Emma Strubell

strubell@cmu.edu
http://cs.cmu.edu/~estrubel

Education

University of Massachusetts Amherst 2012 — 2019
Ph.D. in Computer Science
M.S. in Computer Science, 2012 — 2015

University of Maine 2008 — 2012
B.S. in Computer Science, *summa cum laude*
Minor: Mathematics & Statistics

Research & Professional Experience

Assistant Professor September 2020
Language Technologies Institute, Carnegie Mellon University

Visiting Researcher September 2019 — present
Facebook AI Research

Research Assistant, University of Massachusetts Amherst September 2012 — May 2019

Research Intern, Google AI Language May — August 2017

Research Intern, Amazon Alexa Science June — September 2016

Research Assistant, University of Maine May 2011 — August 2012

Awards

- Best Long Paper Award, EMNLP 2018
- IBM PhD Fellowship Award, 2017–2018
- Yahoo Outstanding Accomplishments in Search and Mining Award, 2016
- Grace Hopper Conference Scholarship Grant (21% acceptance), 2015
- EMC CRA-W Grad Cohort Scholarship Award, 2015
- Outstanding Paper Award, ACL 2015
- UMaine College of Liberal Arts & Sciences Undergraduate Research Fellowship, 2011–2012

Publications

[1] **Emma Strubell**, Ananya Ganesh, and Andrew McCallum. Energy and Policy Considerations for Modern Deep Learning. In *Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI invited abstract)*, New York, USA, February 2020.

[2] **Emma Strubell**, Ananya Ganesh, and Andrew McCallum. Energy and Policy Considerations for Deep Learning in NLP. In *Annual Meeting of the Association for Computational Linguistics (ACL short)*, Florence, Italy, July 2019.

[3] Sheshera Mysore, Zach Jensen, Edward Kim, Kevin Huang, Haw-Shiuan Chang, **Emma Strubell**, Jeffrey Flanigan, Andrew McCallum, and Elsa Olivetti. The Materials Science Procedural Text Corpus: Annotating Materials Synthesis Procedures with Shallow Semantic Structures. In *LAW XIII 2019: The 13th Linguistic Annotation Workshop (ACL WS)*, Florence, Italy, July 2019.

[4] Edward Kim, Zach Jensen, Alexander van Grootel, Kevin Huang, Matthew Staib, Sheshera Mysore, Haw-Shiuan Chang, **Emma Strubell**, Andrew McCallum, Stefanie Jegelka, and Elsa Olivetti. Inorganic Materials Synthesis Planning with Literature-Trained Neural Networks. In *Journal of Chemical Information and Modeling, in press*, 2020.

[5] **Emma Strubell**, Patrick Verga, Daniel Andor, David Weiss, and Andrew McCallum. Linguistically-Informed Self-Attention for Semantic Role Labeling. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Brussels, Belgium, October 2018. ***Best long paper award.***

[6] **Emma Strubell** and Andrew McCallum. Syntax Helps ELMo Understand Semantics: Is Syntax Still Relevant in a Deep Neural Architecture for SRL? In *Proceedings of the Workshop on the Relevance of Linguistic Structure in Neural Architectures for NLP (ACL WS)*, pages 19–27. Association for Computational Linguistics, 2018.

[7] Patrick Verga, **Emma Strubell**, and Andrew McCallum. Simultaneously Self-attending to All Mentions for Full-Abstract Biological Relation Extraction. In *Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)*, New Orleans, Louisiana, June 2018.

[8] Vittorio Perera, Tagyoung Chung, Thomas Kollar, and **Emma Strubell**. Multi-Task Learning For Parsing The Alexa Meaning Representation Language. In *Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, New Orleans, Louisiana, February 2018.

[9] Sheshera Mysore, Edward Kim, **Emma Strubell**, Ao Liu, Haw-Shiuan Chang, Srikrishna Kompella, Kevin Huang, Andrew McCallum, and Elsa Olivetti. Automatically Extracting Action Graphs From Materials Science Synthesis Procedures. In *NIPS Workshop on Machine Learning for Molecules and Materials*, Long Beach, California, December 2017. ***Spotlight talk.***

[10] Patrick Verga, **Emma Strubell**, Ofer Shai, and Andrew McCallum. Attending to All Mention Pairs for Full Abstract Biological Relation Extraction. In *6th Workshop on Automated Knowledge Base Construction (AKBC)*, Long Beach, California, December 2017.

[11] **Emma Strubell**, Patrick Verga, David Belanger, and Andrew McCallum. Fast and Accurate Entity Recognition with Iterated Dilated Convolutions. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Copenhagen, Denmark, September 2017.

[12] **Emma Strubell** and Andrew McCallum. Dependency Parsing with Dilated Iterated Graph CNNs. In *2nd Workshop on Structured Prediction for Natural Language Processing (EMNLP WS)*, Copenhagen, Denmark, September 2017.

[13] Edward Kim, Kevin Huang, Alex Tomala, Sara Matthews, **Emma Strubell**, Adam Saunders, Andrew McCallum, and Elsa Olivetti. Machine-learned and codified synthesis parameters of oxide materials. *Nature Scientific Data*, 4, 2017.

[14] David E. Hiebeler, Andrew Audibert, **Emma Strubell**, and Isaac J. Michaud. An epidemiological model of internet worms with hierarchical dispersal and spatial clustering of hosts. *Journal of Theoretical Biology*, 418:8–15, 2017.

[15] Haw-Shiuan Chang, Abdurrahman Munir, Ao Liu, Johnny Tian-Zheng Wei, Aaron Traylor, Ajay Nagesh, Nicholas Monath, Patrick Verga, **Emma Strubell**, and Andrew McCallum. Extracting Multilingual Relations under Limited Resources: TAC 2016 Cold-Start KB construction and Slot-Filling using Compositional Universal Schema. In *Text Analysis Conference (Knowledge Base Population Track) '16 Workshop (TAC KBP)*, Gaithersburg, Maryland, USA, November 2016.

[16] Patrick Verga, David Belanger, **Emma Strubell**, Benjamin Roth, and Andrew McCallum. Multilingual Relation Extraction using Compositional Universal Schema. In *Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)*, San Diego, California, June 2016.

[17] **Emma Strubell**, Luke Vilnis, Kate Silverstein, and Andrew McCallum. Learning Dynamic Feature Selection for Fast Sequential Prediction. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, Beijing, China, July 2015. ***Outstanding paper award.***

[18] Benjamin Roth, Nicholas Monath, David Belanger, **Emma Strubell**, Patrick Verga, and Andrew McCallum. Building Knowledge Bases with Universal Schema: Cold Start and Slot-Filling Approaches. In *Text Analysis Conference (Knowledge Base Population Track) '15 Workshop (TAC KBP)*, Gaithersburg, Maryland, USA, November 2015.

[19] Benjamin Roth, **Emma Strubell**, Katherine Silverstein, and Andrew McCallum. Minimally Supervised Event Argument Extraction using Universal Schema. In *4th Workshop on Automated Knowledge Base Construction (AKBC)*, NIPS '14, Montreal, Quebec, Canada, December 2014.

[20] **Emma Strubell**, Luke Vilnis, and Andrew McCallum. Training for Fast Sequential Prediction Using Dynamic Feature Selection. In *NIPS Workshop on Modern Machine Learning and NLP (NIPS WS)*, Montreal, Quebec, Canada, December 2014.

[21] Benjamin Roth, **Emma Strubell**, John Sullivan, Lakshmi Vikraman, Katherine Silverstein, and Andrew McCallum. Universal Schema for Slot-Filling, Cold-Start KBP and Event Argument Extraction: UMassIESL at TAC KBP 2014. In *Text Analysis Conference (Knowledge Base Population Track) '14 Workshop (TAC KBP)*, Gaithersburg, Maryland, USA, November 2014.

Teaching

Team Mentor, University of Massachusetts Amherst — January — May 2017, 2018
Course: MS Industry Mentorship Program with Chan Zuckerberg Initiative

Guest Lecturer, University of Massachusetts Amherst — October 2017
Course: Neural Networks with Erik Learned-Miller

Teaching Assistant, University of Massachusetts Amherst — September — December 2012
Course: Introduction to Computation with Neil Immerman

Teaching Assistant, University of Maine — January — May 2012
Course: Automata, Computability and Languages with Larry Latour

Service & Outreach

- Panelist, *Succeeding in Graduate School: Advice from Outstanding Students in the Sciences*, 2017.
- Co-organizer and mentor, Cross-cultural Graduate Peer Mentoring Program ($1000 Welcoming the World to Amherst Grant), 2017
- Co-organizer, CS Women Travel Grant Program ($5000 Women for UMass Grant), 2016–2017.
- Co-organizer, CS Women Technical Skills Workshops ($1000 NCWIT Seed Grant), 2015–2016.
- Co-chair and treasurer, UMass CS Women, 2015–2017.
- Senator, UMass Graduate Student Senate, 2014–2015.
- Mentor, Girls Inc. Eureka! Workshop (Programming in Scratch), 2014, 2015.
- Mentor, CAITE Women in Engineering & Computing Career Day, 2013, 2015.
- Steward, UMass Graduate Employee Organization (UAW 2322), 2013–2014.
- Reviewer, NCWIT Aspirations in Computing Scholarship, 2013, 2014.
- Panelist, *UMaine Outstanding Students in the College of Liberal Arts and Sciences*, 2012.
- Program Committees: ACL 2017, 2018 (top reviewer); EMNLP 2015, 2017, 2018; SCiL 2019; ICWSM 2019; PLDI 2016.
- Student volunteer, NAACL 2018.

Professional Development

- WW2A Cross-Cultural Mentoring Workshop, 2016.
- CIRTL Research Mentor Training Course, 2016.
- UMass OPD Workshop on Engaging Students in Effective Discussions and Active Learning, 2015.

Mentoring

- Juhi Shah (2018). Undergraduate independent study at Mt. Holyoke College. *Automatically detecting fake news using neural networks.*
- Ananya Ganesh (2018). Masters project at University of Massachusetts Amherst. *Improved representation learning for semantic role labeling.*
- Sheshera Mysore (2017). Masters project at University of Massachusetts Amherst. *Kernelized matrix completion for predicate-argument schema induction and extraction.*
- Aditya Shastry (2017). Masters project at University of Massachusetts Amherst. *Automatic header field extraction from research articles.*
- Molly McMahon (2016). Masters project at University of Massachusetts Amherst. *Automatic citation field extraction from research articles*
- Abdurrahman Munir (2016). Louis Stokes Alliance for Minority Participation (LSAMP) undergraduate research fellow at University of Massachusetts Amherst. *Character-level modeling for Arabic tokenization and named entity recognition.*
- Katherine Silverstein (2015). Undergraduate independent study at University of Massachusetts Amherst. *Fast and accurate models for named entity recognition.*