

Ritam Dutt

+1-412-450-2916 ◊ rdutt@cs.cmu.edu ◊ <https://shorit.github.io>

Research Interests

Generalization, Evaluation, Personalization, Agentic AI, Large Language Models, NLP, Graph Representation Learning, Discourse and Pragmatics, Information Extraction, Social Networks, Question Answering

Education

Carnegie Mellon University

2019 - 2026

Ph.D. in Language and Information Technologies, advised by [Prof. Carolyn Rosé](#)

GPA: **4.01**

Indian Institute of Technology, Kharagpur

2014 - 2019

Bachelor's and Master's degrees in Computer Science & Engineering

CGPA: **9.71/10.0**

Research Experience

Senior Applied Scientist, Microsoft Outlook Science

Apr 2026 - Present

- Designing self-adaptive evaluation frameworks for LLM agents across several tasks and scenarios. We explore diverse ways of incorporating user feedback for iterative rubric refinement and grader calibration in an automated fashion.
- Investigating personalized knowledge acquisition and representation for AI agents, including how heterogeneous user information can be structured and incorporated via knowledge graphs to improve downstream task performance.
- Developing responsible AI systems via (i) red-teaming to identify and mitigate vulnerabilities to external attacks and (ii) stress-testing against semantically equivalent queries with variations in linguistic expression, dialect, and style.

Applied Science Intern, AWS AI

Jun 2022 - Sep 2022

- Proposed the idea of graph isomorphisms to characterize the complexity of questions in KBQA (Knowledge Base Question Answering) datasets and observed that current KBQA datasets are biased towards simple questions.
- Leveraged graph isomorphisms to create GrailQA++, a benchmark with a balanced distribution of simple and complex questions for KBQA. Observed a significant drop in the generalization performance of SOTA KBQA models on GrailQA++.
- Investigated non-generalizability in KBQA systems along the linguistic dimensions of explicitness and coherence.

Applied Science Intern, AWS AI

Jun 2021 - Aug 2021

- Proposed a setting for question answering over knowledge graphs to address issues of privacy, scalability, and generalizability.
- Performed inductive reasoning by learning the node embeddings of individual knowledge graphs (KGs). Incorporating this structural information improved the generalization performance on academic and internal datasets by 6.5% and 10.5%.

Research Intern, University of Southern California (USC)

May 2018 - Jul 2018

- Worked with Prof. Emilio Ferrara at the Information Sciences Institute, USC, through the IUSSTF-Viterbi Program, 2018.
- Performed linguistic and temporal analysis of Facebook ads allegedly purchased by Russia. We studied the effectiveness of the ads across different dimensions and inferred that sowing division was the primary motive of their dissemination.

Workshop Participant, Artificial Social Intelligence at Microsoft Research, India

Jun 2017 - Jul 2017

- Runner-up at Artificial Social Intelligence [ASI](#), a competitive summer workshop held at Microsoft Research India. Our project was 'Utilizing Social Media for coordinating post-disaster relief operations'.
- Developed unsupervised NLP techniques to extract actionable needs and availabilities from tweets posted during natural disasters, and subsequently match the corresponding needs with availabilities.

Publications

- Z. Wu, **R. Dutt**, L. Breitfeller, A. Nourbakhsh, S. Parekh, C. Rosé, "R²-CoD: Understanding Text-Graph Complementarity in Relational Reasoning via Knowledge Co-Distillation", **AACL-IJCNLP**, 2025. [\[Paper\]](#)[\[Outstanding Paper\]](#)
- **R. Dutt**, C. Rosé, M. Sap, "SOCIAL SCAFFOLDS: A Generalization Framework for Social Understanding Tasks", **EMNLP**, 2025. [\[Paper\]](#)
- The Big-Science Evaluation Team, "SHADES: Towards a Multilingual Assessment of Stereotypes in Large Language Models", **NAACL**, 2025. [\[Paper\]](#)
- **R. Dutt**, Z. Wen, K. Shi, D. Sheth, P. Gupta, C. Rosé, "Leveraging Machine-Generated Rationales to Facilitate Social Meaning Detection in Conversations", **ACL**, 2024. [\[Paper\]](#)
- L. Weissweiler, V. Hofmann, A. Kantharuban, A. Cai, **R. Dutt**, A. Hengle, A. Kabra, A. Kulkarni, A. Vijayakumar, H. Yu, H. Schuetze, K. Oflazer, & D. R. Mortensen, "Counting the Bugs in ChatGPT's Wugs: A Multilingual Investigation into the Morphological Capabilities of a Large Language Model", **EMNLP**, 2023. [\[Paper\]](#)

- **R. Dutt**, S. Khosla, V. B. Kumar & R. Gangadharaiah, “GrailQA++: A Challenging Zero-Shot Benchmark for Knowledge Base Question Answering”, **AAACL-IJCNLP**, 2023. [\[Paper\]](#)
- S. Gururaja, **R. Dutt**, T. Liao, & C. Rosé, “Linguistic representations for fewer-shot relation extraction across domains”, **ACL**, 2023. [\[Paper\]](#)
- **R. Dutt**, K. Bhattacharjee, R. Gangadharaiah, D. Roth, & C. Rosé, “PERKGQA: Question Answering over Personalized Knowledge Graphs”, Findings of **NAACL**, 2022. [\[Paper\]](#)
- S. Vashishth, D. Newman-Griffis, R. Joshi, **R. Dutt**, & C. Rosé, “Improving broad-coverage medical entity linking with semantic type prediction and large-scale datasets”, **Journal of Biomedical Informatics**, 2021. [\[Paper\]](#)
- M. Das, P. Saha, **R. Dutt**, P. Goyal, A. Mukherjee, & B. Mathew, ““You too Brutus!” Trapping Hateful Users in Social Media: Challenges, Solutions & Insights”, **HyperText**, 2021. [\[Paper\]](#) [\[Ted Nelson Newcomer Award\]](#)
- A. Ghosh, **R. Dutt**, & C. Wilson, “When fair ranking meets uncertain inference”, **SIGIR**, 2021. [\[Paper\]](#)
- **R. Dutt**, G. Hillaire, A. Fang, L. Larke, C. Rosé, & J. Reich, “Investigating Adoption and Collaboration with Digital Clinical Simulations by Teacher Educators”, **SITE**, 2021. [\[Paper\]](#) [\[Outstanding Paper Award\]](#)
- **R. Dutt***, S. Sinha*, R. Joshi, S. S. Chakraborty, M. Riggs, X. Yan, H. Bao & C. Rosé, “ResPer: Computationally Modelling Resisting Strategies in Persuasive Conversations”, **EACL**, 2021. [\[Paper\]](#)
- **R. Dutt**, R. Joshi, & C. Rosé, “Keeping up Appearances: Computational Modeling of Face Acts in Persuasion Oriented Discussions”, **EMNLP**, 2020. [\[Paper\]](#)
- R. Bal, S. Sinha, S. Dutta, R. Joshi, S. Ghosh, & **R. Dutt**, “Analysing the Extent of Misinformation in Cancer Related Tweets”, **ICWSM**, 2020. [\[Paper\]](#)
- B. Mathew*, **R. Dutt***, S. Maity, P. Goyal, & A. Mukherjee, “Deep Dive into Anonymity: Large Scale Analysis of Quora Questions.” **SocInfo**, 2019. [\[Paper\]](#) [\[Best Paper Nomination\]](#)
- **R. Dutt***, M. Basu*, K. Ghosh, & S. Ghosh, “Utilizing microblogs for assisting post-disaster relief operations via matching resource needs and availabilities”, **Journal of Information Processing & Management**, 2019. [\[Paper\]](#)
- B. Mathew, **R. Dutt**, P. Goyal, & A. Mukherjee, “Spread of hate speech in online social media.” **WebSci**, 2019. [\[Paper\]](#) [\[Best Paper Honorable Mention\]](#)
- A. Mullick, S. Ghosh*, **R. Dutt***, A. Ghosh*, & A. Chakraborty, “Public Sphere 2.0: Targeted Commenting in Online News Media”, **ECIR**, 2019. [\[Paper\]](#) [\[Best Poster Award\]](#)
- **R. Dutt**, A. Deb, & E. Ferrara, ““Senator, We Sell Ads”: Analysis of the 2016 Russian Facebook Ads Campaign”, **ICIIT**, 2018. [\[Paper\]](#)
- M. Singh, P. Dogga*, S. Patro*, D. Barnwal*, **R. Dutt***, R. Haldar, P. Goyal, & A. Mukherjee, “CL Scholar: The ACL Anthology Knowledge Graph Miner.” **NAACL Demonstrations**, 2018. [\[Paper\]](#)

Skills and Expertise

- **Ph.D. Courses:** Question Answering, Multilingual NLP, Quantitative Evaluation, Neural Networks for NLP, Computational Modelling of Discourse, Language and Statistics, Dialogue and Discourse Processing, Intermediate Statistics
- **Graduate-Level Courses:** Social Computing, Information Retrieval, Natural Language Processing, Deep Learning
- **Languages:** Python, R, C, C++, Java, SQL
- **Primary Libraries:** Hugging Face, NumPy, scikit-learn, DeepSpeed, PyTorch, Keras, TensorFlow, igraph, pandas
- **Tools and Systems:** Jupyter, GitHub, RStudio, LaTeX, Linux, Bash, Slurm

Workshop Publications

- A. Kierans, **R. Dutt**, K. Rittichier, S. Dori-Hacohen, A. Ghosh, “Position: Evaluations of AI Moral Reasoning Still Miss Half of the Picture.” **EvalEval@ACL**, 2026. [\[Paper\]](#)
- **R. Dutt**, S. Sural, & C. Rosé, “Can dependency parses facilitate generalization in language models? A case study of cross-lingual relation extraction.” **KnowledgeNLP@NAACL**, 2025. [\[Paper\]](#)
- **R. Dutt***, S. Choudhury*, V. Rao, C. Rosé, & V. Vydiswaran. “Investigating the Generalizability of Pretrained Language Models across Multiple Dimensions: A Case Study of NLI and MRC.” **GenBench@EMNLP**, 2024. [\[Paper\]](#) [\[Outstanding Paper\]](#)
- Z. Yu, **R. Dutt**, & C. Rosé, “Evaluating Large Language Models on Social Signal Sensitivity: An Appraisal Theory Approach.”, **HuCLLM@ACL**, 2024. [\[Paper\]](#)
- S. Khosla*, **R. Dutt***, V. B. Kumar, & R. Gangadharaiah, “Exploring the Reasons for Non-generalizability of KBQA systems.”, **Insights from Negative Results in NLP@EACL**, 2023. [\[Paper\]](#)

- S. Bansal, S. Tripathi, S. Agarwal, S. Gururaja, A. Veerubhotla, **R. Dutt**, T. Mitamura, & E. Nyberg, “R3: Refined Retriever-Reader Pipeline for Multidoc2dial”, **Doc2Dial@ACL**, 2022. [\[Paper\]](#) [\[Best Performance\]](#)
- S. Khosla, J. Lovelace, **R. Dutt**, & A. Pratapa, “Team jars: Dialdoc subtask 1-improved knowledge identification with supervised out-of-domain pretraining”, **Doc2Dial@ACL**, 2021. [\[Paper\]](#)
- **R. Dutt***, S. Khosla*, & C. Rosé, “A pipelined approach to Anaphora Resolution in Chemical Patents.” **CLEF**, 2021. [\[Paper\]](#) [\[Best Performance\]](#)
- S. Khosla*, R. Joshi*, **R. Dutt***, A. W. Black, & Y. Tsvetkov, “LTIatCMU at SemEval-2020 Task 11: Incorporating multi-level features for multi-granular propaganda span identification.” **SemEval**, 2020. [\[Paper\]](#)
- K. Hiware*, **R. Dutt***, S. Sinha, S. Patro, K. Ghosh, & S. Ghosh, “NARMADA: Need and Available Resource Managing Assistant for Disasters and Adversities.” **SocialNLP@ACL**, 2020 [\[Paper\]](#)
- **R. Dutt**, K. Hiware, A. Ghosh, & R. Bhaskaran, “Savitr: A system for real-time location extraction from microblogs during emergencies.” **SMERP@WWW**, 2018. [\[Paper\]](#)

Preprints

- **R. Dutt**, Z. Wu, D. Ling, D. Gupta, C. Rosé, “IsoKBQA: Graph Isomorphisms as Structural Scaffolds for Question Answering over Knowledge Bases”. [\[Paper\]](#)
- F. Iqbal, **R. Dutt**, S. Das, A. Verma, S. Choudhury, “Evaluating Multi-Dimensional Generalization of Large Language Models in Temporal Extraction Tasks.” [\[Paper\]](#)
- A. Kierans, K. Rittichier, **R. Dutt**, S. Dori-Hacohen, A. Ghosh, “Position: AI Should Align with Reasoning Norms.” [\[Paper\]](#)
- N. Suh, **R. Dutt**, A. Williams, A. Kuzminykh, “Towards Proactive Personalized AI Systems” [\[Paper\]](#)
- **R. Dutt**, Z. Wu, C. Rosé, “POLIRAT: Investigating Politeness Principles in LLMs using Rationales.” [\[Paper\]](#)

Notable Achievements

Academic Honors

- Outstanding Paper Award at AACL-IJCNLP, 2025.
- Outstanding Paper at GenBench, 2nd GenBench Workshop, at EMNLP 2024.
- Achieved the highest score in the Shared Task at DialDoc, 2nd DialDoc Workshop, at ACL 2022.
- Achieved the highest score in the Shared Task at CLEF, Conference and Labs of the Evaluation Forum, 2021.
- Awarded the Ted Nelson Newcomer Award at Hypertext, 32nd ACM Conference on Hypertext and Social Media, 2021.
- Outstanding Paper Award at SITE, Society for Information Technology & Teacher Education, 2021.
- Best Paper Award (Honorable Mention) at WebSci, 10th ACM Conference on Web Science, 2019.
- Best Poster Award at the 41st European Conference on Information Retrieval (ECIR), 2019.
- Nomination for Best Paper at SocInfo, Social Informatics, 2019.
- One of the 5 recipients of the SIGWEB student travel grant for attending WebSci, 2019.
- Batch Topper: Joint Recipient of the Institute Silver Medal in Computer Science and Engineering at IIT Kharagpur.
- Batch Topper: Secured distinction with 97.5% in ISC, (High School or Class 12).

Professional Service

- Organizer of the 2nd and 3rd Workshops on Social Influence in Conversations (SiCon), at ACL 2025 and EMNLP 2024.
- Reviewer at COLM (2026), NeurIPS (2026), EMNLP (2023, 2024, 2026), ACL (2021, 2022, 2023, 2024, 2025), NAACL (2022, 2024), EACL (2021, 2023), WebSci (2020), WWW (2020), AAI (2020) and affiliated workshops.
- Student volunteer at ICWSM (2020) and WebSci (2019).
- Graduate Teaching Assistant for Ethics, Social Biases, and Positive Impact in Language Technologies at Carnegie Mellon University in Spring, 2024.
- Graduate Teaching Assistant for Applied Machine Learning at Carnegie Mellon University in Fall, 2022.
- Graduate Teaching Assistant for Applied Machine Learning at Carnegie Mellon University in Fall, 2021.
- Graduate Teaching Assistant for Neural Networks for NLP at Carnegie Mellon University in Spring, 2021.
- Graduate Teaching Assistant for the Deep Learning course at IIT Kharagpur in Spring, 2019.
- Graduate Teaching Assistant for the Speech and Natural Language Processing course at IIT Kharagpur in Autumn, 2018.