Institute, University of Amsterdam Computer Science
- Amirhossein Kazemnejad (Computer Science, Jan 2021 – Sep 2023 $\longrightarrow$ RA in my group)
- Aristides Milios (Computer Science, since Sep 2021 – Dec 2023; co-advised with Dzmitry Bahdanau) $\longrightarrow$ PhD at Polytechnique Montréal
- **Undergrad students:**
  - Jasper Jian (NSERC SURA Computer Science, Apr 2022 – May 2023) $\longrightarrow$ PhD at Stanford University
  - Sivan Milton (Honors Linguistics, since Apr 2021) $\longrightarrow$ PhD at Edinburgh University
  - Elinor Poole-Dayan (Fall 2022; Winter 2021/23) $\longrightarrow$ Masters at MIT Media Lab
  - Neshma Metri (Honors Cognitive Science, Sep 2021 – Apr 2022) $\longrightarrow$ Data Analyst, Invera
  - Michelle Yang (Linguistics, Independent Project; Sep 2021 – Dec 2021; co-supervised with Morgan Sonderegger) $\longrightarrow$ Back to McGill
  - Nooshin Soheililangroodi (Computer Science, Sep 2024 – Dec 2024) $\longrightarrow$ Back to McGill
- **Interns/Visting Researchers:**
  - Imene Kerboua (visiting PhD student, Aug 2025 – Nov 2025)
  - Hanseok Oh (Intern, Jun 2025 – Nov 2025)
  - Dongchan Shin (Intern, Apr 2025 – Aug 2025) $\longrightarrow$ PhD at CS, University of Waterloo
  - Hadia Samil (RA, Sep 2024 – Jan 2025) $\longrightarrow$ RA with David Adelani
  - Sara Marjanović (visiting PhD student from the University of Copenhagen, Feb 2025 – May 2025)
  - Aravind Krishnan (visiting MSc student from Saarland University, Oct 2024 – Jan 2025)
  - Alejandra Zambrano (visiting PhD student from the Concordia University, Sep 2024 – Feb 2025)
  - Zdeněk Kasner (visiting PhD student from Charles University, Feb 2023 – May 2023)
  - Xuanda Chen (McGill Linguistics Eval 2, Sep 2023 – Apr 2024; co-advised with Tim O'Donnell)
  - Laurestine Bradford (McGill Linguistics Eval 1, Jan 2023 – Aug 2023; co-advised with Tim O'Donnell)
  - Nouha Dziri (visiting PhD student from Alberta, 2021 – 2022) $\longrightarrow$ Postdoc at AI2
  - Emanuele Bugliarello (visiting PhD student from Copenhagen from Dec 2021 – May 2022) $\longrightarrow$ Google DeepMind Internship
  - Victor Prokhorov (visiting PhD student from Cambridge from Aug 2021 – Dec 2021 $\longrightarrow$ Back to University of Cambridge
  - Lucas Torroba Hennigen (Internship from Oct 2020 – Jul 2021) $\longrightarrow$ PhD at MIT

Computer Science

- Arjun Akula (remote PhD student at UCLA from Oct 2019 – Apr 2021) ⟶ Google Research, LA
- Moin Nadeem (remote undergradue student at MIT from Oct 2019 – Apr 2020) ⟶ Back to MIT

**Students/Postdocs with Fellowships**

- Austin Kraft, Vanier Scholar (2025–2028; CAD $150,000)
- Arkil Patel, NSERC CGS D scholarship (2025–2028; CAD $120,000)
- Mehar Bhatia, FRQNT doctoral scholarship (2025–2028; CAD $100,000)
- Benno Krojer, Vanier Scholar (2023–2026; CAD $150,000)
- Nicholas Meade, NSERC CGS D scholarship (2023–2026; CAD $120,000)
- Xing Han Lu, NSERC CGS D scholarship (2023–2026; CAD $120,000)
- Arkil Patel, NSERC CGS Masters scholarship (2023; CAD $17,500);
- Gaurav Kamath, FRQSC doctoral scholarship (2023–2027; CAD $84,000)
- Laurestine Bradford, FRQSC doctoral scholarship (2023–2027; CAD $84,000)
- Parishad BehnamGhader, Mila EDI Excellence scholarship (2022-2024; CAD $10000)
- Edoardo Ponti, IVADO postdoctoral fellowship 2022 (CAD $70,000)
- Aristides Milios, NSERC CGS Masters (2021; CAD $17,500); McGill Trottier (2021; CAD $5000); IVADO MSc excellence scholarship (2022-2023; CAD $20,000)
- Laurestine Bradford, CRBLM (2022; CAD $6000)
- Xing Han Lu, IVADO MSc excellence scholarship (2021-2023; CAD $40,000); Quebec Hydro Fellowship (2021; CAD $10,000)
- Nicholas Meade, IVADO MSc excellence scholarship (2020-2022; CAD $40,000)

**Academic Service**

**Local Chair**
COLM 2025

**Program Chair**
IVADO Fall 2025 Thematic Semester on Autonomous Agents
People: Siva Reddy (PC), Yoshua Bengio (PC), Dawn Song, Yu Su , Alane Suhr, Victor Zhong, and Christopher Pal

**Action Editor:**
TACL Journal (2025 – current)
ACL Rolling Review (2021 – current)

**Ethics Chair**
COLM 2024

**Tutorial Chair**
ACL 2023

**Senior Area Chair**
EMNLP 2024
NAACL 2022

**Area Chair:**
EACL 2023
EMNLP 2022

AKBC 2022
EMNLP 2021
NAACL 2021
EMNLP 2020
ACL 2020
ACL 2019
AKBC 2019
CoNLL 2018

**Student research workshop: Faculty advisors**
ACL 2022

**Organizer**
Workshop: Capabilities and Safety of Agents, Oct 2025, IVADO
Bootcamp: Focusing on the Current State of Agents, Aug 2025, IVADO
Safety-Guaranteed LLMs, Simons Institute, UC Berkeley, 2025
Workshop on Vision Language Models For All, CVPR, 2025
Simons and IVADO Montreal Joint Program on LLMs and Agents, 2025
Bellairs workshop on Risks of Large Language Models, Mar 2024
Birds of Feather Session, ACL 2021
Document-grounded Dialogue and Conversational QA, ACL 2021
Structured and Unstructured KBs, AKBC 2020
Birds of Feather Session, ACL 2020
Deep Generative Models for Highly Structured Data, ICLR 2019
Stanford NLP Seminar Series, Jul 2017 – Aug 2018
Machine Learning for NLP Reading Group, Dec 2013 - Jun 2014

**Program Committee Member:**
SustaiNLP, EMNLP 2020
IntEx-SemPar, Interactive and Executable Semantic Parsing, EMNLP 2020
Natural Language Interfaces, ACL 2020
RepL4NLP, ACL 2020
MRQA: Machine Reading for Question Answering, EMNLP 2019
Machine Reading for Question Answering, ACL 2018
New Forms of Generalization in Deep Learning and NLP, NAACL 2018
Universal Dependencies, EMNLP 2018
Ethics in NLP, NAACL 2018
Knowledge Base Construction, Reasoning and Mining, WSDM 2018
Automated Knowledge Base Construction, NIPS 2017
Language Grounding for Robotics, ACL 2017
Representation Learning for NLP, ACL 2017
Universal Dependencies, NoDaLiDa 2017
Sense, Concept and Entity Representations and their Applications, EACL 2017
Ethics in NLP, EACL 2017
Web as Corpus, ACL 2016
Vector Space Modeling for NLP, NAACL 2015
Semantic Parsing, ACL 2014

**Reviewer**
NeurIPS 2025
ICLR 2025
TACL Journal (2021 – 2024)
ACL Rolling Review (2021 – current)

SCiL 2022,
ACM Computing Surveys Journal 2022
Regular reviewer of ACL, EMNLP, NAACL, EACL conferences until ARR in 2021
AKBC 2021
ICLR 2021
ICML 2019
AAAI 2018

**Leader:**
NLP Axis Leader for IVADO Apogee Canada Research Excellence for Robust, Intelligent and Responsible Artificial Intelligence, 2024-2026

| | |
|---|---|
| **Grant Agencies Reviewing** | CIFAR Catalyst Grants in AI Safety, 2025 |
| | Canada Research Chair, 2025 |
| | Ivado-Simons Travel Grants, 2025 |
| | NSERC Discovery Grant, 2024 |
| | NSERC Discovery Grant, 2023 |
| | NSERC Discovery Grant, 2021 |
| | Dutch Research Council, 2022 |
| | Israel Science Foundation, 2024 |
| | Israel Science Foundation, 2022 |
| | Israel Science Foundation, 2020 |
| | Canada Foundation for Innovation, 2024 |
| | Canada Foundation for Innovation, 2023 |
| | Canada Foundation for Innovation, 2020 |
| | MITACS Accelerate, 2022 |
| | MITACS Accelerate, 2021 |
| | MITACS Accelerate, 2020 |
| | IVADO PhD Excellence Scholarships, 2021 |
| | IVADO PhD Excellence Scholarships, 2020 |

| | | |
|---|---|---|
| **Other Service** | Pan-Canadian AI Strategy National Program Committee | 2022–2023 |
| | McGill Faculty of Arts Nominating Committee | 2023–current |
| | McGill Faculty of Arts EDI Committee Member, | Fall 2023 |
| | McGill SOCS PhD Committee | |
| | McGill SOCS Faculty Hiring Committee (2 positions) | 2022–2023 |
| | McGill SOCS Faculty Hiring Committee | 2021–2022 |
| | McGill SOCS AI/ML/Data Science Committee | 2022–2025 |
| | McGill SOCS EDI Committee member | 2021–2024 |
| | McGill SOCS PhD Committee | |
| | McGill Linguistics EDI Committee member | 2022–2023 |
| | McGill Linguistics Ombudsman Selection Committee | Fall 2022 |
| | McGill Linguistics Graduate Students Admissions | 2020–2025 |
| | Co-director, McGill-NLP | 2020–current |
| | Active member, cliq.ai (Computational Linguistics in Quebec) Consortium | 2024–current |
| | Mila Infrastructure and Data Team (IDT) | 2023–current |
| | Mila Core Faculty Committee | 2020–current |

**Final Thesis Examiner**

**PhD**
Jonathan Pilault, Polytechnique Montréal, 2025
Oscar Mañas, Université de Montréal, 2025
Clemencia Siro, University of Amsterdam, 2025
Lin Sheng-Chieh, University of Waterloo, 2024
Karin Sevegnani, Heriot-Watt University, 2023
Lucas Caccia, McGill University, 2023
Haji Mohammad Saleem, McGill University, 2021
Ali Emmani, McGill University, 2021
Inigo Lopez-Gazpio, University of the Basque Country, 2018

**Masters**
Yash More, McGill University, 2025
Shun Shao, University of Edinburgh, UK, 2023
Avinash Bhat, McGill University, 2023
Jingyi He, McGill University, 2022
Dora Jambor, McGill University, 2022
Zhi Wen, McGill University, 2021
Alika Utepova, McGill University, 2020
Jin Dong, McGill University, 2020
Matthew Grenander, McGill University, 2020

**Thesis Committee**

**PhD**
Soumya Sharma, PhD Committee Member, McGill University, 2025
Damaris Stephanie Ndjebayi, PhD Committee Member, McGill University, 2025
Senyu Li, PhD Committee Member, McGill University, 2025
Le Zhang, Predoc Examiner, Université de Montréal, 2025
Marawan Gamal, Predoc Examiner, Université de Montréal, 2024
Arjun Vaithalingam, PhD Comprehensive and Predoc Exam, Polytechnique, 2024
Zachary Yang, PhD Thesis Proposal Examiner, McGill University, 2024
Khaoula Chehbouni, PhD Committee Member, McGill University, 2024
Esmail Moghiseh, PhD Defense Committee Member, McGill University
Oscar Manas, Predoc Examiner, Université de Montréal, 2023
Maziar Sargordi, PhD Proposal Examiner, Polytechnique Montreal, 2023
Kellin Pelrine, PhD Committee Member, McGill University, 2023
Ian Porada, PhD Proposal Examiner, McGill University, 2023
Yang Ji, Pro-dean, McGill University, 2023
Saeed Najafi, PhD Committee Member, University of Alberta, 2022
Koustuv Sinha, PhD Committee Member, McGill University, 2022
Alexander Krolik, Pro-dean, McGill University, 2022
Kushal Arora, PhD Committee Member, McGill University, 2021
Jiang Bing'er, PhD Defense Chair, McGill University, 2021
Alexander Krolik, PhD Defense Chair, McGill University, 2020