revision date: November 9, 2026

João Sedoc

Department of Technology, Operations, and Statistics
New York University
44 West Fourth Street
New York, NY 10012

jsedoc@stern.nyu.edu
Phone: 773-257-3362
Office: Kaufman Management Center 8-85

Research Interests

Natural Language Understanding, Conversational Agents, Deep Learning

Employment

New York University	New York, NY
Associate Professor, Tenure-Track, Department of Technology, Operations, and Statistics (2026 -)	
New York University	New York, NY
Assistant Professor, Department of Technology, Operations, and Statistics	(2020 - 2026)
New York University	New York, NY
Affiliated Faculty, Center for Data Science	(2021 -)
Johns Hopkins University	Baltimore, MD
Adjunct Assistant Professor, Center for Speech and Language Processing	(2020 - 2023)
Johns Hopkins University	Baltimore, MD
Assistant Research Professor, Computer Science Department	(2019 - 2020)
Amazon Core Machine Translation	Pittsburgh
Applied Scientist Intern,	(Summer 2017)
Chivalric Trading LLC / Chambertin LLC	Chicago
Managing Member and Founder,	(2010 - 2015)
Millennium Partners	New York
Portfolio Manager,	(2007 - 2009)
Cooper Neff / BNP Paribas	King of Prussia/ Paris / New York
Global Head of High Frequency Research, Director	(2004 - 2007)
Researcher, Vice President	(2001 - 2004)
Researcher Programmer, Analyst	(1999 - 2001)

Education

2019	PhD	Computer Science	University of Pennsylvania, Philadelphia, PA
1999	BA	Mathematics	University of Pennsylvania, Philadelphia, PA
1999	BSE	Computer Science	University of Pennsylvania, Philadelphia, PA

Publications

Unless specified, all papers are the senior author last
(A–Z): alphabetical order
(EF): effort order
(SA): senior author last
(RJP): resubmitted as a journal paper
(RCP): resubmitted as a conference paper
(NYU–S): NYU PhD student co-author
(*): student under direct supervision
(_): primary contributors

Journal Publications

1. (SA) Peter Novello, Rose Weeks, João Sedoc, Yinzhou Ou, Pooja Sangha, Emily Miller, Amelia Jamison,Smisha Agarwal, "What can chatbot conversations reveal about vaccine concerns? An observational topic modeling study for public health infoveillance", BMJ Digital Health and AI, 2026.
2. Huy Vu, Huy Anh Nguyen, Adithya V Ganesan, Swanie Juhng, Oscar NE Kjell, João Sedoc, Margaret L Kern, Ryan L Boyd, Lyle Ungar, H Andrew Schwartz, Johannes C Eichstaedt, "PsychAdapter: Adapting LLM Transformers to Reflect Traits, Personality and Mental Health", npj Artificial Intelligence, 2026.
3. (SA) Pooja Sangha, Rose Weeks, Emily Miller, João Sedoc, Helen Lacey, Christine Liang, Smisha Agarwal, "Do we message with empowerment, empathy, or collectivity? A social media campaign microtargeting young Americans to encourage engagement with a bilingual COVID-19 chatbot", Journal of Social Marketing, 2025.
4. (SA) Aadesh Salecha, Molly E. Ireland, Shashanka Subrahmanya, João Sedoc, Lyle H. Ungar, Johannes C. Eichstaedt, "Large language models display human-like social desirability biases in Big Five personality surveys" PNAS Nexus, 2024.
   Note: Previous arXiv pre-print title "Large language models show human-like social desirability biases in survey responses"
5. (SA) Salvatore Giorgi, Kelsey Isman, Tingting Liu, Zachary Fried, João Sedoc, Brenda Curtis, "Evaluating generative AI responses to real-world drug-related questions", Psychiatry Research, 2024.
6. (SA) Michelle Hoang Nguyen, João Sedoc, Casey Overby Taylor, "Usability, Engagement, and Report Usefulness of Chatbot-Based Family Health History Data Collection: Mixed Methods Analysis", Journal of Medical Internet Research, 2024.
7. (SA) Elizabeth C. Stade, Shannon Wiltsey Stirman, Lyle H. Ungar, Cody L. Boland, H. Andrew Schwartz, David B. Yaden, João Sedoc, Robert J. DeRubeis, Robb Willer, Johannes C. Eichstaedt, "Large language models could change the future of behavioral healthcare: a proposal for responsible development and evaluation," npj Mental Health Research, 2024.
   Note: previous pre-print title "Artificial Intelligence Will Change the Future of Psychotherapy: A Proposal for Responsible, Psychologist-led Development".
8. (SA) Rose Weeks, Pooja Sangha, Lyra Cooper, João Sedoc, Sydney White, Shai Gretz,

Assaf Toledo, Dan Lahav, Anna-Maria Hartner, Nina M Martin, Jae Hyoung Lee, Noam Slonim, Naor Bar-Zeev, "Usability and Credibility of a COVID-19 Vaccine Chatbot for Young Adults and Health Workers in the United States: Formative Mixed Methods Study," JMIR Human Factors, 2023.

9. (SA) Rose Weeks, Lyra Cooper, Pooja Sangha, João Sedoc, Sydney White, Assaf Toledo, Shai Gretz, Dan Lahav, Nina Martin, Alexandra Michel, Jae Hyoung Lee, Noam Slonim, Naor Bar-Zeev, "Piloting a COVID-19 vaccine chatbot with young adults and health workers in the U.S. to validate usability, credibility, and intention to use.," Journal of Medical Internet Research Human Factors, 2022.
10. (SA) Nina M. Martin, João Sedoc, Lisa Poirier, Andrew J. Rosenblum, Melissa M. Reznar, Joel Gittelsohn, Daniel J. Barnett, "Harnessing Artificial Intelligence to Improve Food Assistance: A Scoping Review of Machine Learning Tools," Public Health Nutrition, 2022.
11. (EF) Smisha Agarwal, Pritika Parmar, Jina Ryu, Shivani Pandya, João Sedoc, "Health-Focused Conversational Agents in Person-Centered Care: A Review of Apps," NPJ Digital Medicine 2022.
12. (SA) Seolhwa Lee, Kisu Yang, Chanjun Park, João Sedoc, Heuiseok Lim, "Who Speaks Like a Style of Vitamin: Towards Syntax-Aware Dialogue Summarization Using Multi-Task Learning," IEEE Access 2021.

    NOTE: earlier version presented at Workshop on Widening NLP at EMNLP, 2021 as "Towards Syntax-Aware Dialogue Summarization using Multi-task Learning" and NeurIPS Affinity Workshop: WiML Workshop as "Syntax-enhanced Dialogue Summarization using Syntax-aware Information"
13. (SA) Sunny Tang, Reno Kriz, Sunghye Cho, Suh Jung Park, Jenna Harowitz, Raquel Gur, Mahendra Bhati, Daniel Wolf, João Sedoc, Mark Liberman, "Natural Language Processing Methods are Sensitive to Sub-Clinical Linguistic Differences in Schizophrenia Spectrum Disorders," NPJ Schizophrenia, 2021.
14. (SA) Michelle R. Kaufman, Alicia T. Bazell, Anne Collaco, João Sedoc, "This show hits really close to home on so many levels': An analysis of Reddit comments about HBO's Euphoria to understand viewers' experiences of and reactions to substance use and mental illness," Drug and Alcohol Dependence 2020.

Refereed Conference Publications

1. (NYU–S; SA)*Ilias Triantafyllopoulos, Renyi Qu, Salvatore Giorgi, Brenda Curtis, Lyle H Ungar, João Sedoc, "Knowing When Not to Answer: Lightweight KB-Aligned OOD Detection for Safe RAG", Proceedings of the Association for Computational Linguistics (ACL) 2026.

    NOTE: previous title "Simple Methods Defend RAG Systems Well Against Real-World Attacks".
2. (EF; NYU-S) *Baotong Zhang, João Sedoc, "Human-LLM Deliberation as an Interactive Zero-knowledge Proof Protocol", International Conference on Information Systems (ICIS) 2025.
3. (SA) Craig Thomson, Ehud Reiter, João Sedoc, Anya Belz, "Evolving Stances on Reproducibility: A Longitudinal Study of NLP and ML Researchers' Views and Experience of Reproducibility", Findings of EMNLP 2025.
4. (SA) Tingting Liu, Salvatore Giorgi, Ankit Aich, Allison Lahnala, Brenda Curtis, Lyle

Ungar, João Sedoc, "The Illusion of Empathy: How AI Chatbots Shape Conversation Perception", AAAI 2025.

NOTE: presented at WITS 2024, best paper nominee.

5. (SA) *Britney Ngaw, Grishma Jena, João Sedoc, Aline Normoyle, "Towards Authoring Open-Ended Behaviors for Narrative Puzzle Games with Large Language Model Support", International Conference on the Foundations of Digital Games 2024.
6. (SA) Salvatore Giorgi, Tingting Liu, Ankit Aich, Kelsey Isman, Garrick Sherman, Zachary Fried, João Sedoc, Lyle H Ungar, Brenda Curtis, "Explicit and Implicit Large Language Model Personas Generate Opinions but Fail to Replicate Deeper Perceptions and Biases", Findings of EMNLP 2024.
7. (SA) Salvatore Giorgi, Tingting Liu, Ankit Aich, Kelsey Isman, Garrick Sherman, Zachary Fried, João Sedoc, Lyle H Ungar, Brenda Curtis, "Modeling Human Subjectivity in LLMs Using Explicit and Implicit Human Factors in Personas", Findings of EMNLP 2024.
8. (EF; NYU-S) *Lining Zhang, João Sedoc, Natalia Levina, "Back to Principles: Theory-driven Evaluation of AI-based Conversational Agents", International Conference on Information Systems (ICIS) 2024.
9. (SA) Nikita Soni, H. Andrew Schwartz, João Sedoc, Niranjan Balasubramanian, "Large Human Language Models: A Need and the Challenges", Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL) 2024.
10. (SA) Marcel Nawrath, Agnieszka Nowak, Tristan Ratz, Danilo Walenta, Juri Opitz, Leonardo Ribeiro, João Sedoc, Daniel Deutsch, Simon Mille, Yixin Liu, Sebastian Gehrmann, Lining Zhang, Saad Mahamood, Miruna Clinciu, Khyathi Chandu, Yufang Hou, "On the Role of Summary Content Units in Text Summarization Evaluation", Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL) 2024.
11. (SA) Salvatore Giorgi, Douglas Bellew, Daniel Roy Sadek Habib, Garrick Sherman, João Sedoc, Chase Smitterberg, Amanda Devoto, McKenzie Himelein-Wachowiak, Brenda Curtis, "Lived Experience Matters: Automatic Detection of Stigma on Social Media Toward People Who Use Substances," The International AAAI Conference on Web and Social Media (ICWSM) 2024. Note: previous pre-print title "Lived Experience Matters: Automatic Detection of Stigma toward People Who Use Substances on Social Media".
12. (SA) *Renyi Qu, Lyle Ungar, João Sedoc, "Conditioning on Dialog Acts improves Empathy Style Transfer," Findings of Empirical Methods in Natural Language Processing 2023.
13. (SA) *Li Yifei, Lyle Ungar, João Sedoc, "Conceptor-Aided Debiasing of Large Language Models," Empirical Methods in Natural Language Processing (EMNLP) 2023.
14. (SA) Young Min Cho, Sunny Rai, Lyle Ungar, João Sedoc, Sharath Chandra Guntuku, "An Integrative Survey on Mental Health Conversational Agents to Bridge Computer Science and Medical Perspectives," Empirical Methods in Natural Language Processing (EMNLP) 2023.
15. (SA) *Lining Zhang, Simon Mille, Yufang Hou, Sebastian Gehrmann, Daniel Deutsch, Elizabeth Clark, Yixin Liu, Miruna Clinciu, Saad Mahamood, Khyathi Chandu,

João Sedoc, "Needle in a Haystack: An Analysis of Finding Qualified Workers on MTurk for Summarization," Association for Computational Linguistics (ACL), 2023.

16. (SA) *Seunggun Lee, Alexandra DeLucia, Ryan Guan, Rubing Li, Nikita Nangia, Shalaka Vaidya, Lining Zhang, Zijun Yuan, Praneeth Ganedi, Britney Ngaw, Aditya Singhal, João Sedoc, "Common Law Annotations: Investigating the Stability of Dialog Annotations," Findings of Association for Computational Linguistics (ACL), 2023.

    Note: previous published: "Common Law Annotations: Investigating the Stability of Dialog Annotations", HCOMP Abstracts 2023.
17. (NYU–S; EF) Prasanna Parasurama, João Sedoc, "Gendered Information in Resumes and its Role in Algorithmic and Human Hiring Bias," Academy of Management Proceedings 2022.
18. (SA) Jeevesh Juneja, Rachit Bansal, Kyunghyun Cho, João Sedoc, Naomi Saphra, "Linear Connectivity Reveals Generalization Strategies," Eleventh International Conference on Learning Representations (ICLR) 2023.
19. (SA) Shai Gretz, Assaf Toledo, Roni Friedman, Dan Lahav, Rose Weeks, Naor Bar-Zeev, João Sedoc, Pooja Sangha, Yoav Katz, Noam Slonim, "Benchmark Data and Evaluation Framework for Intent Discovery Around COVID-19 Vaccine Hesitancy," Findings of European Chapter of the Association for Computational Linguistics (EACL) 2023.
20. (SA) *Yukun Feng, Patrick Xia, Benjamin Van Durme, João Sedoc, "Automatic Document Selection for Efficient Encoder Pretraining," Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2022.
21. (SA) *Zetian Wu, Yilin Geng, Tejas Srivastava, Roshan Santhosh, Lyle Ungar, João Sedoc, "Inducing Generalizable and Interpretable Lexica," Findings of Empirical Methods in Natural Language Processing (EMNLP) 2022.
22. (SA) Ann-Katrin Reuel, Sebastian Peralta, João Sedoc, Garrick Sherman, Lyle Ungar, "Measuring the Language of Self-Disclosure across Corpora," Findings of the Association for Computational Linguistics (ACL), 2022.
23. (NYU–S; EF) Prasanna Parasurama, João Sedoc, "Gendered Language in Resumes – An Empirical Analysis of Gender Norm Violation and Hiring Outcomes," International Conference on Information Systems (ICIS) 2021.
24. (SA) Qizhen Zhang, Kelvin KW Ng, Charles Kazer, Shen Yan, João Sedoc, Vincent Liu, "MimicNet: Fast Performance Estimates for Data Center Networks with Machine Learning," ACM SIGCOMM 2021 Conference, 2021.

    Note: workshop publication as "Mimics: Fast Data Center Network Simulation Through Approximation".

    Note: prior workshop publication – Charles W. Kazer, João Sedoc, Kelvin K.W. Ng, Vincent Liu, Lyle Ungar, "Fast Network Simulation Through Approximation or: How Blind Men Should Describe Elephants," Workshop on Hot Topics in Networks (HotNets) 2018.
25. (SA) Huda Khayrallah, João Sedoc, "Measuring the 'I don't know' Problem through the Lens of Gricean Quantity," 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2021.
26. (SA) Patrick Xia, João Sedoc, Ben van Durme, "Incremental Neural Coreference Resolution in Constant Memory," Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020. Note: pre-print previous title "Revisiting Memory-Efficient Incremental Coreference Resolution".

27. (SA) Nathaniel Wier, João Sedoc, Ben van Durme, "COD3S: Diverse Generation with Discrete Semantic Signatures," Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020.
28. (SA) Huda Khayrallah, João Sedoc, "SMRT Chatbots: Improving Non-Task-Oriented Dialog with Simulated Multiple Reference Training," Findings in EMNLP 2020.
29. (SA) João Sedoc*, Sven Buechel*, Yoni Nachmany, Anneke Buffone, Lyle Ungar, "Generating Empathy and Distress Lexica from User Responses," LREC 2020.
30. (SA) Daphne Ippolito, Reno Kriz, Maria Kustikova, João Sedoc, Chris Callison-Burch, "Comparison of Diverse Decoding Methods from Conditional Language Models," Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL) 2019.
31. (SA) Tianyuan Zhou, João Sedoc, Jordan Rodu, "Getting in Shape: Word Embedding SubSpaces," International Joint Conference on Artificial Intelligence (IJCAI) 2019.
32. (SA) João Sedoc*, Daphne Ippolito*, Arun Kirubarajan, Jai Thirani, Lyle Ungar, Chris Callison-Burch, "ChatEval: A Tool for Chatbot Evaluation," 2019 Demo System, Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2019.

    Note prior workshop publication as "ChatEval: A Tool for the Systematic Evaluation of Chatbots," Workshop on Intelligent Interactive Systems and Language Generation (2IS&NLG) 2018.
33. (SA) *Tianlin Liu, Lyle Ungar, João Sedoc, "Continual Learning for Sentence Representations Using Conceptors," 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2019.
34. (SA) Reno Kriz, João Sedoc, Marianna Apidianaki, Carolina Zheng, Gaurav Kumar, Eleni Miltsakaki, Chris Callison-Burch, "Complexity-Weighted Loss and Diverse Reranking for Sentence Simplification," 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2019.
35. (SA) *Tianlin Liu, Lyle Ungar, João Sedoc, "Unsupervised Post-processing of Word Vectors via Conceptor Negation," Thirty-Third AAAI Conference on Artificial Intelligence (AAAI) 2019.
36. (SA) Sven Buechel, Anneke Buffone, Barry Slaff, Lyle Ungar, João Sedoc, "Modeling Empathy and Distress in Reaction to News Stories," Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2018.
37. (SA) João Sedoc, Daniel Preoţiuc-Pietro, Lyle Ungar, "Lexicon Expansion using Distributional Representations and Signed Clustering," Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL) 2017.
38. (SA) João Sedoc, Jean Gallier, Lyle Ungar, Dean Foster, "Semantic Word Clusters Using Signed Spectral Clustering," Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL) 2017.

Refereed Workshop Publications

1. (EF) Sofie Goethals, João Sedoc, Foster Provost, "What If the Prompt Were Different? Counterfactual Explanations for the Characteristics of Generative Outputs", Adjunct Proceedings of the 33rd ACM Conference on User Modeling, Adaptation and Personalization, 2025.

2. (EF) Andrew Smart, Ben Hutchinson, Lameck Mbangula Amugongo, Suzanne Dikker, Alex Zito, Amber Ebinama, Zara Wudiri, Ding Wang, Erin van Liemt, João Sedoc, Seyi Olojo, Stanley Uwakwe, Edem Wornyo, Sonja Schmer-Galunder, Jamila Smith-Loud, "Socially Responsible Data for Large Multilingual Language Models", Workshop on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO), 2024.
3. (SA) Aline Normoyle, João Sedoc, Funda Durupinar, "Using LLMs to Animate Interactive Story Characters with Emotions and Personality", IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) 2024.
4. (SA) Emma O'neil, João Sedoc, Diyi Yang, Haiyi Zhu, Lyle Ungar, "Automatic Reflection Generation for Peer-to-Peer Counseling", GEM Workshop, 2023.
5. (EF) Pedro Rodriguez, Phu Mon Htut, John P. Lalor, João Sedoc, "Clustering Examples in Multi-Dataset Benchmarks with Item Response Theory," Proceedings of the Third Workshop on Insights from Negative Results in NLP 2022.
6. (SA; NYU–S) John Mendonça, Lining Zhang, Rahul Mallidi, Alon Lavie, Isabel Trancoso, Luis Fernando D'Haro, João Sedoc,"Overview of Dialog System Evaluation Track: Dimensionality, Language, Culture and Safety at DSTC 12", Proceedings of the Twelfth Dialog System Technology Challenge (DSTC), 2025.
7. (A–Z) Sebastian Gehrmann, Abhik Bhattacharjee, Abinaya Mahendiran, Alex Wang, Alexandros Papangelis, Aman Madaan, Angelina McMillan-Major, Anna Shvets, Ashish Upadhyay, Bingsheng Yao, Bryan Wilie, Chandra Bhagavatula, Chaobin You, Craig Thomson, Cristina Garbacea, Dakuo Wang, Daniel Deutsch, Deyi Xiong, Di Jin, Dimitra Gkatzia, Dragomir Radev, Elizabeth Clark, Esin Durmus, Faisal Ladhak, Filip Ginter, Genta Indra Winata, Hendrik Strobelt, Hiroaki Hayashi, Jekaterina Novikova, Jenna Kanerva, Jenny Chim, Jiawei Zhou, Jordan Clive, Joshua Maynez, João Sedoc, Juraj Juraska, Kaustubh Dhole, Khyathi Raghavi Chandu, Leonardo FR Ribeiro, Lewis Tunstall, Li Zhang, Mahima Pushkarna, Mathias Creutz, Michael White, Mihir Sanjay Kale, Moussa Kamal Eddine, Nico Daheim, Nishant Subramani, Ondrej Dusek, Paul Pu Liang, Pawan Sasanka Ammanamanchi, Qi Zhu, Ratish Puduppully, Reno Kriz, Rifat Shahriyar, Ronald Cardenas, Saad Mahamood, Salomey Osei, Samuel Cahyawijaya, Sanja Štajner, Sebastien Montella, Shailza Jolly, Simon Mille, Tahmid Hasan, Tianhao Shen, Tosin Adewumi, Vikas Raunak, Vipul Raheja, Vitaly Nikolaev, Vivian Tsai, Yacine Jernite, Ying Xu, Yisi Sang, Yixin Liu, Yufang Hou, "GEMv2: Multilingual NLG Benchmarking in a Single Line of Code", GEM, 2022.
8. (EF) Ryuichiro Higashinaka, Luis F. D'Haro, Bayan Abu Shawar, Rafael E Banchs, Kotaro Funakoshi, Michimasa Inaba, Yuiko Tsunomori, Tetsuro Takahashi, João Sedoc, "Overview of the dialogue breakdown detection challenge 4," 10th International Workshop on Spoken Dialogue Systems, 2021.
9. (SA) *Alexandra DeLucia, Aaron Mueller, Xiang Lisa Li, João Sedoc, "Decoding Methods for Neural Narrative Generation," GEM Workshop at ACL 2021.
10. (SA; NYU–S) Prasanna Parasurama, João Sedoc, "Gendered Language in Resumes and its Implications for Algorithmic Bias in Hiring", Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP), 2021.
11. (SA) *Shuang Gao, Shivani Pandya, Smisha Agarwal, João Sedoc, "Topic Modeling for Maternal Health Using Reddit," 12th International Workshop on Health Text Mining and Information Analysis (LOUHI) 2021.
12. (SA) *Darren Edmonds, João Sedoc, "Multi-Emotion Classification for Song Lyrics,"

11th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA) 2021.

13. (SA) Sven Buechel, João Sedoc, H. Andrew Schwartz, Lyle Ungar, "Learning Neural Emotion Analysis from 100 Observations: The Surprising Effectiveness of Pre-Trained Word Representations." Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES) 2020.

    NOTE: previous arXiv title "Learning Emotion from 100 Observations: Unexpected Robustness of Deep Learning under Strong Data Limitations".
14. (SA) João Sedoc, Lyle Ungar, "Item Response Theory for Efficient Human Evaluation of Chatbots," EMNLP 2020 Workshop: Evaluation and Comparison of NLP Systems (Eval4NLP) 2020.
15. (SA) Adam Poliak, Max Fleming, Cash Costello, Kenton W Murray, Mahsa Yarmohammadi, Shivani Pandya, Darius Irani, Milind Agarwal, Udit Sharma, Shuo Sun, Nicola Ivanov, Lingxi Shang, Kaushik Srinivasan, Seolhwa Lee, Xu Han, Smisha Agarwal, João Sedoc,"Collecting Verified COVID-19 Question Answer Pairs," EMNLP 2020 Workshop: NLP COVID-19 Workshop (Part 2) 2020.
16. (SA) *Seolhwa Lee, João Sedoc,"Using the Poly-encoder for a COVID-19 Question Answering System," EMNLP 2020 Workshop: NLP COVID-19 Workshop (Part 2) 2020.
17. (SA) Gaurav Kumar, Bhavna Saluja, João Sedoc, Chris Callison-Burch, "Anonymization of Sensitive Information in Medical Health Records," MEDDOCAN Iberian Languages Evaluation Forum Workshop 2019.
18. (SA) *Saket Karve, Lyle Ungar, João Sedoc, "Conceptor Debiasing of Word Representations Evaluated on WEAT," Workshop on Gender Bias for Natural Language Processing 2019.
19. (SA) João Sedoc, Lyle Ungar, "The Role of Target Word Lists in Bias Identification of Contextualized Word Representations," Workshop on Gender Bias for Natural Language Processing 2019.
20. (SA) *Tianlin Liu, João Sedoc, Lyle Ungar, "Correcting the Common Discourse Bias in Linear Representation of Sentences using Conceptors," BioCreative/OHNLP Challenge 2018.
21. João Sedoc, "Hierarchical Methods for a Unified Approach to Discourse, Domain, and Style in Neural Conversational Models," Thirty-Second AAAI Conference on Artificial Intelligence (AAAI) Doctoral Consortium 2018.
22. (SA) Sajal Choudhary, Prerna Srivastava, Lyle Ungar, João Sedoc, "Domain Aware Neural Dialog System," Workshop in Machine Translation, Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics 2017.
23. (SA) Grishma Jena, Mansi Vashisht, Abheek Basu, Lyle Ungar, João Sedoc, "Enterprise to Computer: Star Trek chatbot," WiNLP workshop of the Association for Computational Linguistics 2017.
24. (SA) João Sedoc, Jordan Rodu, Lyle Ungar, Dean Foster, "Estimating Stock Covariance Using Factorial Hidden Markov Models," PGMO-COPI 2014.

In submission

1. *Damilola Omitaomu, Shabnam Tafreshi, Sven Buechel, Chris Callison-Burch, Lyle Ungar, João Sedoc, "Empathic Conversations: A Multi-level Dataset of Contextualized Conversations," in submission: revise and submission Transcations of the Association of Computational Linguistics.

Note: More than 100 citations using the dataset or models. Data already released in Shared Task papers:

"Findings of the WASSA 2024 Shared Task on Empathy and Personality Detection in Interactions"

"Findings of WASSA 2024 Shared Task on Empathy and Personality Detection in Interactions"

2. (EF) João Sedoc and Vasant Dhar, "The Damodaran Bot: Artificial Intelligence for Systematic Long-Term Valuation," in submission Management Science.

   NOTE: revision of (A–Z)Vasant Dhar and João Sedoc, "DBOT: Artificial Intelligence for Systematic Long-Term Investing".
3. Sofie Goethals, João Sedoc, Foster Provost, "Prompt-Counterfactual Explanations for Generative AI System Behavior", in submission MISQ.

   Formerly titled : "Explanations for Generative AI System Behavior"
4. (A–Z) Xiang Cheng, Raveesh Mayya, João Sedoc, "To Err Is Human; To Annotate, SILICON? Reducing Measurement Error in LLM Annotation", Management Science – Feb 15th.

   Note: previous title "From Human Annotation to LLMs: SILICON Annotation Workflow for Management Research"

   Note: accepted in CIST 2025

In Progress

1. (EF) Charles Welch, Allison Lahnala, Vasudha Varadarajan, Lucie Flek, Rada Mihalcea, J Lomax Boyd, João Sedoc, "ISCA: A Framework for Interview-Style Conversational Agents".
2. (EF) João Sedoc and Dean Foster, "Detecting Directions of Chatbot Quality using Lighthouse Variables in Item Response Theory", ISR target October.

   Note: extension of "Item Response Theory for Efficient Human Evaluation of Chatbots".
3. (NYU–S; EF) *Lining Zhang, João Sedoc, Natalia Levina, "Back to Principles: Theory-driven Evaluation of AI-based Conversational Agents," target ISR submission Fall 2026.
4. (NYU–S) Rubing Li; João Sedoc; Arun Sundararajan, "Reasoning and the Trusting Behavior of DeepSeek and GPT: An Experiment Revealing Hidden Fault Lines in Large Language Models".
5. (NYU–S;EF) Prasanna Parasurama, João Sedoc, Anindya Ghose, "Degendering Resumes for Fair Algorithmic Resume Screening".
6. (SA) Salvatore Giorgi, Shreya Havaldar, Farhan Ahmed, Zuhaib Akhtar, Shalaka Vaidya, Gary Pan, Lyle H Ungar, H Andrew Schwartz, João Sedoc, "Psychological Metrics for Dialog System Evaluation."

   Note previous pre-print title "Human-Centered Metrics for Dialog System Evaluation".
7. (SA) Xu Han, Felix Yu, João Sedoc, Benjamin Van Durme, "Baby Bear: Seeking a Just Right Rating Scale for Scalar Annotations".

ArXiv

1. (SA) Roni Friedman*, João Sedoc*, Shai Gretz, Assaf Toledo, Rose Weeks, Naor

Bar-Zeev, Yoav Katz, Noam Slonim, "VIRATrustData: A Trust-Annotated Corpus of Human-Chatbot Conversations About COVID-19 Vaccines".

2. (SA) *Seolhwa Lee, Adam Poliak, João Sedoc, "Towards a Comprehensive Evaluation Design for Dialogue System Evaluation."

   Note: previous pre-print on arXiv as "An Evaluation Protocol for Generative Conversational Systems".

3. (A-Z) Sebastian Gehrmann, Tosin Adewumi, Karmanya Aggarwal, Pawan Sasanka Ammanamanchi, Aremu Anuoluwapo, Antoine Bosselut, Khyathi Raghavi Chandu, Miruna Clinciu, Dipanjan Das, Kaustubh D. Dhole, Wanyu Du, Esin Durmus, Ondřej Dušek, Chris Emezue, Varun Gangal, Cristina Garbacea, Tatsunori Hashimoto, Yufang Hou, Yacine Jernite, Harsh Jhamtani, Yangfeng Ji, Shailza Jolly, Mihir Kale, Dhruv Kumar, Faisal Ladhak, Aman Madaan, Mounica Maddela, Khyati Mahajan, Saad Mahamood, Bodhisattwa Prasad Majumder, Pedro Henrique Martins, Angelina McMillan-Major, Simon Mille, Emiel van Miltenburg, Moin Nadeem, Shashi Narayan, Vitaly Nikolaev, Rubungo Andre Niyongabo, Salomey Osei, Ankur Parikh, Laura Perez-Beltrachini, Niranjan Ramesh Rao, Vikas Raunak, Juan Diego Rodriguez, Sashank Santhanam, João Sedoc, Thibault Sellam, Samira Shaikh, Anastasia Shimorina, Marco Antonio Sobrevilla Cabezudo, Hendrik Strobelt, Nishant Subramani, Wei Xu, Diyi Yang, Akhila Yerukola, Jiawei Zhou, "The GEM Benchmark: Natural Language Generation, its Evaluation and Metrics."
4. (SA) Huda Khayrallah, Zuhaib Akhtar, Edward Cohen, João Sedoc, "How to Choose How to Choose Your Chatbot: A Massively Multi-System MultiReference Data Set for Dialog Metric Evaluation".
5. (EF) *Qi He, João Sedoc, Jordan Rodu, "Trees in transformers: a theoretical analysis of the Transformer's ability to represent trees".

   NOTE: previousle as "Neural Tree Transducers for Tree to Tree Learning".

6. *Shuo Sun, João Sedoc, "An Analysis of BERT FAQ Retrieval Models for COVID-19 Infobot".
7. Joao Sedoc, Derry Wijaya, Masoud Rouhizadeh, Andy Schwartz, Lyle Ungar, "Deriving Verb Predicates By Clustering Verbs with Arguments".

Shared Task Papers

1. John Mendonça, Lining Zhang, Rahul Mallidi, Alon Lavie, Isabel Trancoso, Luis Fernando D'Haro, João Sedoc, "DSTC12: Dialogue System Technology Challenge 12", 26th Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2025.
2. (EF) Simon Mille, João Sedoc, Yixin Liu, Elizabeth Clark, Agnes Johanna Axelsson, Miruna Adriana Clinciu, Yufang Hou, Saad Mahamood, Ishmael Nyunya Obonyo, Lining Zhang, "The 2024 GEM shared task on multilingual data-to-text generation and summarization: Overview and preliminary results", Proceedings of the 17th International Natural Language Generation Conference: Generation Challenges, 2025.
3. (EF) Salvatore Giorgi, João Sedoc, Valentin Barriere, Shabnam Tafreshi, "Findings of WASSA 2024 Shared Task on Empathy and Personality Detection in Interactions", Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2024.
4. Mario Rodríguez-Cantelar, Chen Zhang, Chengguang Tang, Ke Shi, Sarik Ghazarian,

João Sedoc, Luis Fernando D'Haro, Alexander Rudnicky, "Overview of Robust and Multilingual Automatic Evaluation Metrics for Open-domain Dialogue Systems at DSTC 11 Track 4", Proceedings of The Eleventh Dialog System Technology Challenge, 2023.

5. (EF) Valentin Barriere, João Sedoc, Shabnam Tafreshi, Salvatore Giorgi, "Findings of WASSA 2023 Shared Task on Empathy, Emotion and Personality Detection in Conversation and Reactions to News Articles", Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2023.
6. (EF) Valentin Barriere, Shabnam Tafreshi, João Sedoc, Sawsan Alqahtani, "WASSA 2022 Shared Task: Predicting Empathy, Emotion and Personality in Reaction to News Stories", Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, 2022.
7. (EF) Shabnam Tafreshi, Orphee De Clercq, Valentin Barriere, Sven Buechel, João Sedoc, Alexandra Balahur, "WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories", Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2021.
8. (SA) Chen Zhang, João Sedoc, Luis Fernando D'Haro, Rafael Banchs, Alexander Rudnicky, "Automatic evaluation and moderation of open-domain dialogue systems", DSTC 10, 2021.

Workshop & Tutorial Proceedings

1. (EF) Aleksandr Drozd, João Sedoc, Shabnam Tafreshi, Arjun Akula, Raphael Shu, "Proceedings of the Sixth Workshop on Insights from Negative Results in NLP", The Sixth Workshop on Insights from Negative Results in NLP, 2025.
2. (EF) Orphee De Clercq, Valentin Barriere, Jeremy Barnes, Roman Klinger, João Sedoc, Shabnam Tafreshi, "Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis", Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, 2024.
3. Simone Balloccu, Anja Belz, Rudali Huidrom, Ehud Reiter, João Sedoc, Craig Thomson, "Proceedings of the Fourth Workshop on Human Evaluation of NLP Systems (HumEval)@ LREC-COLING 2024", Proceedings of the Fourth Workshop on Human Evaluation of NLP Systems (HumEval)@ LREC-COLING 2024.
4. Adithya V Ganesan, Siddharth Mangalik, Vasudha Varadarajan, Nikita Soni, Swanie Juhng, João Sedoc, H Andrew Schwartz, Salvatore Giorgi, Ryan L Boyd, "From Text to Context: Contextualizing Language with Humans, Groups, and Communities for Socially Aware NLP, Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 5: Tutorial Abstracts) (NAACL) 2024.
5. (EF) Anja Belz, João Sedoc, Craig Thomson, Simon Mille, Rudali Huidrom, "The INLG 2024 Tutorial on Human Evaluation of NLP System Quality: Background, Overall Aims, and Summaries of Taught Units", Proceedings of the 17th International Natural Language Generation Conference: Tutorial Abstract, 2024.
6. Shabnam Tafreshi, Arjun Akula, João Sedoc, Aleksandr Drozd, Anna Rogers, Anna Rumshisky, "Proceedings of the Fourth Workshop on Insights from Negative Results in NLP", Proceedings of the Fourth Workshop on Insights from Negative Results in NLP, 2023.
7. Sebastian Gehrmann, Alex Wang, João Sedoc, Elizabeth Clark, Kaustubh Dhole, Khyathi Raghavi Chandu, Enrico Santus, Hooman Sedghamiz, "Proceedings of the

Third Workshop on Natural Language Generation, Evaluation, and Metrics (GEM)", Proceedings of the Third Workshop on Natural Language Generation, Evaluation, and Metrics (GEM), 2023.

8. John P Lalor, Pedro Rodriguez, João Sedoc, Jose Hernandez-Orallo, "Item Response Theory for Natural Language Processing, Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts (EACL), 2023.
9. Anja Belz, Maja Popović, Ehud Reiter, Craig Thomson, João Sedoc, "Proceedings of the 3rd Workshop on Human Evaluation of NLP Systems", Proceedings of the 3rd Workshop on Human Evaluation of NLP Systems, 2023.
10. (A–Z) Koichiro Yoshino, Yun-Nung Chen, Paul Crook, Satwik Kottur, Jinchao Li, Behnam Hedayatnia, Seungwhan Moon, Zhengcong Fei, Zekang Li, Jinchao Zhang, Yang Feng, Jie Zhou, Seokhwan Kim, Yang Liu, Di Jin, Alexandros Papangelis, Karthik Gopalakrishnan, Dilek Hakkani-Tur, Babak Damavandi, Alborz Geramifard, Chiori Hori, Ankit Shah, Chen Zhang, Haizhou Li, João Sedoc, Luis F D'Haro, Rafael Banchs, Alexander Rudnicky, "Overview of the Tenth Dialog System Technology Challenge: DSTC10", IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023.
11. (EF) Jeremy Barnes, Orphee De Clercq, Valentin Barriere, Shabnam Tafreshi, Sawsan Alqahtani, João Sedoc, Roman Klinger, Alexandra Balahur, "Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis", Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, 2022.
12. (EF) Shabnam Tafreshi, João Sedoc, Anna Rogers, Aleksandr Drozd, Anna Rumshisky, Arjun Akula, "Proceedings of the Third Workshop on Insights from Negative Results in NLP", Proceedings of the Third Workshop on Insights from Negative Results in NLP, 2022.
13. (EF) João Sedoc, Anna Rogers, Anna Rumshisky, Shabnam Tafreshi, "Proceedings of the Second Workshop on Insights from Negative Results in NLP, Proceedings of the Second Workshop on Insights from Negative Results in NLP, 2021.
14. (EF) Orphee De Clercq, Alexandra Balahur, João Sedoc, Valentin Barriere, Shabnam Tafreshi, Sven Buechel, Veronique Hoste, "Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis", Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2021.
15. (EF) Anna Rogers, João Sedoc, Anna Rumshisky, "Proceedings of the First Workshop on Insights from Negative Results in NLP", Proceedings of the First Workshop on Insights from Negative Results in NLP, 2020.

Popular Press

1. "People are falling in love with chatbots For users of companionship apps, AI can be a therapist, friend, and even a lover. These relationships feel real — but are they ethical?", Boston Globe
2. "An AI tool is trying to predict your risk of getting many diseases years in advance – here's how it works", The Conversation
3. "2024 Best Undergraduate Professors", Poets & Quants

Teaching

University of Pennsylvania CTL Teaching Certificate 2019

Co-Instructor — Spring 2024, Spring 2026
Research Seminar: Data Science — New York University

Instructor — Spring 2026
Projects in Programming and AI — New York University

Instructor — Spring & Fall (2021, 2022), Spring (2023, 2024, 2025)
Projects in Programming and Data Science — New York University

Instructor — Spring 2020
Deep Learning Methods for Automated Discourse — Johns Hopkins University

Co-Instructor — Fall 2019
Artificial Intelligence — Johns Hopkins University

Instructor (unofficial) — Spring 2017
Special Topics - Deep Learning Methods for Automated Discourse — University of Pennsylvania

Teaching Assistant — Fall 2014
Graduate Linear Algebra taught by Jean Gallier — University of Pennsylvania

Teaching Assistant — Summer 2014
Graduate Theory of Computation taught by Jean Gallier — University of Pennsylvania

Invited Talks

Expert panel on capabilities and safety, COLM 2025 Workshop on AI Agents, October 2025.
"User Perceptions of Empathy in AI vs. Human Conversations About Distressing News Events," UMD Workshop on AI & Analytics for Social Good, May 2025.
Bias Mitigation in Large Language Models, UT Austin, September 2024.
Workshop on Computational Linguistics and Clinical Psychology (CLPsych), Empathy and Emotion in Conversational AI, March 2024.
Herman Sno Stichting, Developments and Applications of Generative AI, March 2024.
Stony Brook University, New Dimensions for the Evaluation of Conversational Agents, January 2024.
The 2023 Singapore Symposium on Natural Language Processing (SSNLP), New Dimensions for the Evaluation of Conversational Agents, December 2023.
University of Minnesota, New Dimensions for the Evaluation of Conversational Agents, December 2023.
Amazon Science, April 2023.
ADAPT and DCU NLP Seminar Series, Addressing Challenges in Annotating for Large Language Model Evaluation, March 2023.
Amazon, Evaluating Conversational Agents, October 2021.
NYU Langone HiBRID Lab Talks, Perspectives and Evaluation of Conversational Agents in Healthcare Application, September 2021.
JRC Community of Practice on Text Mining Meeting, Modeling Empathy and Distress

in Reaction to News Stories, February 2021.
University of Miami CSC Colloquium, Evaluating Conversational Agents, November 2020.
Rigorous Evaluation for AI Systems Workshop, Evaluating Conversational Agents, October 2020.
International Workshop on Spoken Dialogue Systems, ChatEval & CHANEL Overview, September 2020.
Bloomberg, Multiscale Methods for Covariance Estimation and Assumptions made in Academic Timeseries Forecasting, February 2020.
Capital One, Evaluating Conversational Agents, November 2019.
University of Maryland, Evaluating Conversational Agents, November 2019.
New York Academy of Sciences, Evaluating Conversational Agents, November 2019.
Boston College, Practitioner's Perspective on Quantitative Investing, November 2019.
Spotify, Building and Evaluating Conversational Agents, August 2019.
2nd International Conversational Intelligence Summer School, Evaluating Conversational Agents, June 2019.
NYU Medical School, Building and Evaluating Conversational Agents, May 2019.
Temple University, Building and Evaluating Conversational Agents, April 2019.
Swarthmore College, Chivalric High Frequency Trading Platform Design, April 2019.
University of Pennsylvania e-talks, Modeling Empathy and Distress in Reaction to News Stories, March 2019.
International Convention of Psychological Science, Modeling Empathy and Distress in Reaction to News Stories, March 2019.
Johns Hopkins University, Building and Evaluating Conversational Agents, February 2019.
Columbia University, Building and Evaluating Conversational Agents, February 2019.
NYU Stern, Building and Evaluating Conversational Agents, February 2019.
Didi Labs, Building and Evaluating Conversational Agents, January 2019.
Boston University, Building and Evaluating Conversational Agents, December 2018.
Grammarly, Building and Evaluating Conversational Agents, November 2018.
University of Delaware, Building and Evaluating Conversational Agents, October 2018.
Temple University, Building and Evaluating Conversational Agents, October 2018.
Boston College, Practitioner's Perspective on Quantitative Investing, November 2016.
Boston College, Practitioner's Perspective on Quantitative Investing, November 2015.

Awards

Poets & Quants Top 50 Undergraduate Professors (2024)
Best Paper Award at AOM 2022 for "Gendered Information in Resumes and its Role in Algorithmic and Human Hiring Bias".
Professor Marti Subrahmanyam Honors Program Faculty Service Award (2024)
Fontaine Travel Award (2017, 2019)
AAAI Doctoral Consortium Travel Award (2018)
Fontaine Fellowship (2016-2017)
J.P. Eckert Fellowship (2013)

NSF Young Scholars Program (1993)

Grants

Current Grants

Generative AI for anxiety, depression and psychosis (Lead PI) [$4,039,244] (2026-2028)
OpenAI's AI and Mental Health Grant Program (Lead PI) [$ 100,000] (2026)
Google Research Gift (GEM Benchmark Annotations) [$ 67,000] (2021-)

Past Grants

Wellcome Accelerator / Neuromatch [$ 27,000] (2025)
Amazon AWS Cloud Computing Research Credits [$25,000] (2019)
Microsoft Research Dissertation Grant [$25,000] (2018-2019)

Academic & Professional Service

Editorial Roles

Senior Area Chair (Dialogue & Interactive Systems) ACL 2023
Senior Area Chair (Dialogue & Interactive Systems) ACL 2024
Senior Area Chair (Machine Learning) ACL 2024
Senior Area Chair (Dialogue & Interactive Systems) NAACL 2025
Senior Area Chair (Dialogue & Interactive Systems) ACL 2026
Action Editor Transactions of the Association of Computational Linguistics (TACL) (2023-)
Action Editor special issue of Transactions on Audio, Speech and Language Processing
Action Editor ICIS 2022
Area Chair (Natural Language Generation) EMNLP 2022
Area Chair (Dialogue & Interactive Systems) ACL 2025
Area Chair ACL Rolling Reviews (2024-)

Tutorial, Conference Organizing, & Program Committee Service

Tutorial on Item Response Theory for Natural Language Processing, EACL 2024
Tutorial on From Text to Context: Contextualizing Language with Humans, Groups, and Communities for Socially Aware NLP, NAACL 2024
Tutorial on Human Evaluation of NLP System Quality: Background, Overall Aims, and Summaries of Taught Units, INLG 2024
Organizing committee, The Third & Fourth Version of the Generation, Evaluation & Metrics (GEM) Workshop, 2023, 2025
Organizing committee, The 3rd Workshop on Human Evaluation of NLP Systems, 2023
Organizing committee, The 1st, 2nd, 3rd, 4th, 5th, & 6th Workshop on Insights from Negative Results in NLP (2020-2025)

Organizing committee, 11th, 12th, 13th, 14th, & 15th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis (WASSA) (2021-2024)
Organizing committee, Dialog System Technology Challenge 12 Track 1: Robust and Multilingual Automatic Evaluation Metrics for Open-Domain Dialogue Systems, (2025)
Organizing committee, Dialog System Technology Challenge 11 Track 4: Robust and Multilingual Automatic Evaluation Metrics for Open-Domain Dialogue Systems, (2022-2023)
Organizing committee, Dialog System Technology Challenge 10 Track : Automatic Evaluation and Moderation of Open-domain Dialogue Systems, (2021-2022)
Organizing committee, Neural Conversational AI: Bridging the Gap Between Research and Real World (NeuCAIR) (2021)
Organizing committee, Workshop on Chatbots and Conversational Agents (WOCHAT) (2021)
Organizing committee, Workshop on Chatbots and Conversational Agents and Dialogue Breakdown Detection Challenge (WOCHAT+DBDC) (2020)
Organizing member Workshop on Chatbots and Conversational Agents and Dialogue Breakdown Detection Challenge (WOCHAT+DBDC) (2019)
PC member, Mid-Atlantic Student Colloquium on Speech, Language and Learning (2017)
Organizer, Mid-Atlantic Student Colloquium on Speech, Language and Learning (2016)
Organizer, Computational Linguistics Lunch (2014-2016)

Review Panels

NSF Review Panel (2023)

Reviewing Service

Reviewer: AAAI (2015, 2018)
Reviewer: ACL (2018-2022)
Reviewer: TACL (2021-2023)
Reviewer: Management Science (2023-)
Reviewer: Information Systems Research (2025-)
Reviewer: Management Information Systems Quarterly (2025-)

Other Service

Board member, Herman Sno Stichting (2017-)
Panelist, Brain Inspired NeuroMatch Academy 05: NLP and Generative Models Panel, (2021)
Ran the PennStat Natural Language Reading Group (2014-2024)
Student Mentor Lea Middle School (2015)
Graduate student advisor, Penn Emerging Scholars Program (2014-2018)

Other Service to NYU and Stern

Scientific Mentor to Endless Frontier Lab (2024-)

Fubon Innovation Fellow (2024-)
PhD Recruiting Committee (2022-2024)
Committee member on Teaching and Generative Tools (2023-)
co-organizing IS seminars (2024-2025)
organizing IS seminars (2025-2026)
co-organizing TOPS department seminars (2025-)
participating in SPUR program (2020-)
mentoring Stern Honor's Thesis (2021,2023-)

Student Advising

PhD Student Advising

Ilias Triantafyllopoulos (co-author and currently co-mentor)
Baotong Zhang (co-author and currently primary mentor)
Eunsol Cho (advised in NLP, committee member, and co-author)
Lining Zhang (advised in NLP and co-author)
Rubing Li (advised in NLP and co-author)
Prasanna Parasurama (advised in NLP and co-author)

Thesis Committees

Lavender Jiang, New York University (advisor: Eric Oermann and Kyunghyun Cho).
Jingyi Yang, New York University (advisor: Alexander Tuzhilin).
Eunsol Cho, New York University (advisor: Arun Sundararajan).
Michelle Nguyen, Johns Hopkins University (advisor: Casey Taylor).
Vishakh Padmakumar, New York University (advisor: He He), "Human-AI Collaboration with Large Language Models for Creative Writing", 2025.
Angelica Chen, New York University (advisor: Kyunghyun Cho), "Improving Language Models Through the Lens of Training Dynamics", 2025.
Nikita Nangia, New York University (advisor: Samuel R. Bowman), "Why, How, and When to Effectively Crowdsource Data for Natural Language Processing Research", 2024.
Prasanna Parasurama, New York University (advisor: Anindya Ghose and Panos Ipeirotis), "Bias and Fairness in Algorithmic Hiring Systems", 2023.
Tianyuan Zhou, University of Virginia (advisor: Jordan Rodu), "Research on Regime-Switching Between Different Stochastic Dynamics", 2023.
Alex Wang, New York University (advisor: Samuel R. Bowman and Kyunghyun Cho), "NLP Evaluation in the Time of Large Language Models", 2022.
Semi Min, New York University (advisor: Natalia Levina) "Approach for Understanding How Discourse Shapes the Meaning of Technologies", 2022.
Phu Mon Htut, New York University (advisor: Samuel R. Bowman and Kyunghyun Cho), "Analyzing And Comparing Natural Language Processing Models And Datasets", 2022.

Ruizhong Miao, University of Virginia (advisor: Tianxi Li and Jordan Rodu), "Informative Structures in Complex Networks", 2022.
Shijie Wu, Johns Hopkins University (advisor: Mark Dredze), "How Do Multilingual Encoders Learn Cross-Lingual Representation?", 2021.
Tongfei Chen, Johns Hopkins University (advisor: Benjamin Van Durme), "Ranking And Retrieval Under Semantic Relevance", 2020.
Adam Poliak (advisor: Benjamin Van Durme), "Revisiting Recognizing Textual Entailment For Evaluating Natural Language Processing Systems", 2020.

Undergraduate and Master's Students Advised

Senior Honors Theses Supervised

Daniel Liu, "KRUPPE: Knowledge-based Research Using Pathway and Procedural Engine", 2025.
Jesse Price (co-advised with Clifford Hurvich), "Benchmarking FedFormer Against Traditional Econometric Models: Testing Predictive Memory in Financial Time Series", 2025.
Kyuhyun Shim, "Decomposing Pairs Trading: A Machine Learning Approach to Analyzing Drivers of Profitability", 2024.
Jialiang Chen, "Volatility Forecasting with Physics-Informed Neural Networks (PINNs)", 2022.

Master's Theses Supervised

Shalaka Vaidya, "Deep Learning Methods for Modeling Emotion, Empathy and Personality in conversations ", 2023.
Zietan Wu (JHU), "Research on SHAP Values and Other Masking Methods for Lexicon Creation", 2021.

Student Research Advisor

Master's students (independent study)

Zietan Wu (JHU)
Qi He (JHU)
Anjali Agrawal
Shuang Gao
Zuhaib Akhtar
Lining Zhang
Shalaka Vaidya
Haojun Zhao

Undergraduate students (independent study / research assistant)

Damilola Omitaomu (UPenn)
Uju Duru (JHU)

Madina Diane (JHU)
Darren Edmonds (JHU)
Dylan Lewis (JHU)
Patricia Luk
Adi Singhal
Seunggun Lee
Gary Pan
Brandon Somoza
Jiali Xu
John Wang
Tanishi Mishra
Oscar Wan

Undergraduate students (Stern Program for Undergraduate Research [SPUR])

Shivam Patel
Zachary Nemeth
Britney Ngaw
Praneeth Ganedi
Allen Yuan
Jacob Padilla
Corbin Canavarros
Emily Wang
Kevin Wang
Pranjay Kumar
Suri Wu
Gwyneth Zhou
Siyansh Balyan
Will Wu
Ada Sun
Catherine Jiang
David Lauinger
Marcelo Barrera
Molly Zhang
Angelina Rosalie
Gordon Li
Jessica Ortuno
Linh Phan
Naina Sharma
Sree Gaur
Yi Mei Zhang