

Bhuwan Dhingra

Assistant Professor, Computer Science, Duke
Research Scientist, Google Deepmind
bdhingra@cs.duke.edu

Website: <https://users.cs.duke.edu/~bdhingra/>
Scholar: [link](#) (h-index = 30)

Experience

2021-Present Assistant Professor, Computer Science, Duke University, Durham USA
2020-Present Research Scientist, Google Brain → Deepmind, San Francisco USA
2019-2020 Student Researcher, Google Research, Pittsburgh USA
2013-2015 Systems Engineer, Qualcomm Research, San Diego USA

Education

2015-2020 Ph.D. in Language Technologies, Carnegie Mellon University
 Advisers: William W. Cohen and Ruslan Salakhutdinov ([thesis](#))
2008-2013 B.Tech + M.Tech., Electrical Engineering, Indian Institute of Technology, Kanpur

Funding

- [1] ALTER-Math: AI-augmented Learning by Teaching to Enhance and Renovate Math Learning
Learning Engineering Virtual Institute (2023-27), (Total: \$2.5M, Duke: \$400K)
- [2] Scientific search using concept graphs
Gift from Proctor & Gamble (2022-24), (\$125K)
- [3] Ask the Experts: Generating Question Answer Pairs for Addressing Information Deficits about Vaccines
NSF CISE Core Medium (2022-26), (Total: \$1, 15M, Duke: \$957K)
- [4] Long-Form Question Answering via Collaborative Writing
Amazon AWS AI Award (2022-23), (\$35K)
- [5] Assessing the Legibility of Obfuscated Texts for Robust Natural Language Processing
Arts & Sciences Council Committee on Faculty Research (2022-23), (\$4.8K)
- [6] Using NLP & Other Tools to Address Misinformation with Vaccine-Hesitant Communities
Google Faculty Award (2021-22), (Total: \$80K, Per PI: \$20K)

Teaching

Spring 2024	CS572, Introduction to NLP (<i>graduate, class size: 75</i>)	Duke University
Spring 2024	UNIV103, Let's Talk About Digital You (<i>undergraduate</i>)	Duke University
Fall 2023	CS590, Generative Models (<i>graduate, class size: 24</i>)	Duke University
Fall 2022	CS572, Introduction to NLP (<i>graduate, class size: 56</i>)	Duke University
Spring 2022	CS590, Advanced NLP (<i>graduate, class size: 12</i>)	Duke University
Fall 2021	CS590, Introduction to NLP (<i>graduate, class size: 42</i>)	Duke University

Publications

Journals

- [1] VaxConcerns: A taxonomy of vaccine concerns and misinformation with Crowdsorce-Viability. Rickard Stureborg, Jenna Nichols, Bhuwan Dhingra, Jun Yang, Walter Orenstein, Robert A. Bednarczyk, and Lavanya Vasudevan. Vaccine 42, no. 10 (2024): 2672-2679.
- [2] Time-Aware Language Models as Temporal Knowledge Bases. Bhuwan Dhingra, Jeremy R. Cole, Julian Eisenschlos, Daniel Gillick, Jacob Eisenstein, William W. Cohen. Transactions of the Association of Computational Linguistics (TACL, 2022).
- [3] Evaluating Explanations: How much do explanations from the teacher aid students? Danish Pruthi, Bhuwan Dhingra, Livio Baldini Soares, Michael Collins, Zachary C Lipton, Graham Neubig, William W. Cohen. Transactions of the Association of Computational Linguistics (TACL, 2022).

Refereed Conferences

- [1] Benchmarking LLMs on Extracting Polymer Nanocomposite Samples. Ghazal Khalighinejad, Defne Circi, Cate Brinson, Bhuwan Dhingra. Findings of the 62nd Annual Meeting of Association of Computational Linguistics (ACL, 2024).
- [2] Sequence Reducible Holdout Loss for Language Model Pretraining. Raghuv eer Thirukovalluru, Nicholas Monath, Bhuwan Dhingra, and Sam Wiseman. Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING, 2024).
- [3] Tailoring Vaccine Messaging with Common-Ground Opinions. Rickard Stureborg, Sanxing Chen, Ruoyu Xie, Aayushi Patel, Christopher Li, Chloe Zhu, Tingnan Hu, Jun Yang, Bhuwan Dhingra. Findings of Annual Conference of the North American Chapter of the ACL (NAACL, 2024).
- [4] SumCSE: Summary as a transformation for Contrastive Learning. Raghuv eer Thirukovalluru, Xiaolan Wang, Jun Chen, Shuyang Li, Jie Lei, Rong Jin, Bhuwan Dhingra. Findings of Annual Conference of the North American Chapter of the ACL (NAACL, 2024).
- [5] Selectively Answering Ambiguous Questions. Jeremy Cole, Michael Zhang, Daniel Gillick, Julian Eisenschlos, Bhuwan Dhingra, Jacob Eisenstein. Conference on Empirical Methods in Natural Language Processing (EMNLP, 2023).
- [6] Valla: Standardizing and Benchmarking Authorship Attribution and Verification Through Empirical Evaluation and Comparative Analysis. Jacob Tyo, Bhuwan Dhingra, Zachary Lipton. Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL, 2023).
- [7] Interface Design for Crowdsourcing Hierarchical Multi-Label Text Annotations. Rickard Stureborg, Bhuwan Dhingra, Jun Yang. CHI Conference on Human Factors in Computing Systems (CHI, 2023).
- [8] DIFFQG: Generating Questions to Summarize Factual Changes. Jeremy Cole, Palak Jain, Julian Eisenschlos, Michael Zhang, Eunsol Choi, Bhuwan Dhingra. Conference of the European Chapter of the Association for Computational Linguistics (EACL, 2023).

- [9] Learning the Legibility of Visual Text Perturbations.
Dev Seth, Rickard Stureborg, Danish Pruthi, Bhuwan Dhingra.
Conference of the European Chapter of the Association for Computational Linguistics (EACL, 2023).
- [10] Salient Span Masking for Temporal Understanding.
Jeremy Cole, Aditi Chaudhary, Bhuwan Dhingra, Partha Talukdar.
Conference of the European Chapter of the Association for Computational Linguistics (EACL, 2023).
- [11] ASQA: Factoid Questions Meet Long-Form Answers.
Ivan Stelmakh, Yi Luan, Bhuwan Dhingra, Ming-Wei Chang.
Conference on Empirical Methods in Natural Language Processing (EMNLP, 2022).
- [12] Reasoning over virtual knowledge bases with open predicate relations.
Haitian Sun, Pat Verga, Bhuwan Dhingra, Ruslan Salakhutdinov, William W. Cohen.
The Thirty-eighth International Conference on Machine Learning (ICML, 2021).
- [13] Fool Me Twice: Entailment from Wikipedia Gamification.
Julian Eisenschlos, Bhuwan Dhingra, Jannis Bulian, Benjamin Börschinger, Jordan Boyd-Graber.
19th Annual Conference of the North American Chapter of the ACL (NAACL, 2021).
- [14] Differentiable open-ended commonsense reasoning.
Bill Yuchen Lin, Haitian Sun, Bhuwan Dhingra, Manzil Zaheer, Xiang Ren, William W. Cohen
19th Annual Conference of the North American Chapter of the ACL (NAACL, 2021).
- [15] ToTTo: A Controlled Table-To-Text Generation Dataset.
Ankur P. Parikh, Xuezhi Wang, Sebastian Gehrmann, Manaal Faruqui,
Bhuwan Dhingra, Diyi Yang, Dipanjan Das.
Conference on Empirical Methods in Natural Language Processing (EMNLP, 2020).
- [16] Weakly- and Semi-supervised Evidence Extraction.
Danish Pruthi, Bhuwan Dhingra, Graham Neubig, Zachary C. Lipton.
Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP, 2020).
- [17] Differentiable Reasoning over a Virtual Knowledge Base.
Bhuwan Dhingra, Manzil Zaheer, Vidhisha Balachandran, Graham Neubig,
Ruslan Salakhutdinov, William W. Cohen.
International Conference on Learning Representations (ICLR, 2020).
- [18] Learning to Deceive with Attention-Based Explanations.
Danish Pruthi, Mansi Gupta, Bhuwan Dhingra, Graham Neubig, Zachary C. Lipton.
58th annual meeting of Association of Computational Linguistics (ACL, 2020).
- [19] PubMedQA: A Dataset for Biomedical Research Question Answering
Qiao Jin, Bhuwan Dhingra, Zhengping Liu, William W. Cohen, Xinghua Lu.
Conference on Empirical Methods in Natural Language Processing (EMNLP, 2019).
- [20] Handling Divergent Reference Texts when Evaluating Table-to-Text Generation.
Bhuwan Dhingra, Manaal Faruqui, Ankur Parikh, Ming-Wei Chang,
Dipanjan Das, William W. Cohen.
57th annual meeting of Association of Computational Linguistics (ACL, 2019).
- [21] Combating Adversarial Misspellings with Robust Word Recognition.
Danish Pruthi, Bhuwan Dhingra, Zachary C. Lipton.
57th annual meeting of Association of Computational Linguistics (ACL, 2019).
- [22] Text Generation with Exemplar-based Adaptive Decoding.
Hao Peng, Ankur Parikh, Manaal Faruqui, Bhuwan Dhingra, Dipanjan Das.
17th Annual Conference of the North American Chapter of the ACL (NAACL, 2019).

- [23] GLoMo: Unsupervisedly Learned Relational Graphs as Transferable Representations. Zhilin Yang*, Jake Zhao*, Bhuwan Dhingra, Kaiming He, William W. Cohen, Ruslan Salakhutdinov, Yann LeCun. 32nd Conference on Neural Information Processing Systems (NeurIPS, 2018).
- [24] Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text. Haitian Sun*, Bhuwan Dhingra*, Manzil Zaheer, Kathryn Mazaitis, Ruslan Salakhutdinov, William W. Cohen. Conference on Empirical Methods in Natural Language Processing (EMNLP, 2018).
- [25] Neural Models for Reasoning over Multiple Mentions using Coreference. Bhuwan Dhingra, Qiao Jin, Zhilin Yang, William Cohen, Ruslan Salakhutdinov. 16th Annual Conference of the North American Chapter of the ACL (NAACL, 2018).
- [26] Simple and Effective Semi-Supervised Question Answering. Bhuwan Dhingra*, Danish Pruthi*, Dheeraj Rajagopal*. 16th Annual Conference of the North American Chapter of the ACL (NAACL, 2018).
- [27] Towards End-to-end reinforcement learning of dialogue agents for information access. Bhuwan Dhingra, Lihong Li, Xiujun Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed, Li Deng. 55th annual meeting of Association of Computational Linguistics (ACL, 2017).
- [28] Gated Attention Readers for Text Comprehension. Bhuwan Dhingra*, Hanxiao Liu*, Zhilin Yang, William W. Cohen, Ruslan Salakhutdinov. 55th annual meeting of Association of Computational Linguistics (ACL, 2017).
- [29] Words or Characters? Fine-grained Gating for Reading Comprehension. Zhilin Yang, Bhuwan Dhingra, Ye Yuan, Junjie Hu, William W. Cohen, Ruslan Salakhutdinov. International Conference on Learning Representations (ICLR, 2017).
- [30] Bootstrapping Distantly Supervised IE using Joint Learning and Small Well-structured Corpora. Lidong Bing, Bhuwan Dhingra, Kathryn Mazaitis, Jong Hyuk Park, William W. Cohen. Thirty-First AAAI Conference on Artificial Intelligence (AAAI, 2017).
- [31] Tweet2Vec: Character-Based Distributed Representations for Social Media. Bhuwan Dhingra, Zhong Zhou, Dylan Fitzpatrick, Michael Muehl, William W. Cohen. 54th annual meeting of Association of Computational Linguistics (ACL, 2016).
- [32] Stock market prediction using hidden markov model. Aditya Gupta, Bhuwan Dhingra. IEEE Student's Conference on Engineering & Systems (SCES, 2012).

Refereed Workshops

- [1] Hierarchical Multi-Label Classification of Online Vaccine Concerns. Chloe Qinyu Zhu, Rickard Stureborg, Bhuwan Dhingra. 8th International Workshop On Health Intelligence (W3PHIAI-24), AAAI (2024).
- [2] Retrieval of synthesis parameters of polymer nanocomposites using LLMs. Defne Circi, Ghazal Khalighinejad, Shruti Badhwar, Bhuwan Dhingra, L Brinson. AI for Accelerated Materials Design (AI4Mat), NeurIPS (2023).
- [3] Sequence Reducible Holdout Loss for Language Model Pretraining. Raghuvveer Thirukovalluru, Bhuwan Dhingra, Sam Wiseman. 4th Workshop on Simple and Efficient NLP (SustaiNLP), ACL (2023).

- [4] Hierarchical Multi-Instance Multi-Label Learning for Detecting Propaganda Techniques.
Anni Chen, Bhuwan Dhingra.
8th Workshop on Representation Learning for NLP, ACL (2023).
- [5] Characterizing the Efficiency vs. Accuracy Trade-off for Long-Context NLP Models.
Phyllis Ang, Bhuwan Dhingra, and Lisa Wu Wills.
NLP Power! The First Workshop on Efficient Benchmarking in NLP, ACL (2022).
- [6] Investigating the Effect of Background Knowledge on Natural Questions.
Vidhisha Balachandran, Bhuwan Dhingra, Haitian Sun, Michael Collins, William W. Cohen.
The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures, (DeeLIO, 2021).
- [7] Probing Biomedical Embeddings from Language Models.
Qiao Jin, Bhuwan Dhingra, William W. Cohen, Xinghua Lu.
3rd Workshop for Evaluating Vector Space Representations for NLP, NAACL (2019).
- [8] AttentionMeSH: Simple, Effective and Interpretable Automatic MeSH Indexer.
Qiao Jin*, Bhuwan Dhingra*, William W. Cohen, Xinghua Lu.
6th BioASQ Workshop, EMNLP (2018).
- [9] Embedding Text in Hyperbolic Spaces.
Bhuwan Dhingra, Christopher Shallue, Mohammad Norouzi, Andrew Dai, George Dahl.
12th Workshop on Graph-Based Natural Language Processing, NAACL (2018).
- [10] Answering Cloze-style Software Questions Using Stack Overflow.
Ezra Winston, Bhuwan Dhingra, Kathryn Mazaitis, Graham Neubig, William Cohen.
Workshop on Machine Learning for Programming, FLoC (2018).
- [11] Using Graphs of Classifiers to Impose Constraints on Semi-supervised Relation Extraction.
Lidong Bing, William W. Cohen, Bhuwan Dhingra, Richard C. Wang.
Workshop on Automatic Knowledge Base Completion, NAACL (2016).

Patents

- [1] Smartphone motion classifier.
Vitor Carvalho, Bhuwan Dhingra, Edward Harrison Teague,
Siddika Parlak Polatkan, Shankar Sadasivam, and Carlos Manuel Puig.
U.S. Patent Application No. 14/865,665, 2021.
- [2] Detecting that a mobile device is riding with a vehicle.
Leonard Henry Gropok and Bhuwan Dhingra.
U.S. Patent Application No. 13/619,143, 2020.
- [3] End-to-end learning of dialogue agents for information access.
Lihong Li, Bhuwan Dhingra, Jianfeng Gao, Xiujun Li, Yun-Nung Chen, Li Deng, Faisal Ahmed.
U.S. Patent Application No. 15/406,425, 2018.
- [4] Context aware system with multiple power consumption modes.
Tadeusz Jarosinski, Shankar Sadasivam, Ryan Carey, Jinwon Lee, Bhuwan Dhingra,
Abhijeet Bisain, Vitor Carvalho, Rajeev Jain, Muralidhar Akula, and Ashwin Swaminathan.
U.S. Patent 9,622,177, issued April 11, 2017.

Invited Talks & Posters

- *Learning the Legibility of Visual Text Perturbations.*
Google Robustness Group, April 2023.
- *Visually & Temporally Robust Language Models.*
UNC Greensboro, November 2022.
- *Learning with Structured & Unstructured Knowledge.*
Proctor & Gamble, January 2022.
- *Language Models as Structured KBs.*
UNC Chapel Hill, October 2021.
- *Text as a Virtual Knowledge Base (talk).*
UC Santa Barbara, February 2021.
Stanford NLP Seminar, October 2019.
UC Irvine AI/ML Seminar, October 2019.
IIT Delhi, November 2019.
UMass Amherst ML & Friends Lunch, December 2019.
USC AI Rising Stars Symposium, December 2019.
- *Differentiable Reasoning over a Virtual Knowledge Base (poster).*
Microsoft AI Breakthroughs Workshop, September 2019.
- *Learning Neural Knowledge Representations (talk).*
Siemens Research, Princeton, April 2019.
- *Reading Graphs of Facts and Text for Question Answering (talk).*
Intuit Research, San Diego, July 2018.
- *Improving the Statistical Efficiency of Machine Reading Comprehension (talk).*
ML Lunch, Carnegie Mellon University, April 2018.
- *Neural Architectures for Reading and Reasoning over Documents (talk).*
AI Seminar sponsored by Apple, Carnegie Mellon University, September 2017.
- *Gated Attention Readers for Text Comprehension (poster).*
IBM Cognitive Colloquium, September 2016.

Other Awards

2019	Siemens FutureMakers PhD Fellowship
2019	CMU 3-Minute Thesis Championship Winner
2012	Best Student Paper Award, IEEE SCES
2011	Todai-IIT Undergraduate Student Scholarship

Service

- Program Chair, 1st South NLP Symposium (2024)
- Faculty Search Committee, Computer Science Department, Duke University (2023)
- DARPA AI Forward Workshop (2023)

- NSF Proposal Review Panel (2023)
- Duke CS+ Faculty Lead (2023)
Using LLMs to generate vaccine interventions
- Workshop Organizer
2022: 1st Workshop on Ever Evolving NLP (EMNLP, 2022)
2021: 2nd Workshop on Unstructured & Structured KBs, AKBC 2021
- PhD Admissions Committee (2022), Duke Computer Science
- Duke Code+ Project Stakeholder (2022)
Using ML/NLP to Identify Relationships and Similarities in Grant Proposal Texts
- Duke CS+ Faculty Lead (2021)
Identifying Vaccine Misinformation in Text
- Senior Area Chair
2023: ACL
- Area Chair
2022: NeurIPS
2021: ICLR, ACL-IJCNLP (Interpretability and Analysis of Models)
2020: NAACL (Information Extraction)
- Reviewing
2023: TACL
2022: TACL, ACL Rolling Review (March, Apr, Sept)
2021: JMLR, TACL, ACL Rolling Review (Oct, Nov)
2020: ICML, EMNLP, ACL, NAACL, AAAI, ICLR, JMLR
2019: NAACL (Outstanding Reviewer), ICML, AKBC, ACL, EMNLP.
2018: NeurIPS (Top 30% Reviewer), ACL, EMNLP, SIGIR, CoLing, NLP-CC English.
2017: IJCNLP.

Advising

- PhD advisees: Rickard Stureborg (2025), Raghuvveer Thirukovalluru (2026), Ghazal Khalighinejad (2026), Sanxing Chen (2027), Junlin Wang (2027), Ruoyu Xie (2028), Yukun Huang (2028)
- PhD thesis committees: Entropy Xu (2024), Zonghao Huang (2025)
- PhD Research Initiation Project (RIP) committees: Karthikeyan K (2023), Suhyeon Lee (2022), Entropy Xu (2022), Zonghao Huang (2022)
- MS Project / Thesis committees: Rongze Gui (2023), Xiaoyin Chen (2023), Greg Szumel (2023), Phyllis Ang (2022), Zihao Lin (2022)
- Undergraduate thesis advisees: Angikar Ghosal (2023), Dev Seth (2022), Anni Chen (2022)
- Undergraduate independent study advisees: Angikar Ghosal (Fall 2023), Frankie Willard (Fall 2023), Dev Seth (Fall, 2022), Aakash Kothapally (Spring 2022), Anni Chen (Spring 2022)

References

Available on request.