Ana Marasović, Curriculum Vitae, September 2024

www.anamarasovic.com
ana.marasovic@utah.edu

Merrill Engineering, Office 2166
50 Central Campus Dr
Salt Lake City, UT 84112

RESEARCH INTERESTS

Natural Language Processing, Human-Centered AI, Interpretability

EMPLOYMENT

Assistant Professor, **University of Utah, Kahelrt School of Computing**	Jul 2022 – present
Young Investigator, **Allen Institute for AI**	Mar 2019 – Jun 2022

Courtesy appointment at the University of Washington, Paul G. Allen School of Computer Science & Engineering

EDUCATION

Ph.D. in Computational Linguistics, Heidelberg University	2015 – 2019
Honors: Magna cum laude	
M.Sc. in Applied Mathematics, University of Zagreb	2013 – 2015
Honors: Summa cum laude	
B.Sc. in Mathematics, University of Zagreb	2010 – 2013

HONORS AND AWARDS

- ACL 2023 Best Paper Award
- ACL 2023 Best Senior Area Chair Award
- SoCal NLP Symposium 2022 Best Paper Award
- ACL 2020 Best Paper Honorable Mention Award
- UC Berkeley EECS 2020 Rising Star
- EMNLP 2020 Outstanding Reviewer Award
- EMNLP 2019 Outstanding Reviewer Award

FUNDED PROPOSALS

Allen Institute for AI, *Research Gift*, $100,000	Oct 2022
Title: "Expressing Uncertainty During Explanation Generation"	

PUBLICATIONS

Papers in Rigorously Reviewed Journals & Conferences

Findings papers undergo the same review process as main conference papers and are reviewed simultaneously. Main conference acceptances are typically within the top 23% of submissions, while Findings papers usually fall between the top 23% and 35%. Google Scholar citation count on September 10, 2024 is reported. 4,146 citations in total with an h-index of 17.

[1] Fateme Hashemi Chaleshtori, Atreya Ghosal, Alexander Gill, Purbid Bambroo, and **Ana Marasović**. *On Evaluating Explanation Utility for Human-AI Decision-Making in NLP*. Findings of EMNLP 2024.

[2] Oliver Bentham, Nathan Stringham, and **Ana Marasović**. *Chain-of-Thought Unfaithfulness as Disguised Accuracy.* TMLR 2024. 1 citation.

[3] Ashim Gupta, Rishanth Rajendhran, Nathan Stringham, Vivek Srikumar, and **Ana Marasović**. *Whispers of Doubt Amidst Echoes of Triumph in NLP Robustness.* NAACL 2024. 1 citation. 23.2% out of 2,434 submissions.

[4] Jack Hessel, **Ana Marasović**, Jena D. Hwang, Lillian Lee, Jeff Da, Rowan Zellers, Robert Mankoff, and Yejin Choi. *Do Androids Laugh at Electric Sheep? Humor "Understanding" Benchmarks from The New Yorker Caption Contest.* ACL 2023. 45 citations. 22.08% out of 4,864 submissions. **Received a Best Paper Award given to only 4 out of 1074 accepted papers.** Covered in Pyschology Today, The Debrief, and Cornell Chronicle.

[5] Abhilasha Ravichander, Matt Gardner, and **Ana Marasović**. *CondaQA: A Contrastive Reading Comprehension Dataset for Reasoning about Negation.* EMNLP 2022. 24 citations. 19.8% out of 4,190 submissions. **Received a Best Paper Award at the SoCal NLP Symposium 2022; one of 3 award-winning papers out of 88 accepted papers.**

[6] Alexis Ross, Matthew E. Peters, and **Ana Marasović**. *Does Self-Rationalization Improve Robustness to Spurious Correlations?* EMNLP 2022. 11 citations. 19.8% out of 4,190 submissions.

[7] Shruti Palaskar, Akshita Bhagia, Yonatan Bisk, Florian Metze, Alan W Black, and **Ana Marasović**. *On Advances in Text Generation from Images Beyond Captioning: A Case Study in Self-Rationalization.* Findings of EMNLP 2022. 2 citations. 32.9% out of 4,190 submissions.

[8] **Ana Marasović***, Iz Beltagy*, Doug Downey, and Matthew E. Peters. *Few-Shot Self-Rationalization with Natural Language Prompts.* Findings of NAACL 2022. ***Equal contributions.** 93 citations. 32.3% out of 2,013 submissions.

[9] Sarah Wiegreffe* and **Ana Marasović***. *Teach Me to Explain: A Review of Datasets for Explainable NLP.* NeurIPS Datasets and Benchmarks Track 2021. ***Equal contributions.** 185 citations.

[10] Sarah Wiegreffe, **Ana Marasović**, and Noah A. Smith. *Measuring Association Between Labels and Free-Text Rationales.* EMNLP 2021. 145 citations. 23.3% out of 3,600 submissions.

[11] Jesse Dodge, Maarten Sap, **Ana Marasović**, William Agnew, Gabriel Ilharco, Dirk Groeneveld, Margaret Mitchell, and Matt Gardner. *Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus.* EMNLP 2021. 369 citations. 23.3% out of 3,600 submissions.

[12] Kaiser Sun and **Ana Marasović**. *Effective Attention Sheds Light On Interpretability.* Findings of ACL 2021. 19 citations. 34.8% out of 3,350 submissions.

[13] Alexis Ross, **Ana Marasović**, and Matthew E. Peters. *Explaining NLP Models via Minimal Contrastive Editing (MiCE).* Findings of ACL 2021. 122 citations. 34.8% out of 3,350 submissions.

[14] Alexander M. Hoyle, **Ana Marasović**, and Noah A. Smith. *Promoting Graph Awareness in Linearized Graph-to-Text Generation.* Findings of ACL 2021. 24 citations. 34.8% out of 3,350 submissions

[15] Alon Jacovi, **Ana Marasović**, Tim Miller, and Yoav Godlberg. *Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI.* ACM FAccT 2021. 469 citations. 25.0% out of 328 submissions.

[16] **Ana Marasović**, Chandra Bhagavatula, Jae sung Park, Ronan Le Bras, Noah A. Smith, and Yejin Choi. *Natural Language Rationales with Full-Stack Visual Reasoning: From Pixels to Semantic Frames to Commonsense Graphs.* Findings of EMNLP 2020. 61 citations. 37.9% out of 3,359 submissions.

[17] Qiang Ning, Hao Wu, Pradeep Dasigi, Dheeru Dua, Matt Gardner, Robert L. Logan IV, **Ana Marasović**, and Zhen Nie. *Easy, Reproducible and Quality-Controlled Data Collection with* CrowdAQ. EMNLP System Demonstrations Track 2020. 16 citations. 25% out of 203 submissions.

[18] Suchin Gururangan, **Ana Marasović**, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A. Smith. *Don't Stop Pretraining: Adapt Language Models to Domains and Tasks.* ACL 2020. 2,226 citations. 22.7% out of 3,429 submissions. **Received an Honorable Mention for Best Paper Award; one of 5 award-winning papers out of 779 accepted papers.**

[19] Pradeep Dasigi, Nelson F. Liu, **Ana Marasović**, Noah A. Smith, and Matt Gardner. *Quoref: A Reading Comprehension Dataset with Questions Requiring Coreferential Reasoning.* EMNLP 2019. 191 citations. 23.7% out of 2,876 submissions.

[20] **Ana Marasović** and Anette Frank. *SRL4ORL: Improving Opinion Role Labeling Using Multi-Task Learning with Semantic Role Labeling.* NAACL 2018. 40 citations. 30.4% out of 1,122 submissions.

[21] **Ana Marasović**, Leo Born, Juri Opitz, and Anette Frank. *A Mention-Ranking Model for Abstract Anaphora Resolution.* EMNLP 2017. 36 citations. 22.8% out of 1,418 submissions.

[22] **Ana Marasović**, Mengfei Zhou, Alexis Palmer, and Anette Frank. *Modal Sense Classification At Large: Paraphrase-Driven Sense Projection, Semantically Enriched Classification Models and Cross-Genre Evaluations.* Linguistic Issues in Language Technology (LiLT) – Modality: Logic, Semantics, Annotation, and Machine Learning 14 (3). 2016. 23 citations.

Workshop Papers & Other Published Work

[1] Fateme Hashemi Chaleshtori, Atreya Ghosal, and **Ana Marasović**. *On Evaluating Explanation Utility for Human-AI Decision-Making in NLP.* The workshop "XAI in Action: Past, Present, and Future Applications" co-colated with NeurIPS 2023. 2 citations.

[2] Jacob K. Johnson and **Ana Marasović**. *How Much Consistency Is Your Accuracy Worth?* The Workshop on Analyzing and Interpreting Neural Networks for NLP (BlackboxNLP) co-located with EMNLP 2023. 0 citations.

[3] Markus Zopf, Teresa Botschen, Tobias Falke, Benjamin Heinzerling, **Ana Marasović**, Todor Mihaylov, S. Avinesh P.V., Eneldo Loza Mencía, J. Fürnkranz, and A. Frank. *What's Important in a Text? An Extensive Evaluation of Linguistic Annotations for Summarization.* SNAMS 2018. 10 citations.

[4] **Ana Marasović** and Anette Frank. *Multilingual Modal Sense Classification using a Convolutional Neural Network.* Workshop on Representation Learning for NLP (Repl4NLP) 2016. 24 citations.

Other

- **Ana Marasović**. *NLP's Generalization Problem, and How Researchers Are Tackling It.* The Gradient 2018.

TALKS

Keynote Talks

Challenges in Fostering (Dis)Trust in AI   Oct 2023

Workshop on "Trust Perspectives in Machine Learning, Law, and Public Policy" organized by Oct 2023
the Institute for Data, Econometrics, Algorithms, and Learning; located at Northwestern University

Contrastive Explanations of NLP Models   Nov 2021

EMNLP Workshop on "Analyzing and Interpreting Neural Networks for NLP" (BlackBoxNLP)

**Resolving Abstract Anaphors in Discourse: Uphill Battles with Neural Ranking   Jun 2018
Models and Automatic Data Extraction**

NAACL Workshop on Computational Models of Reference, Anaphora, and Coreference (CRAC)

Invited Talks

Explainability and Trust   Apr 2024

Bellairs Invitational Workshop on Contemporary, Foreseeable and Catastrophic Risks of LLMs; Barbados

Challenges in Fostering (Dis)Trust in AI   Apr 2024

University of Pennsylvania

University of Pittsburgh

"We propose `<insert an explainability method>` to increase human trust in AI": **What does this mean and how to go about it?** Apr 2023

University of Maryland

"AI" That Masters Language Could Reason about Negation Feb 2023

University of Utah Data Science Seminar

Formalizing Trust in AI May 2022

NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES)

Self-Explaining for Intuitive Interaction with AI

Carnegie Mellon University	Apr 2022
NEC Labs Europe	Apr 2022
University of Lisbon - Unbabel Seminar	Apr 2022
UC Irvine	Mar 2022
Johns Hopkins University	Feb 2022
Brown University	Feb 2022
University of Chicago	Feb 2022
Ohio State University	Feb 2022
Purdue University, CS	Feb 2022
University of Utah	Feb 2022

On Reducing Human-Authored Free-Text Explanations for Training Dec 2021

University College London

Explanation Selection Through The Lens of Free-Text & Contrastive Explanations

USC Machine Learning Symposium	Dec 2021
Georgia Tech NLP Seminar	Nov 2021
University of Sheffield Seminar	Nov 2021
Purdue University, ECE	Nov 2021

Explaining with Natural Language Oct 2021

University of Edinburgh

Reflecting on Interpretability Desiderata with Visual Commonsense Reasoning Nov 2019

University of Washington, Department of Linguistics Seminar

Teaching Machines Language Understanding Apr 2019

Allen School Women's Research Day, University of Washington

Resolving Abstract Anaphors in Discourse: Uphill Battles with Neural Ranking Models and Automatic Data Extraction

Allen Institute for AI	Oct 2018
Potsdam University, Applied Computational Linguistics	Jun 2018
Bochum University, Sprachwissenschaftliches Institut	May 2018

Confrence Presentations

- EMNLP 2022, Poster
- NAACL 2022, Poster
- EMNLP 2019, Poster
- NAACL 2018, Poster

Teaching

CS 7xxx: Mechanistic Interpretability Seminar	Spring 2025
CS 5340/6340: Natural Language Processing	Spring, Fall 2024
CS 7931: Advances in Generative AI Seminar	Fall 2024
CS 5966/6966: Local Explanations for Deep Learning Models	Fall 2023, 2022
CS 5140/6140: Data Mining	Spring 2023

Guest Lectures & Conference Tutorials & Presentations at Local Events

NAACL'24 Tutorial "Explanation in the Era of Large Language Models"	June 2024
Ingredients of Generative AI	
University of Utah, Research Challenges in Computing Seminar, *Host*: Saday Sadayappan	Nov 2023
University of Utah, DELPHI Data Science Symposium	Oct 2023
University of Utah, The Osher Lifelong Learning Institute Lunch and Learn	Oct 2023
NLP with (Almost) No Training Data	
University of Utah, MATH 4100 / COMP 5360, *Host*: Kate Isaacs	March 2023
University of Utah, CS 5340/6340, *Host*: Ellen Riloff	Nov 2022
Towards Trustworthy NLP: On Intuitive Interaction and Rigorous Validation	
University of Utah, CS 3020 Research Forum, *Host*: Kobus Van Der Merwe	Sep 2022
A Quick Tour of NLP Explainability	
University of Washington, *Host*: Yulia Tsvetkov	May 2022
Detecting Artifacts with Local Explanations	
University of British Columbia, *Host*: Vered Shwartz	March 2022
University of Washington, *Host*: Zaid Harchaoui	March 2022
Methods for Analyzing (Current) NLP Models	
University of Washington, *Host*: Noah A. Smith	March 2020

Mentoring

Utah PhD Students	
Nathan Stringham	Aug 2023 – present
Fateme Hashemi Chaleshtori	Jan 2023 – present
Utah PhD Student Collaborators	
Zhichao Xu	Dec 2023 - present
Michael Clemens → New Jersey Institute of Technology PhD	Jul 2023 – present
Atreya Ghosal	Apr 2023 – Jun 2024
Oliver Bentham	Jan 2024 - Feb 2024
Ashim Gupta	Oct 2022 – Dec 2023
Akansha Kalra	Nov 2022 – Dec 2022
Utah MS Researchers	
Adwait Shinganwade	Aug 2024 – present
Yash Lele	Aug 2024 – present
Alex Gill	Jan 2024 – present
Purbid Bambroo	Jan – May 2024; Sep 2024 – present
Rishanth Rajendhran → University of Maryland PhD	Nov 2022 – May 2024
Mugdha Abhyankar	Jan 2023 – May 2023
Utah BS Researchers	

Braden Fiedel	Aug 2024 – present
Casey Miller	Aug 2024 – present
Kishan Thambu	Aug 2024 – present
Byron Liu (co-mentoring with the PhD student **Maitrey Mehta**)	Jul 2023 – May 2024
Jordan Tan	Jul 2023 – Aug 2023
Dana Ahn	Mar 2023 – Aug 2023
Jacob Johnson → Utah PhD → Weber State University MS	Nov 2022 – May 2024

UW BS/MS Researchers

Kaiser Sun with Noah A. Smith → META Residency → JHU PhD	Mar 2020 – present
Millicent Li with Noah A. Smith → META Residency → Northeastern PhD	Mar 2020 – Aug 2021

AI2 Predoctoral Researchers

Alexis Ross with Matthew E. Peters → MIT PhD	July 2020 – June 2022

AI2 Interns (PhD students)

Abhilasha Ravichander – Carnegie Mellon University – with Matt Gardner	Jun 2021 – June 2022
Shruti Palaskar – Carnegie Mellon University	May 2021 – June 2022
Mimansa Jaiswal – University of Michigan	Sep 2021 – Dec 2021
Alexander M. Hoyle – University of Maryland – with Noah A. Smith	Jun 2020 – Nov 2020
Sarah Wiegreffe – Georgia Tech – **AI2 Intern of the Year** – with Noah A. Smith	May 2020 – Oct 2020
Zhanming Allan Jie – Singapore Uni. of Technology & Design – with Pradeep Dasigi	Jun 2019 – Aug 2019

SERVICE

Department & University Service

Faculty Upskilling in AI (U-AI) instructor and consultant	Fall 2024
AI Track Director	Jul 2023 – present
One Utah RAI Healthcare Working Group	Feb 2024 – Jul 2024
Faculty Hiring Committee	Nov 2023 – May 2024
PhD (AI) Admissions Committee	Dec 2022 – Jan 2023
Presented "Preparing a Research Statement" at 2023 KSoC Graduate Open House	Nov 2023
Organized and moderated "KSoC Women Alumni Panel" during the graduate visit days	Feb 2023

PhD Committee Member

Laura Vietz
Ashim Gupta
Maitrey Mehta
Jakob Johnson
Mattia Medina-Grespan
Zhichao Xu
Peter Hase (UNC Chapel Hill)
Zhenduo Wang
Marcos Treviso (University of Lisbon)
Prince Osei Aboagye
Tianyu Jiang

Co-organizer

TeachingNLP: Workshop on Teaching Natural Language Processing @ ACL 2024
SustaiNLP: Workshop on Simple and Efficient Natural Language Processing @ EMNLP 2021
AllenNLP 2021 Hackathon for Minoritized Communities
Host of the NLP Highlights Podcast from 2021-2022
AIPHES Scientfic Workshop 2017

Conferences I have served for (ACL, EMNLP, NAACL, EACL, ICML, AAAI) typically have an acceptance rate of up to 23%.

Senior Area Chair

    ACL 2024

    ACL 2023 (**Best Senior Area Chair**)

    EMNLP 2022

Monograph Reviewer

    Reviewer for "Synthesis Lectures in Human Language Technologies: Explainable Natural Language Processing" in 2021

Journal Reviewer

    TACL 2017 (secondary reviewer), 2021; *in 2021 TACL had an impact factor of 9.194*

Conference Reviewer

    NAACL 2024

    EMNLP 2018, 2019 (**outstanding reviewer**), 2020 (**outstanding reviewer**), 2021, 2024

    ACL 2017 (secondary reviewer), 2019, 2020, 2021

    EACL 2021, 2023

    ICML 2020

    AAAI 2020

Workshop Reviewer

    Representation Learning for NLP (Repl4NLP) @ ACL 2018, ACL 2019

    Deep Learning Approaches for Low-Resource Natural Language Processing (DeepLo) @ EMNLP 2019

    Learning from Limited Labeled Data (LLD) Workshop @ ICLR 2019

    Evaluation and Comparison of NLP Systems (EvalNLP) @ EMNLP 2020

    Insights from Negative Results in NLP @ EMNLP 2020

    Computational Approaches to Discourse (CODI) @ EMNLP 2020

    Domain Adaptation for NLP (Adapt-NLP) @ EACL 2021

Panelist

    Moderator of the panel "AI in Pop Culture vs. Reality" at the Leonardo in Nov 2024

    Panelist at the Utah Digital Health Initiative "Generative AI in Health Sciences Research" panel in Nov 2023

    Panelist at the SCI's 30th Anniversary event in April 2024

    Panelist at a career panel at 2023 EACL Student Research Workshop (SRW)

    Panelist for the screening of the movie "After Yang" for the STEM Education Film Series by Park City Film in Sep 2022

    Panelist at ACL Mentorship: "Building NLP datasets" in Aug 2022

    Panelist at SoCal ML & NLP Symposium 2021 (Networking and career paths for under-represented minorities)

    Panelist at undergraduate panel session @ ACL 2020

Moderator at Conferences

    Moderator of a meetup on "Interpretability" @ ACL 2021

    Moderator of a meetup on "Whether to do a PhD; and how to apply for and choose a PhD program" @ EMNLP 2020

    Moderator of a meetup on "Interpretability and Analysis of Models for NLP" at EMNLP 2020

    Moderator of a roundtable on "Natural Language Generation" at WiNLP @ ACL 2020

Mentoring at Conferences

    Mentor at WiNLP @ ACL 2020

    "Big sibling" at EMNLP 2019, NAACL 2019

Other Service

    Leonardo AI Advisory Committee Member, Jun 2024 – present

    Co-Lead of the AI2 Employee Resource Group (ERG) for Women and Gender Minorities in 2022

    Member of the AI2 Diversity, Equity and Inclusion Committee in 2020 and 2021

    Reader of 2020 PhD applications at the University of Washington

OPEN-SOURCE SOFTWARE AND DATASETS

- Few-Shot Self-Rationalization with Natural Language Prompts

https://github.com/allenai/feb

- Natural Language Rationales with Full-Stack Visual Reasoning: From Pixels to Semantic Frames to Commonsense Graphs
    https://github.com/allenai/visual-reasoning-rationalization
- Don't Stop Pretraining: Adapt Language Models to Domains and Tasks
    Contributor
    https://github.com/allenai/dont-stop-pretraining
- Quoref Dataset
    Contributor
    https://allenai.org/data/quoref
- SRL4ORL: Improving Opinion Role Labeling Using Multi-Task Learning With Semantic Role Labeling
    https://github.com/amarasovic/naacl-mpqa-srl4orl
- A Mention-Ranking Model for Abstract Anaphora Resolution
    https://github.com/amarasovic/neural-abstract-anaphora
    https://github.com/amarasovic/abstract-anaphora-data
- Multilingual Modal Sense Classification using a Convolutional Neural Network
    https://github.com/amarasovic/modal-sense-classification
    Dataset: http://projects.cl.uni-heidelberg.de/modals/
- Matplotlib Examples for Visualizing Paper Results
    https://github.com/amarasovic/matplotlib4papers

References

- Noah A. Smith, nasmith@cs.washington.edu
- Yejin Choi, yejin@cs.washington.edu
- Anette Frank, frank@cl.uni-heidelberg.de
- Doug Downey, dougd@allenai.org