

Hannah Rashkin

Research Scientist, Google Deepmind
Email: hrashkin <at> google.com
Website: hrashkin.github.io

WORK EXPERIENCE

Current:

Research Scientist, Google Deepmind, 2020 - Present

Prior:

Graduate research and teaching assistant, UW, 2014-Spring 2020

Microsoft Research Intern, MSR, Summer 2019

AI2 Research Intern, Allen Institute for Artificial Intelligence, Fall 2018-Spring 2019

Facebook AI Research Intern, FAIR, Summer 2018

National Security Internship Program, Pacific Northwest National Laboratory, . . Summer 2016

EDUCATION

University of Washington, Fall 2014 - Spring 2020

PhD, Paul G. Allen School of Computer Science & Engineering.

Advisor: Yejin Choi

MS, Paul G. Allen School of Computer Science & Engineering.

Received Spring 2016

University of Illinois at Urbana-Champaign, Fall 2010 - Spring 2014

BS with Highest Honors in Computer Science.

Minor in International Engineering in Japanese Studies.

RESEARCH INTERESTS

- Evaluating faithfulness and attribution in knowledge-grounded LLM output
- Designing more controllable and interactive NLP systems
- Applications of NLP technologies for positive social impact
- Incorporating social dynamics and reasoning for improving story generation, narrative understanding, empathetic response generation

PUBLICATIONS

- *Investigating Content Planning for Navigating Trade-offs in Knowledge-Grounded Dialogue*. Kushal Chawla, Hannah Rashkin, Gaurav Singh Tomar, David Reitter. EACL 2024.
- *Measuring Attribution in Natural Language Generation Models*. Hannah Rashkin*, Vitaly Nikolaev*, Matthew Lamm, Lora Aroyo, Michael Collins, Dipanjan Das, Slav Petrov, Gaurav Singh Tomar, Iulia Turc and David Reitter. Computational Linguistics, 2023.
- *Evaluating Attribution in Dialogue Systems: The BEGIN Benchmark*. Nouha Dziri*, Hannah Rashkin*, Tal Linzen, and David Reitter. TACL 2022.
- *CONQRR: Conversational Query Rewriting for Retrieval with Reinforcement Learning*. Zeqiu Wu, Yi Luan, Hannah Rashkin, David Reitter, Hannaneh Hajishirzi, Mari Ostendorf, Gaurav Singh Tomar. EMNLP 2022
- *Increasing Faithfulness in Knowledge-Grounded Dialogue with Controllable Features*. Hannah Rashkin, David Reitter, Gaurav Singh Tomar, and Dipanjan Das. In Proceedings of ACL, 2021.

- *PowerTransformer: Unsupervised Controllable Revision for Biased Language Correction*. Xinyao Ma*, Maarten Sap*, Hannah Rashkin, and Yejin Choi. In Proceedings of EMNLP, 2020.
- *PlotMachines: Outline-Conditioned Generation with Dynamic Plot State Tracking*. Hannah Rashkin, Asli Celikyilmaz, Yejin Choi, and Jianfeng Gao. In Proceedings of EMNLP, 2020.
- *Abductive Commonsense Reasoning*. Chandra Bhagavatula, Ronan Le Bras, Chaitanya Malaviya, Keisuke Sakaguchi, Ari Holtzman, Hannah Rashkin, Doug Downey, Scott Wen-tau Yih, and Yejin Choi. In Proceedings of ICLR, 2020.
- *Social IQA: Commonsense Reasoning about Social Interactions*. Maarten Sap*, Hannah Rashkin*, Derek Chen, Ronan LeBras, and Yejin Choi. In Proceedings of EMNLP, 2019.
- *Defending Against Neural Fake News*. Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. In Proceedings of NeurIPS, 2019.
- *Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset* Hannah Rashkin, Eric Michael Smith, Margaret Li, Y-Lan Boureau. In Proceedings of ACL, 2019.
- *COMET: Commonsense Transformers for Automatic Knowledge Graph Construction*. Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. In Proceedings of ACL, 2019.
- *ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning*. Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A Smith, and Yejin Choi. In Proceedings of AACL, 2019.
- *Modeling Naive Psychology of Characters in Simple Commonsense Stories*. Hannah Rashkin, Antoine Bosselut, Maarten Sap, Kevin Knight and Yejin Choi. In Proceedings of ACL, 2018.
- *Event2Mind: Commonsense Inference on Events, Intents, and Reactions*. Hannah Rashkin*, Maarten Sap*, Emily Allaway, Noah A. Smith, and Yejin Choi. In Proceedings of ACL, 2018.
- *Connotation Frames of Agency and Power in Modern Films*. Maarten Sap, Marcella Cindy Prasettio, Ari Holtzman, Hannah Rashkin, and Yejin Choi. In Proceedings of EMNLP - Short, 2017.
- *Truth of Varying Shades: On Political Fact-Checking and Fake News*. Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova and Yejin Choi. In Proceedings of EMNLP - Short, 2017.
- *Multilingual Connotation Frames: A Case Study on Social Media for Targeted Sentiment Analysis and Forecast*. Hannah Rashkin, Eric Bell, Yejin Choi, and Svitlana Volkova. In Proceedings of ACL - Short, 2017.
- *Connotation Frames: A Data Driven Investigation*. Hannah Rashkin, Sameer Singh, and Yejin Choi. In Proceedings of ACL, 2016.
- *Document-Level Sentiment Inference with Social, Factual, and Discourse Context*. Eunsol Choi, Hannah Rashkin, Luke Zettlemoyer and Yejin Choi. In Proceedings of ACL, 2016.

SERVICE

Organizing Committee	NeuralGen Workshop (2019)
Program Committee/Reviewer	ACL, EMNLP, NAACL, TACL, ARR, SIGDial
Area Chair	NAACL 21, Coling 22, EMNLP 21,22,24, ACL 23-24, COLM 24

TEACHING EXPERIENCE

Teaching Assistant, Winter 2019
University of Washington
Course title: Natural Language Processing (Graduate Level)
Held weekly office hours, helped students with their projects, and graded assignments and projects.

Teaching Assistant, Spring 2017
University of Washington
Course title: Natural Language Processing Capstone.
Held weekly meetings with undergraduate students to help them design and implement NLP research projects.

Undergraduate Teaching Assistant and Lab Instructor, Fall 2012-Spring 2014
University of Illinois at Urbana-Champaign
Course title: Computer Architecture.
Responsible for leading weekly lab sections, held regular office hours.

ACADEMIC AWARDS

NSF Graduate Research Fellowship 2015-2020
Bronze Tablet, University of Illinois Spring 2014
James Scholar, University of Illinois 2010-2014
Engineering International Scholarship, University of Illinois 2010-2014