Jonathan David Louis May

Contact Information

4676 Admiralty Way, Suite 1001
Marina Del Rey, CA 90292

Phone: (310) 448-9157
Email: jonmay@isi.edu

Work Experience

USC Information Sciences Institute, Marina Del Rey, California

Research Associate Professor, Department of Computer Science *July 2022 – present*

Research Assistant Professor, Department of Computer Science *July 2016 – June 2022*

Classes:

CSCI 662, Advanced Natural Language Processing	*Fall 2019–25*
CSCI 544, Applied Natural Language Processing	*Fall 2017–18*

Principal Scientist, Artificial Intelligence Division *July 2021 – present*

Lead Scientist, Artificial Intelligence Division *July 2018 – June 2021*

Computer Scientist, Intelligent Systems Division *May 2014 – June 2018*

Programs:

DARPA-CODORD (Co-PI)	2025–
ARPA-H CARE (Co-PI)	2024–
MITRE LCNLP (PI)	2024–2025
DARPA-FACT AIE (PI)	2024–2025
Army SBIR (PI, sub to Inferlink)	2024–
IARPA-HIATUS	2022–
DARPA-KMASS (Co-PI)	2022–2025
DARPA-SHADE AIE (PI, sub to UMD)	2022–2023
DARPA-Civil Sanctuary AIE (PI)	2021–2023
DARPA-INCAS	2021–2023
DARPA-REMATH	2020–2022
Amazon Alexa Prize (PI)	2020–2021
NIH-EVIDENCEEXTR (PI)	2020–2021
IARPA-BETTER	2019–2023
DARPA-KAIROS (Co-PI)	2019–2023
DARPA-MCS	2019–2022
DARPA-LWLL (Co-PI)	2019–2023
DARPA-ASED	2019–2022
IARPA-MATERIAL (Co-PI)	2017–2021
DARPA-LORELEI (Co-PI, PI)	2016–2022
MITRE DEEPLANG (PI)	2015
DARPA-BOLT	2014–2015
IARPA-Metaphor	2014–2015

Amazon, Seattle, Washington

Consultant, Search *Oct. 2021 – Sep. 2022*

Scholar, Search *Oct. 2022 – present*

SDL Language Weaver, Los Angeles, California

Research Scientist *Aug. 2010 – May 2014*

Government Projects Technical Lead:

FP7-FAUST (Feedback Analysis for User-Adaptive Statistical Translation)
DARPA-BOLT (Broad Operation Language Translation)
CTTSO/TSWG-MTIL (Machine Translation for Informal Language)

BBN Technologies, Cambridge, Massachusetts

Staff Scientist, Natural Language Processing group *Sep. 2001 – Jun. 2004*

Education

University of Southern California, Los Angeles, California

Ph.D, Computer Science. *Sep. 2004 – Aug. 2010*

Thesis: "Weighted Tree Automata and Transducers for Syntactic Natural Language Processing" (Committee: Kevin Knight (chair), Daniel Marcu, David Chiang, Sven Koenig, Shri Narayanan, Fernando Pereira)

University of Pennsylvania, Philadelphia, Pennsylvania

M.S.E., Computer and Information Science *Feb. 2000–May 2001*

B.S.E., Computer Science Engineering. *Sep. 1997–May 2001*

Grants and Gifts Awarded

"Can Language Models Hide What They Know?" Jonathan May, PI. 2025–2026. Open Philanthropy.

"CHAT-HPV: Conversational Health AI Tool for HPV." Mona Guo (Keck), PI. 2025–2027. The Physicians Foundation.

"Machine Translation for Indo-pacific Low Resource Languages." Jonathan May, PI; sub to Inferlink. 2025. SBIR Phase 2.

"ALICE: Assessing LLM Integrity for Clinical Engagement." Marjorie Freedman, PI. Elizabeth Boschee, Jonathan May, Souti Chattopadhyay, Co-PIs. ChLA (Michael Neely, Matt Keefer, Elizabeth Burgener, Maria del Carmen Reyes, Kathryn Smith, Grace Anaya), Subs. 2024. ARPA-H CARE.

"Long-Context Natural Language Policy Understanding (LCNLP)." Jonathan May, PI. Xuezhe Ma, Co-PI. 2024. MITRE.

"Complementing Human Intelligence to Recognize Opponent Narratives (CHIRON)." Jonathan May, PI. Souti Chattopadhyay, Co-PI. UMD (Jordan Boyd-Graber) and University of Sydney (Jonathan Kummerfeld), Subs. 2024–2025. DARPA FACT AIE.

"Machine Translation for Indo-pacific Low Resource Languages." Jonathan May, PI; sub to Inferlink. 2024. SBIR Phase 1.

"Detecting Bias in the Law." Jonathan May and Jonathan Choi (Gould Law School), Co-PIs. 2024. ISI Research Award.

"Multi-document Newsworthy Event Monitoring and Forecasting." Muhao Chen and Jonathan May, Co-PIs. 2023. ISI Research Award.

"Bloomberg Fellowship Support Award." Jonathan May, PI. 2021. Bloomberg.

"ALLAN: Agents Learning Lying And Negotiation." Jonathan May, PI; sub to UMD.

"DARMA: Dialogue Agent for Reducing Malicious Acts." Jonathan May, PI. Kristina Lerman, Co-PI.

"UPSCALE: Universal Population Segmentation and Characterization Algorithms for OnLine Environments." Emilio Ferrara, PI, Allessandro Flamini and Jacob Shapiro, Co-PIs.

"SMELLCPS: Symbolic Math Expressions from Low-level Logic in Cyber-physical Systems." Luis Garcia, PI, Christophe Hauser, Co-PI. DARPA.

"Viola: Neural Generation with Improvised Dialogues and Common Sense Reasoning." Jonathan May, PI. 2020–2021. Amazon.

"Bloomberg Fellowship Support Award." Jonathan May, PI. 2020. Bloomberg.

"Universal Translators for Asylum Seekers at the Border." Jonathan May, PI. 2020. ISI Research Award.

"CLEAR: Cross-Lingual Event & Argument Retrieval." Elizabeth Boschee, PI, Scott Miller, Co-PI. 2019–2022. IARPA.

"LESTAT: Discovering Schemas from Diverse Data." Ralph Weischedel, PI, Marjorie Freedman, Jonathan May, Co-PIs. 2019–2023. DARPA.

"Discovering Common Sense from Video, Images, Text and Knowledge Bases." Ralph Weischedel, PI, Marjorie Freedman, Co-PI. 2019–2022. DARPA.

"CORAL: Combined Representations for Adept Learning." Aram Galstyan (2019–2020), Wael Abd-Almageed (2021–2022), PI, Wael Abd-Almageed (2019–2020), Jonathan May (2021–2022), Pedro Szekely, Ralph Weischedel, Co-PIs. 2019–2022. DARPA.

"Summarization and domain-Adaptive Retrieval Across Languages." Scott Miller, PI. Jonathan May, Elizabeth Boschee, Co-PIs. 2017–2021. IARPA.

"Exploiting Language Information for Situational Awareness." Kevin Knight (2015–2018), Jonathan May (2018–2021), PI. Daniel Marcu (2015–2017), Jonathan May (2017–2018), Co-PI. 2015–2021.

"DEEPLANG". Jonathan May, PI. 7/27/15–9/30/15. MITRE.

Patents

"Efficient online domain adaptation".

Felix Hieber and Jonathan May. US 9,213,694. Awarded December 15, 2015.

"Personalized machine translation via online adaptation".

Daniel Marcu and Jonathan May. US 9,152,622. Awarded October 6, 2015.

"Modification of annotated bilingual segment pairs in syntax-based machine translation".

Wei Wang and Jonathan May and Kevin Knight. US 8,825,466. Awarded September 2, 2014.

"Systems and methods for tuning parameters in statistical machine translation".

Mark Hopkins and Jonathan May. US 8,694,303. Awarded April 8, 2014.

Students Supervised and Mentored

Ph.D Students

Xusen Yin	(from Kevin Knight in 2018) Ph.D. Summer 2021;	2016–2021
Nada Aldarrab	(from Kevin Knight in 2018) Ph.D Fall 2022;	2017–2022
Thamme Gowda	Ph.D Summer 2022;	2018–2022
Meryem M'Hamdi	Ph.D Winter 2024;	2019–2024
Mozhdeh Gheini	Ph.D Winter 2025;	2019–2025
Alex Spangher	Ph.D Summer 2025; (co-advised with Emilio Ferrara)	2019–2025
Justin Cho		2020–
Katy Felkner		2020–
Jacob Bremerman	(co-advised with Xiang Ren)	2021–
Tenghao Huang	(from Muhao Chen in 2023)	2023–
Dhananjay Ashok		2024–
Yanze Wang		2024–
Zhejian Zhou		2025–
Ruth-Ann Armstrong		2025–
Ryan Lee		2025–

MS Students

Harshavardhan Alimi		2025–
Ryan Lee	(co-advised with Xuezhe Ma) MS Spring 2025;	2024–2025
Joseph Wang	MS Spring 2025;	2024–2025
Apoorva Sharma	(co-advised with Kristina Lerman) MS Spring 2023;	2022–2023
Darpan Jain	MS Spring 2023;	2021–2023
Kartik Shenoy	MS Spring 2022;	2021–2022
Shuai Liu	MS Spring 2021;	2020–2021
David Jiang-Gorsline	MS Spring 2020;	2019–2020
Tian Xie	MS Spring 2018;	2017–2018
Jibiao Shen	MS Fall 2018;	2017–2018
Mozhdeh Gheini	MS Fall 2018;	2017–2018
Bashar Alhafni	(co-advised with Nanyun Peng) MS Spring 2019;	2018–2019

Undergraduate Students

Timothy Wang		2022–2023
Justin Cho	(HKUST) BS Spring 2019;	2019

Summer Interns

2025: Tanner Spendlove (BYU)
2024: Dhananjay Ashok (Toronto), Yanze Wang (USC), Joseph Wang (USC)
2022: Taiwei Shi (Georgia Tech) (with Xuezhe Ma), Jonne Saleva (Brandeis), Jacqueline He (Princeton) (with Xuezhe Ma)
2021: Shira Wein (Georgetown), Zeyu Liu (University of Washington), Shanxiu He (UCLA) (with Xuezhe Ma and Muhao Chen)

2020: Omar Shaikh (Georgia Tech), Ugur Yavuz (Dartmouth College), Weiqiu You (University of Pennsylvania), Naitian Zhou (University of Michigan)
2019: Justin Cho (Hong Kong Univ. of Science and Technology), Denis Emelin (Edinburgh), Zhifeng Hu (Fudan Univ.) Angelina McMillan-Major (UW), Prince Wang (UCSB), Shufan Wang (UMass) (with Nanyun Peng)
2018: Ronald Cardenas (Charles University in Prague), Mozhdeh Gheini (USC), Xiaolei Huang (Univ. of Colorado), Allison Limke (Wartburg College), James Mullenbach (Georgia Tech), Xinyu Wang (CMU) (with Nanyun Peng)
2017: Yining Chen (Dartmouth), Leon Cheung (UCSD), Sorcha Gilroy (Edinburgh), Nelson Liu (UW), Alexandra (Sasha) Mayn (Carleton College) (with Kevin Knight)
2016: Nada Aldarrab (USC), Angeliki Laziradou (U. Trento), Xiang Li (U. Chicago), Sabrina Mielke (Dresden Univ. Technology), Ke Tran (U. Amsterdam) (with Kevin Knight and Daniel Marcu)

Publications

Refereed Conference Papers (95)

Dhananjay Ashok and Jonathan May. "Language Models Can Predict Their Own Behavior". *Proc. NeurIPS* 2025. To Appear.

Tenghao Huang, Sihao Chen, Muhao Chen, Jonathan May, Longqi Yang, Mengting Wan, Pei Zhou. "Teaching Language Models To Gather Information Proactively". *Findings of EMNLP* 2025. To Appear.

Dhananjay Ashok, Ashutosh Chaubey, Hirona J. Arai, Jonathan May, Jesse Thomason. "Can VLMs Recall Factual Associations From Visual References?" *Findings of EMNLP* 2025. To Appear.

Tenghao Huang, Dong Hee Lee, John Sweeney, Jiatong Shi, Emily Steliotes, Matthew Lange, Jonathan May, Muhao Chen. "FoodPuzzle: Toward Developing Large Language Models as Autonomous Flavor Scientists". *Proc. KDD* (19%) 2025.

Tenghao Huang, Kinjal Basu, Ibrahim Abdelaziz, Pavan Kapanipathi, Jonathan May, Muhao Chen. "R2D2: Remembering, Replaying and Dynamic Decision Making with a Reflective Agentic Memory". *Proc. ACL* (20.3%) 2025.

Dhananjay Ashok, Jonathan May. "A Little Human Data Goes A Long Way". *Proc. ACL* (20.3%) 2025.

Alexander Spangher, Michael Lu, Sriya Kalyan, Hyundong Justin Cho, Tenghao Huang, Weiyan Shi, Jonathan May. "NewsInterview: a Dataset and a Playground to Evaluate LLMs' Ground Gap via Informational Interviews". *Proc. ACL* (20.3%) 2025.

Hyundong Justin Cho, Spencer Lin, Tejas Srinivasan, Michael Saxon, Deuksin Kwon, Natali T. Chavez, Jonathan May. "Can Vision Language Models Understand Mimed Actions?". *Findings of ACL* (37%) 2025.

Abraham Israeli, Shuai Liu, Jonathan May, David Jurgens. "The Million Authors Corpus: A Cross-Lingual and Cross-Domain Wikipedia Dataset for Authorship Verification". *Findings of ACL* (37%) 2025.

Wichayaporn Wongkamjan, Yanze Wang, Feng Gu, Denis Peskoff, Jonathan K. Kummerfeld, Jonathan May, Jordan Lee Boyd-Graber. " Should I Trust You? Detecting Deception in Negotiations using Counterfactual RL". *Findings of ACL* (37%) 2025.

Feng Gu, Wichayaporn Wongkamjan, Jordan Lee Boyd-Graber, Jonathan K. Kummerfeld, Denis Peskoff, Jonathan May. "Personalized Help for Optimizing Low-Skilled Users' Strategy". *Proc. NAACL* (22%), 2025.

Shuai Liu, Jonathan May. "Style Transfer with Multi-iteration Preference Optimization". *Proc. NAACL* (22%), 2025.

Hyundong Justin Cho, Karishma Sharma, Nicolaas Paul Jedema, Leonardo F. R. Ribeiro, Jonathan May, Alessandro Moschitti. "Tuning-Free Personalized Alignment via Trial-Error-Explain In-Context Learning". *Findings of NAACL* (37%), 2025.

Shanxiu He, Mutasem Al-Darabsah, Suraj Nair, Jonathan May, Tarun Agarwal, Tao Yang and Choon Hui Teo. "Token Pruning Optimization for Efficient Dense Retrieval with Multi-Vector Representations". *Proc. ECIR* (23%), 2025.

Mengtian Guo, Mutasem Al-Darabsah, Choon Hui Teo, Jonathan May, Tarun Agarwal, and Rahul Bhagat. "Learning to Rewrite Negation Queries in Product Search". *Prof. ICCL Industry Track* (42%), 2025.

Xuezhe Ma, Xiaomeng Yang, Wenhan Xiong, Beidi Chen, LILI YU, Hao Zhang, Jonathan May, Luke Zettlemoyer, Omer Levy, and Chunting Zhou. "Megalodon: Efficient LLM Pretraining and Inference with Unlimited Context Length". *Proc. NeurIPS* (25.8%), 2024.

Alexander Spangher, James Youn, Matt DeButts, Nanyun Peng, Emilio Ferrara, and Jonathan May. "Explaining Mixtures of Sources in News Articles". *Findings of EMNLP* (37.7%), 2024.

Yufei Tian, Tenghao Huang, Miri Liu, Derek Jiang, Alexander Spangher, Muhao Chen, Jonathan May, and Nanyun Peng. "Are Large Language Models Capable of Generating Human-Level Narratives?". *Proc. EMNLP* (20.8%), 2024. **Outstanding Paper Award.**

Hyundong Justin Cho, Nicolaas Paul Jedema, Leonardo F. R. Ribeiro, Karishma Sharma, Pedro Szekely, Alessandro Moschitti, Ruben Janssen, and Jonathan May. "Speechworthy Instruction-tuned Language Models". *Proc. EMNLP* (20.8%), 2024.

Hyundong Justin Cho, Thamme Gowda, Yuyang Huang, Zixun Lu, Tianli Tong, and Jonathan May. "BotEval: Facilitating Interactive Human Evaluation". *Proc. ACL Demo Sessions*, 2024.

Wichayaporn Wongkamjan, Feng Gu, Yanze Wang, Ulf Hermjakob, Jonathan May, Brandon M. Stewart, Jonathan K. Kummerfeld, Denis Peskoff, and Jordan Lee Boyd-Graber. "More Victories, Less Cooperation: Assessing Cicero's Diplomacy Play". *Proc. ACL* (21.3%), 2024.

Virginia K. Felkner, Jennifer A. Thompson, and Jonathan May. "GPT is Not an Annotator: The Necessity of Human Annotation in Fairness Benchmark Construction". *Proc. ACL* (21.3%), 2024.

Alexander Spangher, Serdar Tumgoren, Ben Welsh, Nanyun Peng, Emilio Ferrara, and Jonathan May. "Tracking the Newsworthiness of Public Documents". *Proc. ACL* (21.3%), 2024.

Meryem M'hamdi, Jonathan May, Franck Dernoncourt, Trung Bui, and Seunghyun Yoon. "Multilingual Meta-Distillation Alignment for Semantic Retrieval". *Proc. SIGIR* (20.1%), 2024.

Hyundong Justin Cho, Shuai Liu, Taiwei Shi, Darpan Jain, Basem Rizk, Yuyang Huang, Zixun Lu, Nuan Wen, Jonathan Gratch, Emilio Ferrara, and Jonathan May. "Can Language Model Moderators Improve the Health of Online Discourse?". *Proc. NAACL* (23%), 2024.

Meryem M'hamdi and Jonathan May. "Leitner-Guided Memory Replay for Cross-lingual Continual Learning". *Proc. NAACL* (23%), 2024.

Alexander Spangher, Zihan Xue, Te-Lin Wu, Mark Hansen, and Jonathan May. "LegalDiscourse: Interpreting When Laws Apply and To Whom". *Proc. NAACL* (23%), 2024.

Zihao He, Jonathan May, and Kristina Lerman. "CPL-NoViD: Context-Aware Prompt-based Learning for Norm Violation Detection in Online Communities". *Proc. ICWSM* (20%), 2024.

Linghao Jin, Jacqueline He, Jonathan May, and Xuezhe Ma. "Challenges in Context-Aware Neural Machine Translation". *Proc. EMNLP* (21.3%), 2023.

Hyundong Justin Cho, Andrea Madotto, Zhaojiang Lin, Khyathi Chandu, Satwik Kottur, Jing Xu, Jonathan May, and Chinnadhurai Sankar. "Continual Dialogue State Tracking via Example-Guided Question Answering". *Proc.EMNLP* (21.3%), 2023.

Jihyung Moon, Dong-Ho Lee, Hyundong Justin Cho, Woojeong Jin, Chan Young Park, Minwoo Kim, Jonathan May, Jay Pujara, and Sungjoon Park. "Analyzing Norm Violations in Live-Stream Chat". *Proc.EMNLP* (21.3%), 2023.

Alexander Spangher, Nanyun Peng, Emilio Ferrara, and Jonathan May. "Identifying Informational Sources in News Articles". *Proc.EMNLP* (21.3%), 2023.

Rong-Ching Chang, Jonathan May and Kristina Lerman. "Feedback Loops and Complex Dynamics of Harmful Speech in Online Discussions". *Proc. SBP-BRiMS* (42%), 2023.

Kai Chen, Zihao He, Rong-Ching Chang, Jonathan May and Kristina Lerman. "Anger Breeds Controversy: Analyzing Controversy and Emotions on Reddit". *Proc. SBP-BRiMS* (42%), 2023.

Alexander Spangher, James Youn, Jonathan May, and Nanyun Peng. "First Steps Towards a Source Recommendation Engine: Investigating How Sources Are Used in News Articles". *Proc. The Joint Computation + Journalism European Data & Computational Journalism Conference* (46%), 2023.

Meryem M'hamdi, Xiang Ren, and Jonathan May, "Cross-lingual Continual Learning". *Proc. ACL* (22%), 2023.

Virginia Felkner, Ho-Chun Herbert Chang, Eugene Jang, and Jonathan May, "WinoQueer: A Community-in-the-Loop Benchmark for Anti-LGBTQ+ Bias in Large Language Models". *Proc. ACL* (22%), 2023.

Shuai Liu, Hyundong Cho, Marjorie Freedman, Xuezhe Ma, and Jonathan May, "RECAP: Retrieval-Enhanced Context-Aware Prefix Encoder for Personalized Dialogue Response Generation". *Proc. ACL* (22%), 2023.

Mozhdeh Gheini, Xuezhe Ma, and Jonathan May, "Know Where You're Going: Meta-Learning for Parameter-Efficient Fine-Tuning". *Findings of ACL* (41%), 2023.

Pengfei Yu, Jonathan May, and Heng Ji. "Bridging the Gap between Native Text and Translated Text through Adversarial Learning: A Case Study on Cross-Lingual Event Extraction". *Findings of EACL* (41%), 2023.

Xuezhe Ma, Chunting Zhou, Xiang Kong, Junxian He, Liangke Gui, Graham Neubig, Jonathan May, and Luke Zettlemoyer. "Mega: Moving Average Equipped Gated Attention". *Proc. ICLR* (32%), 2023.

Jacob Bremerman, Xiang Ren and Jonathan May. "Machine Translation Robustness to Natural Asemantic Variation". *Proc. EMNLP* (20%), 2022.

Nada Aldarrab and Jonathan May. "Segmenting Numerical Substitution Ciphers". *Proc. EMNLP* (20%), 2022.

Jiao Sun, Swabha Swayamdipta, Jonathan May and Xuezhe Ma. "Investigating the Benefits of Free-Form Rationales". *Findings of EMNLP* (33%), 2022.

Hyundong Cho, Chinnadhurai Sankar, Christopher Lin, Kaushik Ram Sadagopan, Shahin Shayandeh, Asli Celikyilmaz, Jonathan May and Ahmad Beirami. "Know Thy Strengths: Comprehensive Dialogue State Tracking Diagnostics". *Findings of EMNLP* (33%), 2022.

Ashutosh Joshi, Shankar Vishwanath, Choon Hui Teo, Vaclav Petricek, Vishy Vishwanathan, Rahul Bhagat, and Jonathan May. "Augmenting Training Data for Massive Semantic Matching Models in Low-Traffic E-commerce Stores". *Proc. NAACL Industry Track* (33%), 2022.

Alexander Spangher, Xiang Ren, Jonathan May, and Nanyun Peng. "NewsEdits: A Dataset of News Article Revision Histories and a Novel Document-Level Reasoning Challenge". *Proc. NAACL* (22%), 2022. **Outstanding Paper Award.**

Kushal Chawla, Gale Lucas, Jonathan May, and Jonathan Gratch. "Opponent Modeling in Negotiation Dialogues by Related Data Adaptation". *Findings of NAACL* (31%), 2022.

Hengameh Mirzaalian, Mohamed E. Hussein, Leonidas Spinoulas, Jonathan May and Wael Abd-Almageed. "Explaining Face Presentation Attack Detection Using Natural Language". *Proc. IEEE International Conference on Automatic Face and Gesture Recognition* (35%), 2021.

Xuezhe Ma, Xiang Kong, Sinong Wang, Chunting Zhou, Jonathan May, Hao Ma and Luke Zettlemoyer. "Luna: Linear Unified Nested Attention". *Proc. NeurIPS* (26%), 2021.

Alexander Spangher, Jonathan May, Sz-Rung Shiang and Lingjia Deng. "Multi-task Semi-Supervised Learning for Class-Imbalanced Discourse Classification". *Proc. EMNLP* (23%), 2021.

Xiyang Zhang, Muhao Chen and Jonathan May. "Salience-Aware Event Chain Modeling for Narrative Understanding". *Proc. EMNLP* (23%), 2021.

Mozhdeh Gheini, Xiang Ren and Jonathan May. "Cross-Attention is All You Need: Adapting Pretrained Transformers for Machine Translation". *Proc. EMNLP* (23%), 2021.

Thamme Gowda, Zhao Zhang, Chris Mattmann and Jonathan May. "Many-to-English Machine Translation Tools, Data, and Pretrained Models". *Proc. ACL Demo Track* (32%), 2021.

Karen Hambardzumyan, Hrant Khachatrian and Jonathan May. "WARP: Word-level Adversarial ReProgramming". *Proc. ACL* (21%), 2021.

Nada Aldarrab and Jonathan May. "Can Sequence-to-Sequence Models Crack Substitution Ciphers?". *Proc. ACL* (21%), 2021.

Thamme Gowda, Weiqiu You, Constantine Lignos and Jonathan May. "Macro-Average: Rare Types Are Important Too". *Proc. NAACL* (26%), 2021.

Meryem M'hamdi, Doo Soon Kim, Franck Dernoncourt, Trung Bui, Xiang Ren and Jonathan May. "X-METRA-ADA: Cross-lingual Meta-Transfer learning Adaptation to Natural Language Understanding and Question Answering". *Proc. NAACL* (26%), 2021.

Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May and Jonathan Gratch. "CaSiNo: A Corpus of Campsite Negotiation Dialogues for Automatic Negotiation Systems". *Proc. NAACL* (26%), 2021.

Xusen Yin, Ralph Weischedel and Jonathan May. "Learning to Generalize for Sequential Decision Making". *Findings of EMNLP* (38%), 2020.

Thamme Gowda and Jonathan May. "Finding the Optimal Vocabulary Size for Neural Machine Translation". *Findings of EMNLP* (38%), 2020.

Yonatan Bisk, Ari Holtzman, Jesse Thomason, Jacob Andreas, Yoshua Bengio, Joyce Chai, Mirella Lapata, Angeliki Lazaridou, Jonathan May, Aleksandr Nisnevich, Nicolas Pinto and Joseph Turian. "Experience Grounds Language". *Proc. EMNLP* (22%), 2020.

Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers and Clare Voss. "Connecting the Dots: Event Graph Schema Induction with Path Language Modeling". *Proc. EMNLP* (22%), 2020.

Justin Cho and Jonathan May. "Grounding Conversations with Improvised Dialogues". *Proc. ACL* (23%), 2020.

Di Lu, Ananya Subburathinam, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, and Clare Voss. "Cross-lingual Structure Transfer for Relation and Event Extraction". *Proc. LREC* (60%), 2020.

James Mullenbach, Jonathan Gordon, Nanyun Peng, and Jonathan May. "Do Nuclear Submarines Have Nuclear Captains? A Challenge Dataset for Commonsense Reasoning over Adjectives and Objects". *Proc. EMNLP* (26%), 2019.

Ananya Subburathinam, Di Lu, Heng Ji, Jonathan May, Shih-Fu Chang, Avirup Sil, and Clare Voss. "Cross-lingual Structure Transfer for Relation and Event Extraction". *Proc. EMNLP* (26%), 2019.

Xiaolei Huang, Jonathan May, and Nanyun Peng. "What Matters for Neural Cross-Lingual Named Entity Recognition: An Empirical Analysis". *Proc. EMNLP* (26%), 2019.

Meryem M'hamdi, Marjorie Freedman, and Jonathan May. "Contextualized Cross-Lingual Event Trigger Extraction with Minimal Resources". *Proc. CoNLL* (22%), 2019.

Elizabeth Boschee, Joel Barry, Jayadev Billa, Marjorie Freedman, Thamme Gowda, Constantine Lignos, Chester Palen-Michel, Michael Pust, Banriskhem Kayang Khonglah, Srikanth Madikeri, Jonathan May and Scott Miller. "SARAL: A Low-Resource Cross-Lingual Domain-Focused Information Retrieval System for Effective Rapid Document Triage". *Proc. ACL Demo Track* (35%), 2019.

Nima Pourdamghani, Nada Aldarrab, Marjan Ghazvininejad, Kevin Knight and Jonathan May. "Translating Translationese: A Two-Step Approach to Unsupervised Machine Translation". *Proc. ACL* (23%), 2019.

Xusen Yin and Jonathan May. "Comprehensible Context-driven Text Game Playing". *Proc. CoG* (41%), 2019.

Ronald Cardenas, Ying Lin, Heng Ji and Jonathan May. "A Grounded Unsupervised Universal Part-of-Speech Tagger for Low-Resource Languages". *Proc. NAACL* (26%), 2019.

Lifu Huang, Heng Ji and Jonathan May. "Cross-lingual Multi-Level Adversarial Transfer to Enhance Low-Resource Name Tagging". *Proc. NAACL* (26%), 2019.

Ulf Hermjakob, Jonathan May, Michael Pust, and Kevin Knight. "Translating a Language You Don't Know In the Chinese Room". *Proc. ACL Demo Track* (30%), 2018.

Ulf Hermjakob, Jonathan May, and Kevin Knight. "Out-of-the-box Universal Romanization Tool *uroman*". *Proc. ACL Demo Track* (30%), 2018. **Best Demo Award.**

Boliang Zhang, Ying Lin, Xiaoman Pan, Di Lu, Jonathan May, Kevin Knight and Heng Ji. "ELISA-EDL: A Cross-lingual Entity Extraction, Linking and Localization System". *Proc. NAACL Demo Track* (51%), 2018.

Yining Chen, Sorcha Gilroy, Andreas Maletti, Jonathan May, and Kevin Knight. "Recurrent Neural Networks as Weighted Language Recognizers". *Proc. NAACL* (30%), 2018. **Outstanding Paper Award.**

Pavlos Papadopoulos, Ruchir Travadi, Colin Vaz, Nikolaos Malandrakis, Ulf Hermjakob, Nima Pourdamghani, Michael Pust, Boliang Zhang, Xiaoman Pan, Di Lu, Ying Lin, Ondřej Glembek, Murali Karthick Baskar, Martin Karafiát, Lukáš Burget, Jonathan May, Heng Ji, Kevin Knight, and Shrikanth Narayanan. "Team ELISA System for DARPA LORELEI Speech Evaluation 2016". *Proc. Interspeech* (51%), 2017.

Xiaoman Pan, Boliang Zhang, Jonathan May, Joel Nothman, Kevin Knight and Heng Ji. "Cross-lingual Name Tagging and Linking for 282 Languages". *Proc. ACL* (23%), 2017.

Barret Zoph, Deniz Yuret, Jonathan May and Kevin Knight. "Transfer Learning for Low-Resource Neural Machine Translation". *Proc. EMNLP* (24%), 2016.

Barret Zoph, Ashish Vaswani, Jonathan May and Kevin Knight. "Simple, Fast Noise-Contrastive Estimation for Large RNN Vocabularies". *Proc. NAACL* (26%), 2016.

Eunsol Choi, Matic Horvat, Jonathan May, Kevin Knight and Daniel Marcu. "Extracting Structured Scholarly Information from the Machine Translation Literature". *Proc. LREC* (60%), 2016.

Michael Pust, Ulf Hermjakob, Kevin Knight, Daniel Marcu, and Jonathan May. "Parsing English into Abstract Meaning Representation Using Syntax-Based Machine Translation". *Proc. EMNLP* (26%), 2015.

Jonathan May, Yassine Benjira, and Abdessamad Echihabi. "An Arabizi-English Social Media Statistical Machine Translation System". *Proc. AMTA*, 2014.

Alberto Barrón-Cedeño, Lluís Màrquez, Carlos A. Henríquez Q., Lluís Formiga, Enrique Romero, and Jonathan May. "Identifying Useful Human Correction Feedback from an On-line Machine Translation Service". *Proc. IJCAI* (28%), 2013.

Mark Hopkins and Jonathan May. "Models of Translation Competitions". *Proc. ACL* (26%), 2013.

Daniel Pighin, Lluís Màrquez, and Jonathan May. "An Analysis (and an Annotated Corpus) of User Responses to Machine Translation Output". *Proc. LREC* (69%), 2012

Mark Hopkins and Jonathan May. "Tuning as Ranking". *Proc. EMNLP* (24%), 2011.

Jonathan May, Kevin Knight, and Heiko Vogler. "Efficient Inference Through Cascades of Weighted Tree Transducers". *Proc. ACL* (25%), 2010.

Jonathan May and Kevin Knight. "Syntactic Re-Alignment Models for Machine Translation". *Proc. EMNLP* (27%), 2007.

Johanna Högberg, Andreas Maletti, and Jonathan May. "Bisimulation Minimisation for Weighted Tree Automata". *Proc. International Conference on Developments in Language Theory (DLT)*, 2007.

Johanna Högberg, Andreas Maletti, and Jonathan May. "Backward and Forward Bisimulation Minimisation of Tree Automata". *Proc. International Conference on Implementation and Application of Automata (CIAA)* (29%), 2007.

Jonathan May and Kevin Knight. "Tiburon: A Weighted Tree Automata Toolkit". *Proc. International Conference on Implementation and Application of Automata (CIAA)*, 2006.

Jonathan May and Kevin Knight. "A Better N-Best List: Practical Determinization of Weighted Finite Tree Automata". *Proc. NAACL-HLT* (24%), 2006.

Refereed Journal Articles (6)

Ulf Hermjakob, Qiang Li, Daniel Marcu, Jonathan May, Sebastian J. Mielke, Nima Pourdamghani, Michael Pust, Xing Shi, Kevin Knight, Tomer Levinboim, Kenton Murray, David Chiang, Boliang Zhang, Xiaoman Pan, Di Lu, Ying Lin, Heng Ji. "Incident-Driven Machine Translation and Name Tagging for Low-resource Languages". *Machine Translation* (2.91 IF). October, 2017.

Lifu Huang, Jonathan May, Xiaoman Pan, Heng Ji, Xiang Ren, Jiawei Han, Lin Zhao, James A. Hendler. "Liberal Entity Extraction: Rapid Construction of Fine-Grained Entity Typing Systems". *Big Data* (2.55 IF). 5(1). March, 2017.

Wei Wang, Jonathan May, Kevin Knight, and Daniel Marcu. "Re-Structuring, Re-Labeling, and Re-Aligning for Syntax-Based Machine Translation". *Computational Linguistics* (1.32 IF). 36(2). June, 2010.

Johanna Högberg, Andreas Maletti, and Jonathan May. "Backward and Forward Bisimulation Minimisation of Tree Automata". *Theoretical Computer Science* (0.72 IF). 410(37). September, 2009.

Jonathan Graehl, Kevin Knight, and Jonathan May. "Training Tree Transducers". *Computational Linguistics* (1.32 IF), 34(3). September, 2008.

Jonathan May, Ada Brunstein, Prem Natarajan, and Ralph Weischedel. "Surprise! What's in a Cebuano or Hindi Name?". *ACM Transactions on Asian Language Information Processing* (1.65 IF), 2(3). September, 2003.

Refereed Workshop Papers (22)

Jaspreet Ranjit, Hyundong J. Cho, Claire J. Smerdon, Yoonsoo Nam, Myles Phung, Jonathan May, John R. Blosnich, and Swabha Swayamdipta. "Uncovering Intervention Opportunities for Suicide Prevention with Language Model Assistants". *Proc. GenAI4Health Workshop at NeurIPS*, 2025.

Shuai Liu, Shantanu Agarwal, and Jonathan May. "Authorship Style Transfer with Policy Optimization". *Proc. CustomNLP4U*, 2024.

Tenghao Huang, Dong Hee Lee, John Sweeney, Jiatong Shi, Emily Steliotes, Matthew Lange, Jonathan May, and Muhao Chen. "FOODPUZZLE: Toward Developing Large Language Models as Flavor Scientist". *Proc. NLP4Science*, 2024.

Hamed Bonab, Ashutosh Joshi, Ravi Bhatia, Ankit Gandhi, Vijay Huddar, Juhi Naik, Mutasem Al-Darabsah, Choon Hui Teo, Jonathan May, Tarun Agarwal, and Vaclav Petricek. "Blend and Match: Distilling Semantic Search Models with Different Inductive Biases and Model Architectures". *Proc. ISIR-eCom*, 2023.

Pengfei Yu, Zixuan Zhang, Clare Voss, Jonathan May, and Heng Ji. "Building an Event Extractor with Only a Few Examples". *Proc. DeepLo*, 2022.

Virginia K. Felkner, Ho-Chun Herbert Chang, Eugene Jang, and Jonathan May. "Towards WinoQueer: Developing a Benchmark for Anti-Queer Bias in Large Language Models". *Proc. Queer in AI @ NAACL*, 2022.

Nicolaas Weideman, Virginia K. Felkner, Wei-Cheng Wu, Jonathan May, Christophe Hauser and Luis Garcia. "PERFUME: Programmatic Extraction and Refinement for Usability of Mathematical Expression". *Proc. Checkmate*, 2021.

Xusen Yin, Li Zhou, Kevin Small, and Jonathan May. "Summary-Oriented Question Generation for Informational Queries". *Proceedings of the 1st Workshop on Document-grounded Dialogue and Conversational Question Answering (DialDoc 2021)*, 2021.

Aurelie Herbelot, Xiaodan Zhu, Alexis Palmer, Nathan Schneider, Jonathan May and Ekaterina Shutova (Editors). *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, 2020.

Alexander Spangher, Nanyun Peng, Jonathan May, and Emilio Ferrara. "Enabling Low-Resource Transfer Learning across COVID-19 Corpora by Combining Event-Extraction and Co-Training". *Proc. Workshop on NLP for COVID-19 at ACL 2020*, 2020.

Karin Verspoor, Kevin Bretonnel Cohen, Mark Dredze, Emilio Ferrara, Jonathan May, Robert Munro, Cecile Paris, and Byron Wallace (Editors). *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*, 2020.

Alexander Spangher, Jonathan May, Emilio Ferrara, and Nanyun Peng. "'Don't quote me on that': Finding Mixtures of Sources in News Articles". *Proc. Computation + Journalism Symposium*, 2020.

Alexander Spangher, Emilio Ferrara, Jonathan May, and Nanyun Peng. "Modeling 'Newsworthiness' for Lead-Generation Across Corpora". *Proc. Computation + Journalism Symposium*, 2020.

Xiaoman Pan, Thamme Gowda, Heng Ji, Jonathan May, and Scott Miller. "Cross-lingual Joint Entity and Word Embedding to Improve Entity Linking and Parallel Sentence Mining". *Proc. DeepLo*, 2019.

Jonathan May, Ekaterina Shutova, Auerlie Herbelot, Xiaodan Zhu, Marianna Apidianaki, and Saif M. Mohammad (Editors). *Proceedings of the 13th International Workshop on Semantic Evaluation*. 2019.

Nanyun Peng, Marjan Ghazvininejad, Jonathan May, and Kevin Knight. "Towards Controllable Story Generation". *Proc. of the 1st Workshop on Storytelling*, 2018.

Marianna Apidianaki, Saif M. Mohammad, Jonathan May, Ekaterina Shutova, Steven Bethard, and Marine Carpuat (Editors). *Proceedings of the 12th International Workshop on Semantic Evaluation*. 2018.

Jonathan May and Jay Priyadarshi. "SemEval-2017 Task 9: Abstract Meaning Representation Parsing and Generation". *Proc. SemEval*, 2017.

Jonathan May. "SemEval-2016 Task 8: Meaning Representation Parsing". *Proc. SemEval*, 2016.

Jonathan Gordon, Jerry Hobbs, Jonathan May, Michael Mohler, Fabrizio Morbini, Bryan Rink, Marc Tomlinson, and Suzanne Wertheim. "A Corpus of Rich Metaphor Annotation". *Proc. Workshop on Metaphor in NLP*, 2015.

Jonathan Gordon, Jerry Hobbs, Jonathan May, and Fabrizio Morbini. "High-Precision Abductive Mapping of Multilingual Metaphors". *Proc. Workshop on Metaphor in NLP*, 2015.

Matthias Büchse, Jonathan May, and Heiko Vogler. "Determinization of Weighted Tree Automata Using Factorizations". *Proc. ATANLP*, 2009.

Abstracts (2)

Ece Kamar, Jonathan May, Hal Daumé III, and Maria Gini. "Large Language Models: Helpful Assistants, Romantic Partners, or Con Artists?" Panel at *AAAS Conference*, 2024.

Jennifer Thompson, Matthew Boxer, Virginia Kathryn Felkner, Gregg Drinkwater, David Schraub, and Jonathan May. "Defining and disrupting antisemitism: Perspectives from artificial intelligence, sociology, history, law, and ethics". Hybrid Roundtable at *Assoc. Jewish Studies*, 2023.

Book Chapters (2)

Jonathan May and Joseph Dane. "Evidence and Artificial Intelligence". In Joseph A. Dane et al., *Begging The Question: Chauceriana, Book History, and Humanistic Inquiry (Mythodologies II)*. Los Angeles: Loyola Marymount Univ. Press, 2019.

Kevin Knight and Jonathan May. "Applications of Weighted Automata in Natural Language Processing". In M. Droste, W. Kuich, and H. Vogler, editors, *Handbook of Weighted Automata*. Springer-Verlag, 2009.

Other Publications (4)

Jonathan May, Virginia K. Felkner, and Jennifer Thompson. "We Can Have AI without Antisemitism—If We Want It". *AJS Perspectives*. Summer, 2024.

Jonathan May. "ChatGPT is great – you're just using it wrong". *The Conversation*. February 2, 2023.

Jinxi Xu, Ana Licuanan, Jonathan May, Scott Miller, and Ralph Weischedel. "Answer Selection and Confidence Estimation". *New Directions in Question Answering, Papers from 2003 AAAI Spring Symposium*, Stanford University, Stanford, CA AAAI Press, 2003.

Jinxi Xu, Ana Licuanan, Jonathan May, Scott Miller, and Ralph Weischedel. "TREC 2002 QA at BBN: Answer Selection and Confidence Estimation". *Proc. TREC*, 2002.

Service

Screening Exam Committee: Mary Kennedy (2024)

Qualifying Exam Committee: Nikos Malandrakis (2019), Brendan Kennedy (2020), Jacqueline Brixey (2020), Karan Singla (2020), Sami Abu-El-Haija (2020), Xusen Yin (Chair, 2020), Nada Aldarrab (Chair, 2020), Thamme Gowda (Chair, 2021), Yury Zemlyanskiy (2021), Meryem M'Hamdi (Chair, 2021), Jiao Sun (2021), Sabyasachee Baruah (2021), Kushal Chawla (2021), Mozhdeh Gheini (Chair, 2022), Wenxuan Zhou (2022), Xisen Jin (2022), Alexander Spangher (Chair, 2022), Nan Xu (2023), Justin Cho (Chair, 2023), Katy Felkner (Chair, 2023), Tenghao Huang (Chair, 2024), Preni Golazizian (2024), Rebecca Dorn (2024), Alireza Salkhordeh (2024), Linghao Jin (2024), Suhaib Abdurahman (2024), Taiwei Shi (2025)

Dissertation Proposal Committee: Austin Matthews (CMU, 2018), Marjan Ghazvininejad (2019), Nima Pourdamghani (2019), Xing Shi (2019), Victor Martinez (2020), Karan Singla (2020), Xusen Yin (Chair, 2021), Sami Abu-El-Haija (2021), Aida Mostafazadeh Davani (2021), Nada Aldarrab (Chair, 2021), Yury Zemlyanskiy (2021), Brendan Kennedy (2021), Thamme Gowda (Chair, 2022), Jacqueline Brixey (2022), Jiao Sun (2023), Sabyasachee Baruah (2023), Mozhdeh Gheini (Chair, 2023), Zihao Huang (2023), Meryem M'Hamdi (Chair, 2023), Kushal Chawla (2023), Nan Xu (2024), Alexander Spangher (Chair, 2024), Lee Kezar (2024), Justin Cho (Chair, 2025), Pegah Jandaghi (2025), Tenghao Huang (2025)

Dissertation Defense Committee: Austin Matthews (CMU, 2019), Marjan Ghazvininejad (2019), Nima Pourdamghani (2019), Xing Shi (2019), Xusen Yin (Chair, 2021), Karan Singla (2021), Sami Abu-el-Haija (2022), Thamme Gowda (Chair, 2022), Nada Aldarrab (Chair, 2022), Wenxuan Zhou (2023), Meryem M'Hamdi (Chair, 2024), Jiao Sun (2024), Kushal Chawla (2024), Lee Kezar (2025), Alexander Spangher (Chair, 2025)

MS Thesis Committee: Nada Aldarrab (2017)

Ph.D Fellowship Committee: 2017–2019

Research Assistant Professor Appointment Committee: 2018–2020
Research Associate Professor Advancement Committee: 2023 , 2024

Professional Activity

Associate Treasurer, ACL 2025–

Organizing Committee, ACL 2026

Best Paper Committee, ACL 2023

Organizing Committee, ICML workshop: What's left to TEACH chatbots? 2023

Senior Area Chair, Semantics, AACL-IJCNLP 2022

Senior Area Chair, Theme track, EMNLP 2022

Area Chair, Machine Translation and Multilinguality, AACL 2020

Publication Chair, *SEM 2020

Area Chair, Semantics, ACL 2020

Area Chair, Multilinguality, ACL 2019

Symposium Co-Organizer, SoCalNLP 2019

Treasurer, NAACL, 2019–2024

Handbook Chair, NAACL HLT 2018

Symposium Co-Organizer, WeCNLP 2018

Area Chair, Semantics, NAACL HLT 2018

Workshop Co-Organizer, SemEval 2018, SemEval 2019

Task Organizer, Task 9 (AMR Parsing and Generation), SemEval 2017

Task Organizer, Task 8 (Meaning Representation Parsing), SemEval 2016

Local Organizer, NACLO 2015–present

Social Media Chair, NAACL HLT 2015

Program Committee, AAAI, ACL, EACL, EMNLP, NAACL, NIPS, AMTA, IJCNLP, CIAA, NACLO, *SEM, SemEval proposals, ACL Rolling Review

Reviewer, MIT Press, TACL, *Computational Linguistics*, *Journal of Machine Learning*, NSF, NSF-BSF, ANR (France), NWO (Netherlands)

Local Chair for NAACL HLT 2010 (with David Chiang, Jason Riesa, Ed Hovy)

Coordinator, ISI Natural Language Seminar (2006-7)

Invited Talks

"Better Communication With Computers"

Saarland University *May 2025*

"LA's AI Secrets Unveiled: AI Pioneers Reveal Hidden Past & Bold Predictions for the Future of AIs" (panel, with Adam Russell, Yolanda Gil, Karl Jacob, and Kristina Lerman).

LA Tech Week *Oct. 2024*

"Large Language Models: Helpful Assistants, Romantic Partners or Con Artists?" (panel, with Ece Kamar, Hal Daumé III, and Maria Gini).

AAAS Conference *Feb. 2024*

"Fairness & Transparency" (panel moderator, with Bobby Ghajar, Nathanael Fast, and Leslie Saxon).

USC Marshall Seminar on Responsible AI in Business *Jan. 2024*

"Natural Language Processing Pedagogy and Deployment"

USC CET Faculty Showcase *Oct. 2023*

"Talk to The Trolls! Dialogue Models to Aid Engaged Moderation in Online Forums"

SICon Workshop, Toronto, Canada *Jul. 2023*

"Multi-Lingual Meta-Learning"

Amazon Machine Learning Conference
Workshop on Multilingual NLP (virtual) *Oct. 2022*

"Rethinking Assumptions in Multilingual NLP"

Amazon Machine Learning Conference
Workshop on Multilingual NLP (virtual) *Oct. 2021*

"Translation For All"

USC AI Futures Symposium
on Artificial Intelligence and Data Science (virtual) *May 2021*

"Theory of Mind and Intent in Procedural Agents: Learning Lessons from Improvisational Theater and Text-based Adventure Games"

University of Pennsylvania (virtual; class guest lecture) *Apr. 2022*

University of Illinois at Urbana-Champaign (virtual) *Jun. 2020*

"Resource-Constrained Neural Machine Translation"

Arizona State University, Phoenix, AZ *Jan. 2020*

"Machine Translation: 350 years of progress and new challenges in the connectionist age"

Reed College, Portland, OR *Feb. 2019*

AIHacks High School Hackathon, Playa del Rey, CA *Jun. 2019*

"How I Learned to Stop Worrying and Love Evaluations"

Johns Hopkins University, Baltimore, MD *Nov. 2016*

University of Pennsylvania, Philadelphia, PA *Nov. 2016*

"The Machine Learning of Machine Translation"

University of Southern California, Los Angeles, CA *Aug. 2015*

"Using Syntax-Based Machine Translation to Parse English into Abstract Meaning Representation"

University of Edinburgh, Edinburgh, UK *Mar. 2015*

"Machine Translation: How it Works, Why it's Hard, and How To Make it Better"

Bloomberg, London, UK *Mar. 2015*

University of Edinburgh, Edinburgh, UK *Mar. 2015*

"Toward User-Focused NLP"

USC-ISI, Marina del Rey, CA *Feb. 2014*

"Models of Translation Competitions"

USC-ISI, Marina del Rey, CA *Aug. 2013*

"Tuning As Ranking"

USC-ISI, Marina del Rey, CA *Jul. 2011*

"Natural Language Processing and Weighted Finite-State Machines"

Umeå University, Umeå, Sweden *Oct. 2008*

"Syntactic Re-Alignment Models for Machine Translation"

University of California, Berkeley, CA *Apr. 2008*

"Bisimulation Minimisation for Weighted Tree Automata"

Dresden University of Technology, Dresden, Germany *Jun. 2007*

Honors and Awards

Outstanding Paper Award, EMNLP, 2024

ISI Research Award ($100,000), "Detecting Bias in the Law", 2024

ISI Research Award ($100,000), "Multi-document Newsworthy Event Monitoring and Forecasting", 2023

Outstanding Paper Award, NAACL, 2022

Outstanding Reviewer, EMNLP, 2020

ISI Research Award ($100,000), "Universal Translators for Asylum Seekers at the Border", 2020

Best Demo Award, ACL, 2018

Outstanding Paper Award, NAACL, 2018

USC School of Engineering Doctoral Fellowship, 2004–2008