

Sachin Kumar

kumar.1145@osu.edu • +1 (412) 494-8510 • <http://shocheen.com>

PROFESSIONAL EXPERIENCE

The Ohio State University, Computer Science and Engineering, Columbus, OH
Assistant Professor

Aug 2024 – Present

Allen Institute for Artificial Intelligence, Seattle, WA
Postdoctoral Researcher

Aug 2023 – Aug 2024

- Advisors: Profs. Noah Smith, Hannaneh Hajishirzi

University of Washington, Seattle, WA
Visiting Researcher

Jun 2021 – Aug 2024

EDUCATION

Carnegie Mellon University, Pittsburgh, PA, USA
Ph.D., Language Technologies Institute (LTI)

Sep 2017 – Sep 2023

- Advisor: Prof. Yulia Tsvetkov
- Thesis: Optimization Methods for Improving Diversity in Language Technologies

Indian Institute of Technology, Kharagpur, India

Bachelor of Technology (Hons) and Master of Technology, Computer Science

Jul 2011 – Jun 2016

- Gold Medalist (*Summa Cum Laude*)

PREPRINTS

Orevaoghene Ahia, Martijn Bartelds, Kabir Ahuja, Hila Gonen, Valentin Hofmann, Siddhant Arora, Shuyue Stella Li, Vishal Puttagunta, Mofetoluwa Adeyemi, Charishma Buchiredy, Ben Walls, Noah Bennett, Shinji Watanabe, Noah A. Smith, Yulia Tsvetkov, **Sachin Kumar**, “BLAB: Brutally Long Audio Bench”, 2025.

Abraham Toluwase Owodunni, **Sachin Kumar**, “Continually Adding New Languages to Multilingual Language Models”, 2025.

Abraham Toluwase Owodunni, Orevaoghene Ahia, **Sachin Kumar**, “FLEXITOKENS: Flexible Tokenization for Evolving Language Models”, 2025.

Tarun Kathuria, **Sachin Kumar**, “Leveraging Pretrained Language Models as Energy Functions for Glauber Dynamics Text Diffusion”, 2025.

Jiwoo Park, Ruoqi Liu, Avani Jagdale, Andrew Srisuwananukorn, Ping Zhang, **Sachin Kumar**, “ClinicalTrialsHub: Bridging Registries and Literature for Comprehensive Clinical Trial Access”, 2025.

CONFERENCE PUBLICATIONS

Zhouhang Xie, Junda Wu, Yiran Shen, Raghav Jain, Yu Xia, Xintong Li, Aaron Chang, Ryan A. Rossi, Tong Yu, **Sachin Kumar**, Bodhisattwa Prasad Majumder, Jingbo Shang, Prithviraj Ammanabrolu, Julian McAuley, “A Survey on Personalized and Pluralistic Preference Alignment in Large Language Models”, Conference on Language Modeling (COLM 2025).

Jaesung Tae, Hamish Ivison, **Sachin Kumar**, Arman Cohan, “SAD-LM: A Large-Scale Generalist Diffusion Language Model”, 2025 Conference of the Association for Computational Linguistics (ACL 2025). Oral (top 8%).

Patrick Queiroz Da Silva, Hari Sethuraman, Dheeraj Rajagopal, Hannaneh Hajishirzi, **Sachin Kumar**, “Steering off Course: Reliability Challenges in Steering Language Models”, 2025 Conference of the Association for Computational Linguistics (ACL 2025). Oral (top 8%); Panel (top 0.8%). Also presented at the Actionable Interpretability Workshop at ICML 2025.

Lester James V. Miranda, Yizhong Wang, Yanai Elazar, **Sachin Kumar**, Valentina Pyatkin, Faeze Brahman, Noah A. Smith, Hannaneh Hajishirzi, Pradeep Dasigi, “Hybrid Preferences: Learning to Route Instances for Human vs. AI Feedback”, 2025 Conference of the Association for Computational Linguistics (ACL 2025).

Sachin Kumar*, Chan Young Park*, Yulia Tsvetkov, Noah A. Smith, Hannaneh Hajishirzi, “ComPO: Community Preferences for Language Model Personalization”. 2025 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2025).

- Harsh Kohli, **Sachin Kumar**, Huan Sun. “GroundCocoa: A Benchmark for Evaluating Compositional & Conditional Reasoning in Language Models”. 2025 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2025).
- Nathan Lambert, Valentina Pyatkin, Jacob Morrison, LJ Miranda, Bill Yuchen Lin, Khyathi Chandu, Nouha Dziri, **Sachin Kumar**, Tom Zick, Yejin Choi, Noah A. Smith, Hannaneh Hajishirzi, “RewardBench: Evaluating Reward Models for Language Modeling”. 2025 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2025) Findings.
- Faeze Brahman*, **Sachin Kumar***, Vidhisha Balachandran, Pradeep Dasigi, Valentina Pyatkin, Abhilasha Ravichander, Sarah Wiegrefe, Nouha Dziri, Khyathi Chandu, Jack Hessel, Yulia Tsvetkov, Noah A. Smith, Yejin Choi, Hannaneh Hajishirzi, “The Art of Saying No: Contextual Noncompliance in Language Models”, Thirty-eighth Conference on Neural Information Processing Systems (NeurIPS 2024): Datasets and Benchmarks.
- Liwei Jiang, Kavel Rao, Seungju Han, Allyson Ettinger, Faeze Brahman, **Sachin Kumar**, Niloofar Mireshghallah, Ximing Lu, Maarten Sap, Yejin Choi, Nouha Dziri, “WildTeaming at Scale: From In-the-Wild Jailbreaks to (Adversarially) Safer Language Models”, Thirty-eighth Conference on Neural Information Processing Systems (NeurIPS 2024).
- Orevaoghene Ahia, **Sachin Kumar**, Hila Gonen, Valentin Hoffman, Tomasz Limisiewicz, Yulia Tsvetkov, Noah A. Smith, “MAGNET: Improving the Multilingual Fairness of Language Models with Adaptive Gradient-Based Tokenization”, Thirty-eighth Conference on Neural Information Processing Systems (NeurIPS 2024).
- Luca Soldaini, Rodney Kinney, Akshita Bhagia, Dustin Schwenk, David Atkinson, Russell Authur, Ben Bogin, Khyathi Chandu, Jennifer Dumas, Yanai Elazar, Valentin Hofmann, Ananya Harsh Jha, **Sachin Kumar**, Li Lucy, Xinxu Lyu, Nathan Lambert, Ian Magnusson, Jacob Morrison, Niklas Muennighoff, Aakanksha Naik, Crystal Nam, Matthew E. Peters, Abhilasha Ravichander, Kyle Richardson, Zejiang Shen, Emma Strubell, Nishant Subramani, Oyvind Tafjord, Pete Walsh, Luke Zettlemoyer, Noah A. Smith, Hannaneh Hajishirzi, Iz Beltagy, Dirk Groeneveld, Jesse Dodge, Kyle Lo, “Dolma: an Open Corpus of Three Trillion Tokens for Language Model Pretraining Research”, 2024 Conference of the Association for Computational Linguistics (ACL 2024). *Best Resource Paper Award*.
- YuHan Liu, Shangbin Feng, Xiaochuang Han, Vidhisha Balachandran, Chan Young Park, **Sachin Kumar**, Yulia Tsvetkov, “What Constitutes a Faithful Summary? Preserving Author Perspectives in News Summarization”, 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024).
- Xiaochuang Han, **Sachin Kumar**, Yulia Tsvetkov, Marjan Ghazvininejad, “SSD-2: Scaling and Inference-time Fusion of Diffusion Language Models”, 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024).
- Sachin Kumar**, Chan Young Park, Yulia Tsvetkov, “Gen-Z: Generative Zero-Shot Text Classification with Contextualized Label Description”, 12th International Conference on Learning Representations (ICLR 2024).
- Orevaoghene Ahia, **Sachin Kumar**, Hila Gonen, Jungo Kasai, David R. Mortensen, Noah A. Smith, Yulia Tsvetkov, “Do All Languages Cost the Same? Tokenization in the Era of Commercial Language Models”, 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023).
- Xiaochuang Han, **Sachin Kumar**, Yulia Tsvetkov, “SSD-LM: Semi-autoregressive Simplex-based Diffusion Language Model for Text Generation and Modular Control”, 2023 Conference on Association of Computational Linguistics (ACL 2023).
- Melanie Sclar, **Sachin Kumar**, Peter West, Alane Suhr, Yejin Choi and Yulia Tsvetkov, “Minding Language Models’ Theory of Mind: A Plug-and-Play Multi-Character Belief Tracker”, 2023 Conference on Association of Computational Linguistics (ACL 2023). *Outstanding Paper Award*.
- Tianxing He*, Jingyu Zhang*, Tianle Wang, **Sachin Kumar**, Kyunghyun Cho, James Glass, Yulia Tsvetkov, “On the Blind Spots of Text Generation Evaluation Metrics Based on Pretrained Language Models”, 2023 Conference on Association of Computational Linguistics (ACL 2023). *equal contribution

Sachin Kumar*, Vidhisha Balachandran*, Lucille Njoo, Antonios Anastasopoulos, Yulia Tsvetkov, “Language Generation Models Can Cause Harm: So What Can We Do About It? An Actionable Survey”, 2023 Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023).
*equal contribution

Sachin Kumar, Biswajit Paria, Yulia Tsvetkov, “Gradient-based Constrained Sampling from Language Models”, 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022).

Melanie Sclar, Peter West, **Sachin Kumar**, Yulia Tsvetkov and Yejin Choi, “Reference-Free Sentence Summarization with Sharper Controllability through Symbolic Knowledge Distillation”, 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022).

Sachin Kumar, Eric Malmi, Aliaksei Severyn, Yulia Tsvetkov, “Controlled Text Generation as Continuous Optimization with Multiple Constraints”, Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021)

Sachin Kumar, Antonios Anastasopoulos, Shuly Wintner, Yulia Tsvetkov, “Machine Translation into Low-Resource Language Varieties, 2021 Conference on Association of Computational Linguistics (ACL 2021)

Sachin Kumar, Shuly Wintner, Noah A. Smith, Yulia Tsvetkov, “Topics to Avoid: Demoting Latent Confounds in Text Classification”, 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP 2019).

Sachin Kumar and Yulia Tsvetkov, “Von Mises-Fisher Loss for Training Sequence to Sequence Models with Continuous Outputs”, 7th International Conference on Learning Representations (ICLR) 2019. (An earlier version presented at ICML Workshop on Theoretical Foundations and Applications of Deep Generative Models, 2018).

Shreshtha Mundra*, **Sachin Kumar***, Manjira Sinha, Sandya Mannarswamy, “Mining & Summarizing E-petitions for Enhanced Understanding of Public Opinion”, In Proceedings of the International Conference on Information and Knowledge Management (CIKM) 2018
*equal contribution

Sachin Kumar, Soumen Chakrabarti, Shourya Roy, “Earth Mover Distance Pooling over Siamese LSTMs for Automatic Short Answer Grading”. In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI) 2017.

WORKSHOP PUBLICATIONS

Carolina Hatanpää, Noah A. Smith, **Sachin Kumar**, “On Distributional Robustness of In-Context Learning for Text Classification”, 2nd Workshop on Test-Time Adaptation: Putting Updates to the Test! (PUT) at ICML 2025.

Monisha Jegadeesan, **Sachin Kumar**, John Wieting, Yulia Tsvetkov, “Improving the Diversity of Unsupervised Paraphrasing with Embedding Outputs”, Workshop on Multilingual Representation Learning at EMNLP 2021.

Lidia Kidane, **Sachin Kumar**, Yulia Tsvetkov, “An Exploration of Data Augmentation Techniques for Improving English to Tigrinya Translation”, The Second AfricaNLP Workshop at EACL 2021.

Sachin Kumar, Yulia Tsvetkov, “End-to-End Differentiable GANs for Text Generation”, I Can’t Believe It’s Not Better! Workshop at NeurIPS 2020.

Zi-Yi Dou, **Sachin Kumar**, Yulia Tsvetkov, “A Deep Reinforced Model for Cross-Lingual Summarization with Bilingual Semantic Similarity Reward”, Workshop on Neural Generation and Translation (WNGT) at ACL 2020.

Gayatri Bhat, **Sachin Kumar**, Yulia Tsvetkov, “A Margin-based Loss with Synthetic Negative Samples for Continuous-output Machine Translation”, Workshop on Neural Generation and Translation (WNGT) at EMNLP 2019.

ADVISING

Abraham Owodunni , PhD at OSU	Aug 2024 –
▪ Multilingual NLP	
Hanane Moussa , OSU PhD Student (co-advised with Huan Sun)	Jan 2025 –
▪ AI4Science	

	Zishuo Zheng , OSU PhD Student ▪ (Test-time Scaling / Reasoning)	Jan 2025 –
	Patrick Da Silva , MS at OSU, PhD at OSU (Fall 2025) ▪ Post Training LMs, AI4Science	Aug 2024 –
	Doeun Lee , MS at OSU ▪ Science Writing Assistance, Medical Benchmarking of LLMs	Sep 2024 –
	Aaron Jencks , MS at OSU ▪ Multilingual NLP, Text-Speech Models	Sep 2024 –
	Jiwoo Park , MS at OSU ▪ Medical Information Extraction	Jan 2025 –
	Nikhil Pavan Kanaka , MS at OSU ▪ Grokking	Aug 2024 – Dec 2024
MENTORING	Yusen Peng, OSU BS Student (Multimodal LMs)	May 2025 – Present
	Jason Su, OSU BS Student (AI Safety)	May 2025 – Present
	Yang Wang, OSU MS Student (AI Safety)	May 2025 – Present
	Roosbeh Nahavandi, OSU PhD Student (RLHF)	Aug 2024 – Present
	Orevaoghene Ahia, UW CSE PhD Student (Multilingual Multimodal LLMs)	Mar 2023 – Present
	Ben Walls, OSU BS Student (Evaluating Speech LMs)	Dec 2024 – Present
	Noah Bennett, OSU BS Student (Evaluating Speech LMs)	Dec 2024 – Present
	Vishal Puttagunta, OSU BS Student (Evaluating Speech LMs)	Dec 2024 – Present
	Charishma Buchireddy, OSU BS Student (Evaluating Speech LMs)	Dec 2024 – Present
	Harsh Kohli, OSU PhD Student (LoRA and Model Merging)	Aug 2024 – Dec 2024
	Hari Sethuraman, UW CSE BS Student (Mechanistic Interpretability)	Sep 2023 – Feb 2025
	Carolina Hatanpaa, Predoctoral Researcher	Sep 2023 – Aug 2024
	Melanie Sclar, UW CSE PhD Student	May 2021 – May 2023
	Daniel Saelid, UW CSE BS Student	Sep 2022 – Dec 2022
	Lidia Kidane, AIMS Rwanda MS Student	Sep 2020 – Apr 2021
	Monisha Jagadeesan, IIT Madras BS Student	Mar 2020 – May 2021
	Jiayun Wu, Tsinghua University BS Student	May 2021 – Sep 2021
	Yue Yin, CMU MS Student	May 2020 – Aug 2020
	Zi-Yi Dou, CMU MS Student	Aug 2019 – Dec 2019
	Aditya Siddhant, CMU MS Student	Aug 2018 – Dec 2018
AWARDS & HONORS	OSU Kirwan Institute SEED Grant	2025
	AI2 Faculty Research Award	2024
	OSU Kirwan Institute SEED Grant	2024
	Best Resource Paper Award (ACL)	2024
	Outstanding Paper Award (ACL)	2023
	Recipient of Google PhD Fellowship	2022
	Google Faculty Research Award (proposal co-written with Prof. Yulia Tsvetkov)	2020
	Amazon Machine Learning Research Award (proposal co-written with Prof. Yulia Tsvetkov)	2019
	Prime Minister of India Gold Medal at IIT Kharagpur (Summa Cum Laude)	2016
	Recipient of B. P. Poddar Scholarship at IIT Kharagpur	2015
TALKS	Improving the Multilingual Fairness of Language Models <i>University of Pennsylvania (NLP Seminar)</i> <i>The Ohio State University (Student Seminar)</i>	March 2025 December 2024
	The Art of Saying No: Contextual Noncompliance in Language Models <i>MilaNLP, Bocconi University</i>	April 2024
	NLP in the time of Generative Models: Centering the Role of People <i>Indian Symposium on Machine Learning (IndoML)</i> <i>The Ohio State University</i> <i>Boston University</i> <i>George Washington University</i>	December 2023 March 2023 March 2023 March 2023

	<i>Pennsylvania State University, Harrisburg</i> <i>Allen Institute for Artificial Intelligence</i>	March 2023 Feb 2023
	Gradient Based Controlled Text Generation <i>EMNLP Conference (virtual)</i> <i>Invited Talk at Google</i> <i>University of Washington NLP Retreat</i>	December 2022 October 2022 September 2022
	Controlled Text Generation as Continuous Optimization with Multiple Constraints <i>NeurIPS Conference (virtual)</i>	December 2021
	Machine Translation into Low-Resource Language Varieties <i>ACL Conference (virtual)</i>	August 2021
	Text Generation into Continuous Spaces <i>Invited Talk at G-Research</i> <i>CMU LTI Student Research Symposium</i>	November 2020 August 2018
INDUSTRY EXPERIENCE	Facebook AI Research, Seattle, WA <i>Research Intern</i> Pretraining masked language models via predicting into continuous spaces. ▪ Mentors: Luke Zettlemoyer, Emma Strubell	May 2020 – Aug 2020
	Facebook Inc, Menlo Park, CA <i>Research Intern, Facebook Assistant</i> Adversarial methods for developing task-oriented dialogue policy. ▪ Mentors: Bing Liu, Pararth Shah, Honglei Liu	May 2019 – Aug 2019
	Xerox Research Center India, Bangalore <i>Pre-doctoral Researcher</i> ▪ Developed training algorithms for automatically grading short essays. ▪ Mining and summarizing e-petitions to identify public opinions about urban issues in Bangalore ▪ Time series modelling of buying patterns of prepaid card users	Jul 2016 – Jul 2017
	Xerox Research Center India, Bangalore <i>Summer Intern</i> ▪ Developed text similarity metrics for grading short textual answers containing special symbols and chemical equations borrowing techniques from chemical information retrieval.	May 2015 – Aug 2015
	IBM India Research Lab, Bangalore <i>Summer Intern</i> ▪ Calibrating binary classifiers via Gaussian Mixture Models. ▪ Awarded best internship project out of 40 interns.	May 2014 – Jul 2014
TEACHING	COURSE INSTRUCTOR The Ohio State University Foundations of Speech and Language Processing (Graduate Course) Advanced Topics in Natural Language Processing (Graduate Seminar Course)	Spring 2025 Fall 2024
	GUEST LECTURES KAIST, South Korea Mitigating Societal Harms in Large Language Models – Introduction to NLP	March 2024
	University of Washington, Seattle Introduction to AI – Undergraduate AI Course Introduction to NLP – AI for Electrical Engineers Course Neural Sequence Labeling – Undergraduate NLP Course	Nov 2023 Nov 2022 April 2022 & Nov 2022
	TEACHING ASSISTANT Carnegie Mellon University Multilingual NLP (Graduate Course), <i>Teaching Assistant</i> , ▪ Guest Lecture on Unsupervised Text Generation Algorithms for NLP (Graduate Course), <i>Teaching Assistant</i> ▪ Guest Lectures on Sequence Labeling and Structured Prediction.	Fall 2020 Fall 2018 & Fall 2019

	Indian Institute of Technology, Kharagpur Software Engineering (Undergraduate Course), <i>Teaching Assistant</i> Compilers (Undergraduate Course), <i>Teaching Assistant</i>	Spring 2016 Fall 2015
TEACHING BEYOND UNIVERSITY	TUTORIAL Mitigating Societal Harms of Large Language Models https://llm-harm-mitigation.github.io/tutorial/ <i>The Web Conference 2022 – co-presenter</i> <i>EMNLP 2023 – co-presenter</i>	April 2022 December 2023
SERVICE	WORKSHOP ORGANIZER TokShop: Workshop on Tokenization ICML 2025 (Upcoming) CustomNLP4U: Workshop on Customizable NLP EMNLP 2024 SENIOR AREA CHAIR ▪ ACL ARR AREA CHAIR ▪ ACL ARR ▪ ICLR REVIEWING Conferences ▪ ACL Rolling Review (ARR) ▪ ICML, NeurIPS ▪ EACL, ICLR ▪ ACL, EMNLP, NAACL ▪ AAAI Journal ▪ Journal for AI Research Workshops ▪ ACL Student Research Workshop ▪ COLING Workshop on Multimodal, Multipurpose, and Massive-Scale Models ▪ ACL Workshop on NLP for Positive Impact ▪ ACL Workshop on Subword and Character Level Models for NLP (SCLeM) ▪ ACL Workshop on Noisy User Generated Data (W-NUT)	July 2025 November 2024 May 2025 June, December 2024 October 2024 2021–2024 2019–2024 2021–2023 2018–2022 2020–2021 2022 2019–2022 2022 2021 2018 2018
	COMMUNITY SERVICE Carnegie Mellon University ▪ LTI Peer Mentoring Program (co-organizer) ▪ CMU AI Undergraduate Mentoring Program ▪ LTI Graduate Admissions Committee ▪ CMU Prospective Graduate Application Support program University of Washington, Seattle ▪ UW NLP Retreat & Research Symposium (co-organizer) Queer in AI ▪ Prospective Graduate Application Support program EMNLP 2021 ▪ Session Chair	Fall 2021–Spring 2023 Spring 2020–Spring 2022 2021 2020, 2021 2022 2021 2021
MISCELLANEOUS	Computer Languages: Python, Java, C++, C Languages: Native speaker of Hindi, proficient in English and Punjabi	