

EMILY ALLAWAY

eallaway@ed.ac.uk

<https://emilyallaway.github.io/>

EMPLOYMENT EXPERIENCE

University of Edinburgh, School of Informatics Edinburgh, Scotland, UK. Chancellor's Fellow.	08/2024 - Present
Columbia University, Department of Computer Science New York, NY, USA. Research and Teaching Assistant with Kathleen McKeown.	09/2018 – 07/2024
Allen Institute for Artificial Intelligence Seattle, WA, USA. Research Intern with MOSAIC team.	05/2021 – 12/2021
Amazon Science New York, NY, USA. Research Intern with Miguel Ballesteros.	05/2020 – 10/2020
University of Washington, Department of Computer Science Seattle, WA, USA. Research Assistant with Yejin Choi.	01/2017 – 06/2018

EDUCATION

Columbia University , New York, NY Ph.D. in Computer Science – Natural Language Processing Advisor: Kathleen McKeown Thesis: Understanding and Reasoning about Implicit Meaning in Language	2024
M.S. in Computer Science	2021
University of Washington , Seattle, WA B.S. in Computer Science	2018
B.S. in Mathematics	2018

ACADEMIC AWARDS

NSF Graduate Research Fellowship National Science Foundation	09/2019
Denice Dee Denton Scholars Endowment Paul G. Allen School of Computer Science and Engineering, University of Washington	09/2016
Jim Greenfield Scholarship Department of Classics, University of Washington	09/2016
Harvey Densmore Scholarship Department of Classics, University of Washington	09/2016

TALKS

Investigating Generics in Language Models 2nd Workshop on Meaning@Edinburgh	05/2025
Exploring the Challenges of Defeasible Reasoning in Large Language Models: A Generics-Based Perspective University of Aberdeen CS Seminar	04/2025

Substantiating Claims in Scientific Documents

Huawei Joint Research Lab - Spring Workshop 04/2025

Huawei Joint Research Lab - Fall Workshop 10/2024

Exceptions, Instantiations, and Overgeneralization: Insights into How Language Models Process Generics

EMNLP Main Conference 11/2024

Understanding and Reasoning about Generalizations in Language

University of Edinburgh NLP Meeting 05/2023

Columbia University NLP Seminar 05/2023

Penguins Don't Fly:

Reasoning about Generics through Instantiations and Exceptions 05/2023

EACL Main Conference

Implicit Meaning and Domain Adaptation in Stance Detection

Columbia University NLP Seminar 03/2021

Stance Detection

Columbia University Department of Computer Science Alumni Day 02/2021

Lexical Connotations & Stance Detection

Demystifying the dissertation talk series, Columbia University Department of Computer Science 08/2020

Word Connotations & NLP

Emerging Scholars Program at Columbia University 11/2019

TEACHING

University of Edinburgh

Natural Language Understanding, Generation, and Machine Translation (NLU-11) – Guest Lecture

Fact-checking in NLP 03/2025

Accelerated Natural Language Processing (ANLP) – Guest Lecture

Analysing Implied Meaning in Text 11/2024

Columbia University

Language Generation and Summarization (COMS 6998) – Guest Lecture

Unsupervised Constrained Decoding 09/2022, 09/2023

Natural Language Processing (COMS W4705) – Guest Lecture

Neural Networks for NLP 09/2021

Natural Language Processing (COMS W4705) – Teaching Assistant Fall 2019, Spring 2021

University of Washington

Computer Programming I/II (CSE 142/143) – Teaching Assistant 2015 – 2017

SUPERVISION

Current MSc Students

University of Edinburgh

Hanna Yakusevych 2025

Xinwei Niu 2025

Jinghong Li 2025

Ignacio Porras Vargas 2025

MSc Students

Columbia University

Lauren Hong 2022

Anish Saha 2021/2022

Lucy Wang 2021/2022

Sahil Jayaram 2020/2021

Amith Ananthram	2020
Malavika Srikanth	2020

Undergraduate Students

Columbia University

Nina Taneja	2022
Noah Bergam	2022
António Câmara	2021
William Chiu	2021
Yuncong Yang	2020

SERVICE

Chairs

WiNLP Workshop Chair	2025
IJCNLP-AAACL Senior Area Chair	2025
ARR Area Chair	2025
Publication Chair - COLING	2025

Reviewing

Conferences

EACL	2023
ACL	2021, 2023
EMNLP	2021, 2022, 2023
LREC	2022
LREC-COLING	2024
ARR	2021, 2022, 2023

Journals

Journal of Natural Language Engineering	2022
Language Resources and Evaluation	2023

Workshops

NAACL Workshop on Deep Learning for Low-resource NLP	2022
EMNLP Workshop on Noisy User-generated Text	2021
COLING Workshop on Noisy User-generated Text	2022
EACL Workshop on Noisy User-generated Text	2024
ACL Workshop on Commonsense Representation and Reasoning	2022

Seminars & Reading Group Organization

University of Edinburgh ILCC/CDT NLP Seminar, Co-Organizer	10/2024 -
Columbia University NLP Seminar, Organizer	01/2020 - 08/2023
NLP Classics Reading Group at Columbia University, Organizer	Fall 2020

PUBLICATIONS

Peer-reviewed conference articles

(C.19) Gustavo Cilleruelo Calderón, **Emily Allaway**, Barry Haddow, Alexandra Birch (2025). *MGen: Millions of Naturally Occurring Generics in Context*. SCiL.

(C.18) **Emily Allaway** and Kathleen McKeown (2025). *Evaluating Defeasible Reasoning in LLMs with DEFREASING*. NAACL.

(C.17) Gustavo Cilleruelo Calderón, **Emily Allaway**, Barry Haddow, and Alexandra Birch (2025). *Generics are Puzzling. Can language models find the missing piece?* COLING.

- (C.16) Jimin Mun, **Emily Allaway**, Akhila Yerukola, Laura Vianna, Sarah-Jane Leslie, Maarten Sap (2023). *Beyond Denouncing Hate: Strategies for Countering Implied Biases and Stereotypes in Language*. Findings of EMNLP.
- (C.15) **Emily Allaway**, Jena D. Hwang, Chandra Bhagavatula, Kathleen McKeown, Doug Downey, Yejin Choi (2023). *Penguins don't Fly: Reasoning about Generics through Instantiations and Exceptions*. EACL
- (C.14) Anish Saha, Amith Ananthram, **Emily Allaway**, Heng Ji, Kathleen McKeown (2022). *Seeded Hierarchical Clustering for Expert-Crafted Taxonomies*. Findings of EMNLP.
- (C.13) Sharon Levy, **Emily Allaway**, Melanie Subbiah, Lydia Chilton, Desmond Patton, Kathleen McKeown, William Yang Wang (2022). *SafeText: A Benchmark for Exploring Physical Safety in Language Models*. EMNLP.
- (C.12) Alex Mei, Anisha Kabir, Sharon Levy, Melanie Subbiah, **Emily Allaway**, John Judge, Desmond Patton, Bruce Bimber, Kathleen McKeown, William Yang Wang (2022). *Mitigating Covertly Unsafe Text within Natural Language Systems*. Findings of EMNLP.
- (C.11) Alicia Parrish, William Huang, Omar Agha, Soo-Hwan Lee, Nikita Nangia, Alex Warstadt, Karmanya Aggarwal, **Emily Allaway**, Tal Linzen, and Samuel R. Bowman (2021). *Does Putting a Linguist in the Loop Improve NLU Data Collection?*. Findings of EMNLP.
- (C.10) Sahil Jayaram and **Emily Allaway** (2021). *Human Rationales as Attribution Priors for Explainable Stance Detection*. EMNLP.
- (C.9) **Emily Allaway**, Shuai Wang, and Miguel Ballesteros (2021). *Sequential Cross-Document Coreference Resolution*. EMNLP.
- (C.8) **Emily Allaway** and Kathleen McKeown (2021). *A Unified Feature Representation for Lexical Connotations*. EACL.
- (C.7) **Emily Allaway**^{*}, Malavika Srikanth^{*}, Kathleen McKeown (2021). *Adversarial Learning for Zero-Shot Stance Detection on Social Media*. NAACL.
- ^{*} denotes equal contribution
- (C.6) Amith Ananthram, **Emily Allaway**, Kathleen McKeown (2020). *Event-Guided Denoising for Multilingual Relation Learning*. COLING.
- (C.5) **Emily Allaway** and Kathleen McKeown (2020). *Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations*. EMNLP.
- (C.4) Manling Li, Ying Lin, Ananya Subburathinam, Spencer Whitehead, Xiaoman Pan, Di Lu¹, Qingyun Wang, Tongtao Zhang, Lifu Huang, Heng Ji, Alireza Zareian, Hassan Akbari, Brian Chen, Bo Wu, **Emily Allaway**, Shih-Fu Chang, Kathleen McKeown, Yixiang Yao, Jennifer Chen, Eric Berquist, Kexuan Sun, Xujun Peng, Ryan Gabbard, Marjorie Freedman, Pedro Szekely, T.K. Satish Kumar, Arka Sadhu, Ram Nevatia, Miguel Rodriguez, Yifan Wang, Yang Bai, Ali Sadeghian, Daisy Zhe Wang (2019). *GAIA at SM-KBP2019 - A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System*. TAC.
- (C.3) Maarten Sap, Ronan Le Bras, **Emily Allaway**, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi (2019). *ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning*. AAAI.
- (C.2) Tongtao Zhang, Ananya Subburathinam, Ge Shi, Lifu Huang, Di Lu, Xiaoman Pan, Manling Li, Boliang Zhang, Qingyun Wang, Spencer Whitehead, Heng Ji, Alireza Zareian, Hassan Akbari, Brian Chen, Ruiqi Zhong, Steven Shao, **Emily Allaway**, Shih-Fu Chang, Kathleen McKeown, Dongyu Li, Xin Huang, Kexuan Sun, Xujun Peng, Ryan Gabbard, Marjorie Freedman, Mayank Kejriwal, Ram Nevatia, Pedro Szekely, T.K. Satish Kumar, Ali Sadeghian, Giacomo Bergami, Sourav Dutta, Miguel Rodriguez, Daisy Zhe Wang (2018). *GAIA - A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System*. TAC.

(C.1) Hannah Rashkin, Maarten Sap, **Emily Allaway**, Noah A. Smith and Yejin Choi (2018). *Event2Mind: Commonsense Inference on Events, Intents, and Reactions*. ACL.

Peer-reviewed journal articles

(J.2) **Emily Allaway**, Chandra Bhagavatula, Jena D. Hwang, Yejin Choi, Kathleen McKeown, and Sarah-Jane Leslie (to appear, 2024). *Exceptions, Instantiations, and Overgeneralization: Insights into How Language Models Process Generics*. Computational Linguistics: Special Issue on Language Learning, Representation, and Processing in Humans and Machines.

(J.1) **Emily Allaway**, Kathleen McKeown (2023). *Zero-Shot Stance Detection: Paradigms and Challenges*. Frontiers of Artificial Intelligence, vol. 5.

Peer-reviewed workshop articles

(W.3) **Emily Allaway**, Nina Taneja, Sarah-Jane Leslie, Maarten Sap (2022). *Towards Countering Essentialism Through Social Bias Reasoning*. EMNLP Workshop on NLP for Positive Impact.

(W.2) Noah Bergam, **Emily Allaway**, Kathleen McKeown (2022). *Legal and Political Stance Detection of SCOTUS Language*. EMNLP Workshop on Natural Legal Language Processing.

(W.1) António Câmara, Nina Taneja, Tamjeed Azad, **Emily Allaway**, Richard Zemel (2022). *Mapping the Multilingual Margins: Intersectional Biases of Sentiment Analysis Systems in English, Spanish, and Arabic*. ACL Workshop on Language Technology for Equality, Diversity, Inclusion.