ORION WELLER

Baltimore Maryland, 21218

http://orionweller.github.io ⋄ wellerorion@gmail.com

EDUCATION

Johns Hopkins University   2021 – Present
Ph.D. in Computer Science   GPA: 4.0/4.0
Center for Language and Speech Processing
Advisors: Benjamin Van Durme and Dawn Lawrie

Brigham Young University   2017 – 2021
Bachelor of Science, Summa Cum Laude (Top 1%)   GPA: 4.0/4.0
Majors: Computer Science, Statistics; Minor in Mathematics

AWARDS

- NSF Graduate Research Fellowship, 2022
- DoD NDSEG Fellowship (declined), 2022
- CRA Outstanding Undergraduate Researcher in North America Award, 2021
- Fulbright Fellowship Semi-Finalist, 2021
- NSF Graduate Research Fellowship Honorable Mention, 2021
- Goldwater Scholarship, 2020

RESEARCH EXPERIENCE

Semantic Scholar, Allen Institute for Artificial Intelligence (AI2)   Summer 2023
Research Intern, with Luca Soldaini, Kyle Lo, and Arman Cohan

Apple AI/ML Research   Summers 2020, 2021
Machine Learning Intern, with Matthias Sperber

AllenNLP, Allen Institute for Artificial Intelligence (AI2)   Winter 2020
Research Intern, with Matt Gardner and Matt Peters

Brigham Young University   2018 – 2021
Undergraduate Research Assistant, with Kevin Seppi and Quinn Snell

Adobe   Fall 2018
Student Researcher, with Ilya Reznik and Chris Challis

PUBLICATIONS

[1] **Orion Weller**, Dawn Lawrie, Benjamin Van Durme, "NevIR: Negation in Neural Information Retrieval," in *European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[2] **Orion Weller**, Kyle Lo, David Wadden, Dawn Lawrie, Benjamin Van Durme, Arman Cohan, Luca Soldaini, "When do Generative Query and Document Expansions Fail? A Comprehensive Study Across Methods, Retrievers, and Datasets," in *Findings of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[3] **Orion Weller**, Marc Marone, Nathaniel Weir, Dawn Lawrie, Daniel Khashabi, Benjamin Van Durme, ""According to..." Prompting Language Models Improves Quoting from Pre-Training Data," in *European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[4] **Orion Weller**, Aleem Khan, Nathaniel Weir, Dawn J Lawrie, Benjamin Van Durme, "Defending Against Poisoning Attacks in Open-Domain Question Answering," in *Findings of the European Chapter of the Association for Computational Linguistics (EACL)*, 2024.

[5] Kangda Wei, Dawn J Lawrie, Benjamin Van Durme, Yunmo Chen, **Orion Weller,** "When Do Decompositions Help for Machine Reading?" In *Empirical Methods in Natural Language Processing (EMNLP)*, 2023.

[8] **Orion Weller**, Marc Marone, Vladimir Braverman, Dawn Lawrie, Benjamin Van Durme, "Pretrained Models for Multilingual Federated Learning," in *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.

[9] **Orion Weller**, Kevin Seppi, Matt Gardner, "When to Use Multi-Task Learning vs Intermediate Fine-Tuning for Pre-Trained Encoder Transfer Learning," in *Association of Computational Linguistics (ACL)*, 2022.

[10] **Orion Weller**, Matthias Sperber, Telmo Pires, Hendra Setiawan, Christian Gollan, Dominic Telaar, Matthias Paulik, "Speech Translation for Code-Switched Speech," in *Findings of the Association of Computational Linguistics (ACL)*, 2022.

[11] Wilson Fearn, **Orion Weller**, Kevin Seppi, "Exploring the Relationship Between Algorithm Performance, Vocabulary, and Run-Time in Text Classification," in *North American Chapter of the Association for Computational Linguistics (NAACL)*, 2021.

[12] **Orion Weller**, Matthias Sperber, Christian Gollan, Joris Kluivers, "Streaming Joint Speech Translation and Transcription," in *European Chapter of the Association for Computational Linguistics (EACL)*, 2021.

[13] **Orion Weller**, Nicholas Lourie, Matt Gardner, Matthew Peters, "Learning from Task Descriptions," in *Empirical Methods in Natural Language Processing (EMNLP)*, 2020.

[14] **Orion Weller**, Jordan Hildebrandt, Ilya Reznik, Christopher Challis, E. Shannon Tass, Quinn Snell, Kevin Seppi, "You Don't Have Time to Read This: an Exploration of Document Level Reading Time Prediction," in *Association of Computational Linguistics (ACL)*, 2020.

[15] **Orion Weller**, Nancy Fulda, Kevin Seppi, "Can Humor Prediction Datasets be used for Humor Generation? Humorous Headline Generation via Style Transfer," in *Second Workshop on Figurative Language Processing @ ACL*, 2020.

[16] **Orion Weller**, Luke Sagers, Carl Hanson, Quinn Snell, Michael Barnes, Shannon Tass, "Predicting Suicidal Thoughts and Behavior among Adolescents using the Risk and Protective Factor Framework: a Large-Scale Machine Learning Approach," in *Plos One*, 2021.

[17] **Orion Weller**, Kevin Seppi, "The rJokes Dataset: a Large Scale Humor Collection," in *Language Resources and Evaluation (LREC)*, 2020.

[18] **Orion Weller**, Kevin Seppi, "Humor Detection: a Transformer Gets the Last Laugh," in *Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

PREPRINTS

[6] James Mayfield, Eugene Yang, Dawn Lawrie, Samuel Barham, **Orion Weller**, Marc Mason, Suraj Nair, Scott Miller, "Synthetic Cross-language Information Retrieval Training Data," *ArXiv*, 2023.

[7] Samuel Barham, **Orion Weller**, Michelle Yuan, Kenton Murray, Mahsa Yarmohammadi, Zhengping Jiang, Siddharth Vashishtha, Alexander Martin, Anqi Liu, Aaron Steven White, "MegaWika: Millions of reports and their sources across 50 diverse languages," *ArXiv*, 2023.

ENGINEERING EXPERIENCE

Qualtrics   Apr 2019 - Aug 2019
Software Engineer Intern, responsible for prototyping on the Data Pipeline team

Digi International   Oct 2017 - Jan 2019
Student Software Engineer, worked on building firmware for Zigbee wireless networking modules

SKILLS

Languages	English (native), Portuguese (intermediate), French (basic)
Programming Languages	Python, C++, R, Javascript, Golang, Java, OCaml, SQL
Frameworks & Tools	PyTorch, Linux, Mechanical Turk, Django, VueJS

SERVICE

Workshop Organizer

- Scholarly Document Processing (SDP) Workshop at ACL, 2024
- Mid-Atlantic Student Colloquium on Speech, Language and Learning, 2022-2024

Teaching

- EN.500.111: Reasoning with ChatGPT, Johns Hopkins University, Fall 2023
- After School Coding Instructor for Elementary School with Code in the Schools, 2021-2022

Program Committee Member

- 2023: ACL, EMNLP, ACL Rolling Review
- 2022: ACL Rolling Review
- 2021: ACL Rolling Review, ACL, EMNLP, NAACL (secondary)
- 2020: Meta-Learning@NeurIPS