

Ana Marasović, Curriculum Vitae, September 2021

Allen Institute for Artificial Intelligence (AI2)
University of Washington
www.anamarasovic.com
anam@allenai.org

2157 N Northlake Way, Suite 110
Paul G. Allen Center, Office 410

RESEARCH INTERESTS

Natural Language Processing, Explainable AI, Multimodality

EDUCATION

| | |
|---|-------------|
| Ph.D. in Computational Linguistics, Heidelberg University | 2015 – 2019 |
| <i>Thesis:</i> “Deep Learning With Sentiment Inference For Discourse-Oriented Opinion Analysis” | |
| <i>Advisor:</i> Anette Frank | |
| <i>Honors:</i> Magna cum laude | |
| M.Sc. in Applied Mathematics, University of Zagreb | 2013 – 2015 |
| <i>Honors:</i> Summa cum laude | |
| B.Sc. in Mathematics, University of Zagreb | 2010 – 2013 |

HONORS AND AWARDS

- ACL 2020 Best Paper Honorable Mention for “Don’t Stop Pretraining: Adapt Language Models to Domains and Tasks”
- Outstanding Reviewer Award at EMNLP 2019 and 2020
- Invited to Rising Stars EECS 2020 at UC Berkeley

PROFESSIONAL EXPERIENCE

| | |
|---|------------------------|
| Postdoctoral Research, University of Washington | January 2020 – present |
| Paul G. Allen School of Computer Science & Engineering | |
| <i>Host:</i> Noah A. Smith | |
| Postdoctoral Research, Allen Institute for Artificial Intelligence | March 2019 – present |
| <i>Mentors:</i> Noah A. Smith, Yejin Choi | |
| Research Intern, Amazon Berlin | Summer 2017 |
| <i>Topic:</i> Neural sentiment domain adaption with multiple source domains | |
| Research Intern, Poslovna Inteligencija | Summer 2014 |
| <i>Topic:</i> Churn prediction in telecom using social network analysis | |

INVITED TALKS

| | |
|--|---------------|
| Title TBD | November 2021 |
| Georgia Tech NLP Seminar | |
| Title TBD | November 2021 |
| BlacboxNLP: Analyzing and interpreting neural networks for NLP @ EMNLP | |
| Reflecting on Interpretability Desiderata with Visual Commonsense Reasoning | November 2019 |
| University of Washington, Department of Linguistics | |

| | |
|--|--------------|
| Teaching Machines Language Understanding | April 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 | October 2018 |
| NAACL Workshop on Computational Models of Reference, Anaphora, and Coreference | June 2018 |
| Potsdam University, Applied Computational Linguistics | June 2018 |
| Bochum University, Sprachwissenschaftliches Institut | May 2018 |

INVITATIONS TO RESEARCH EVENTS

| | |
|---|--------------|
| Trustworthy NLP Workshop, Google Research | October 2021 |
| An event that has aimed to bring 30 academics with 30 Googlers to the discuss trustworthy NLP | |
| Machine Learning Impact Initiative Summit, Northwestern University | June 2021 |
| I helped devise a research roadmap focused on the impacts of machine learning on people | |
| Seminar Series, Google Research | October 2020 |
| A discussion about our paper "Don't stop pretraining" | |

GUEST LECTURES

| | |
|--|-------------|
| Title TBD | Spring 2022 |
| University of Washington, <i>Host</i> : Yulia Tsvetkov | |
| Title TBD | Autumn 2021 |
| University of Washington, <i>Host</i> : Zaid Harchaoui | |
| Methods for Analyzing (Current) NLP Models | March 2020 |
| University of Washington, <i>Host</i> : Noah A. Smith | |

MENTORING

Mentored 9 students (1 BS, 2 MS, and 6 PhD); 5 submitted papers to top-tier conferences (5 accepted to date)

| | |
|--|---------------------------|
| AI2 Predoctoral Researchers | |
| Alexis Ross with Matthew E. Peters | July 2020 – present |
| AI2 Interns (PhD students) | |
| Mimansa Jaiswal – University of Michigan | September 2021 – present |
| Abhilasha Ravichander – Carnegie Mellon University – with Matt Gardner | June 2021 – present |
| Shruti Palaskar – Carnegie Mellon University | May 2021 – present |
| Alexander M. Hoyle – University of Maryland – with Noah A. Smith | June 2020 – November 2020 |
| Sarah Wiegrefe – Georgia Tech – AI2 Intern of the Year – with Noah A. Smith | May 2020 – October 2020 |
| Zhanming Allan Jie – Singapore Uni. of Technology & Design – with Pradeep Dasigi | June 2019 – August 2019 |
| UW BS/MS Students | |
| Kaiser Sun with Noah A. Smith | March 2020 – present |
| Millicent Li with Noah A. Smith | March 2020 – August 2021 |

SERVICE

Co-organizer

SustainNLP: Workshop on Simple and Efficient Natural Language Processing @ EMNLP 2021
AllenNLP 2021 Hackathon for Marginalized Communities
AIPHES Scientific Workshop 2017

Monograph Reviewer

Synthesis Lectures in Human Language Technologies: *Explainable Natural Language Processing* 2021

Journal Reviewer

Transactions of ACL 2017 (secondary reviewer)

Conference Reviewer

EMNLP 2018, 2019 (**outstanding reviewer**), 2020 (**outstanding reviewer**), 2021
ACL 2017 (secondary reviewer), 2019, 2020, 2021
EACL 2021
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

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 Community Service

Host of the NLP Highlights Podcast in 2021
Member of the AI2 Diversity, Equity and Inclusion Committee in 2020 and 2021
Reader of 2020 PhD applications at the University of Washington

PUBLICATIONS

Peer-Reviewed Publications

- Sarah Wiegrefe, **Ana Marasović**, and Noah A. Smith. *Measuring Association Between Labels and Free-Text Rationales*. EMNLP 2021.
- Jesse Dodge, Maarten Sap, **Ana Marasović**, William Agnew, Gabriel Ilharco, Dirk Groeneveld, and Matt Gardner. *Documenting the English Colossal Clean Crawled Corpus*. EMNLP 2021.

- Sarah Wiegrefe and **Ana Marasović (equal contributions)**. *Teach Me to Explain: A Review of Datasets for Explainable NLP*. NeurIPS 2021.
- Kaiser Sun and **Ana Marasović**. *Effective Attention Sheds Light On Interpretability*. Findings of ACL 2021.
- Alexis Ross, **Ana Marasović**, and Matthew E. Peters. *Explaining NLP Models via Minimal Contrastive Editing (MiCE)*. Findings of ACL 2021.
- Alexander M. Hoyle, **Ana Marasović**, and Noah A. Smith. *Promoting Graph Awareness in Linearized Graph-to-Text Generation*. Findings of ACL 2021.
- 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.
- **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.
- 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 Demo 2020.
- 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. **(Received an Honorable Mention for Best Paper)**
- 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.
- **Ana Marasović** and Anette Frank. *SRL4ORL: Improving Opinion Role Labeling Using Multi-Task Learning with Semantic Role Labeling*. NAACL 2018.
- 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.
- **Ana Marasović**, Leo Born, Juri Opitz, and Anette Frank. *A Mention-Ranking Model for Abstract Anaphora Resolution*. EMNLP 2017.
- **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.
- **Ana Marasović** and Anette Frank. *Multilingual Modal Sense Classification using a Convolutional Neural Network*. Workshop on Representation Learning for NLP (Repl4NLP) 2016.

Blog Posts

- **Ana Marasović**. *NLP's Generalization Problem, and How Researchers Are Tackling It*. The Gradient 2018.
- **Ana Marasović**. *BERT-base Transformer Forward Pass*. Available at <https://www.anamarasovic.com/bert-forward/>.

OPEN-SOURCE SOFTWARE AND DATASETS

- 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>