

Jonathan K. Kummerfeld

Contact E-mail: jkummerf@umich.edu

Information Website: <http://www.jkk.name>

Research **Natural Language Processing:** Dialogue, Executable Semantic Parsing
Interests **Crowdsourcing:** Hybrid Intelligence Systems

Current Postdoctoral Research Fellow September 2016 – Present
Computer Science Engineering, University of Michigan

U.S. Permanent Resident (Green Card Holder), Australian Citizen

Education **University of California, Berkeley**, Berkeley, CA
Ph.D. in Computer Science, August 2016
Advisor: Dan Klein

University of Sydney, Sydney, Australia
B.S. in Computer Science, with University Medal, May 2010
Advisor: James R. Curran

Publications: 19 papers in TACL, ACL, NAACL and EMNLP (top venues in NLP)

Overview Google Scholar metrics:

H-index: 18

H5-index: 15

Total Citations: 990

In the list below, **my name is bold** and the names of students I co-advised are underlined (including graduates, undergraduates, company employees, and interns).

Publications: C. Welch, **J. K. Kummerfeld**, V. Perez-Rosas, R. Mihalcea. 2020. Compositional Demographic Word Embeddings. *Empirical Methods in Natural Language Processing (EMNLP)*.

C. Welch, R. Mihalcea, **J. K. Kummerfeld**. 2020. Improving Low Compute Language Modeling with In-Domain Embedding Initialisation. *Empirical Methods in Natural Language Processing (EMNLP)*.

S. Larson, A. Zheng, A. Mahendran, R. Tekriwal, A. Cheung, E. Guldán, K. Leach, **J. K. Kummerfeld**. 2020. Iterative Feature Mining for Constraint-Based Data Collection to Increase Data Diversity and Model Robustness. *Empirical Methods in Natural Language Processing (EMNLP)*.

Y. Jiang, H. Zhu, **J. K. Kummerfeld**, Y. Li, W. Lasecki. 2020. A Novel Workflow for Accurately and Efficiently Crowdsourcing Predicate Senses and Argument Labels. *Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP)*.

- C. Welch, **J. K. Kummerfeld**, V. Perez-Rosas, R. Mihalcea. 2020. Exploring the Value of Personalized Word Embeddings. *International Conference on Computational Linguistics (CoLing)*.
- S. Larson, A. Cheung, A. Mahendran, K. Leach, **J. K. Kummerfeld**. 2020. Inconsistencies in Crowdsourced Slot-Filling Annotations: A Typology and Identification Methods. *International Conference on Computational Linguistics (CoLing)*.
- J. S. Huffaker, **J. K. Kummerfeld**, W. S. Lasecki, M. S. Ackerman. 2020. Crowdsourced Detection of Emotionally Manipulative Language. *Conference on Human Factors in Computing Systems (CHI)*.
- P. Paquette, Y. Lu, S. Bocco, M. O. Smith, S. Ortiz-Gagné, **J. K. Kummerfeld**, J. Pineau, S. Singh, A. Courville. 2019. No-Press Diplomacy: Modeling Multi-Agent Gameplay. *Advances in Neural Information Processing Systems (NeurIPS)*.
- S. Larson, A. Mahendran, J. J. Peper, C. Clarke, A. Lee, P. Hill, **J. K. Kummerfeld**, K. Leach, M. A. Laurenzano, L. Tang and J. Mars. 2019. An Evaluation for Intent Classification and Out-of-Scope Prediction. *Empirical Methods in Natural Language Processing (EMNLP)*.
- J. K. Kummerfeld**, S. R. Gouravajhala, J. Peper, V. Athreya, C. Gunasekara, J. Ganhotra, S. S. Patel, L. Polymenakos, W. S. Lasecki. 2019. A Large-Scale Corpus for Conversation Disentanglement. *Association for Computational Linguistics (ACL)*.
- J. K. Kummerfeld**. 2019. SLATE: A Super-Lightweight Annotation Tool for Experts. *Association for Computational Linguistics: Demonstrations (ACL: Demo)*.
- S. Larson, A. Mahendran, A. Lee, **J. K. Kummerfeld**, P. Hill, M. Laurenzano, J. Hauswald, L. Tang, J. Mars. 2019. Outlier Detection for Improved Data Quality and Diversity in Dialog Systems. *North American Chapter of the Association for Computational Linguistics (NAACL)*.
- C. Welch, V. Perez-Rosas, **J. K. Kummerfeld**, R. Mihalcea. 2019. Look Who’s Talking: Inferring Speaker Attributes from Personal Longitudinal Dialog. *Computational Linguistics and Intelligent Text Processing (CICLing)*. **Best Student Paper**
- C. Finegan-Dollak, **J. K. Kummerfeld**, L. Zhang, K. R. D. Ramanathan, S. Sadasivam, R. Zhang, D. Radev. 2018. Improving Text-to-SQL Evaluation Methodology. *Association for Computational Linguistics (ACL)*.
- L. Burdick, **J. K. Kummerfeld**, R. Mihalcea. 2018. Factors Influencing the Surprising Instability of Word Embeddings. *North American Chapter of the Association for Computational Linguistics (NAACL)*.
- Y. Jiang, C. Finegan-Dollak, **J. K. Kummerfeld**, W. Lasecki. 2018. Effective Crowdsourcing for a New Type of Summarization Task. *North American Chapter of the Association for Computational Linguistics (NAACL)*.

- Y. Kang, Y. Zhang, **J. K. Kummerfeld**, P. Hill, J. Hauswald, M. A. Laurenzano, L. Tang, J. Mars. 2018. Data Collection for Dialogue System: A Startup Perspective. *North American Chapter of the Association for Computational Linguistics (NAACL: Industry Track)*.
- C. Welch, **J. K. Kummerfeld**, S. Feng, R. Mihalcea. 2018. World Knowledge for Semantic Parsing with Abstract Meaning Representation. *Language Resources and Evaluation Conference (LREC)*.
- G. Durrett, **J. K. Kummerfeld**, T. Berg-Kirkpatrick, R. S. Portnoff, S. Afroz, D. McCoy, K. Levchenko, V. Paxson. 2017. Identifying Products in Online Cyber-crime Marketplaces: A Dataset for Fine-grained Domain Adaptation. *Empirical Methods in Natural Language Processing (EMNLP)*.
- Y. Jiang, **J. K. Kummerfeld**, W. S. Laseck. 2017. Understanding Task Design Trade-offs in Crowdsourced Paraphrase Collection. *Association for Computational Linguistics (ACL)*.
- R. S. Portnoff, S. Afroz, G. Durrett, **J. K. Kummerfeld**, T. Berg-Kirkpatrick, D. McCoy, K. Levchenko, V. Paxson. 2017. Tools for Automated Analysis of Cybercriminal Markets. *International World Wide Web Conferenc (WWW)*.
- J. K. Kummerfeld**, T. Berg-Kirkpatrick, D. Klein. 2015. An Empirical Analysis of Optimization for Max-Margin NLP. *Empirical Methods in Natural Language Processing (EMNLP)*.
- J. K. Kummerfeld**, D. Klein. 2013. Error-Driven Analysis of Challenges in Coreference Resolution. *Empirical Methods in Natural Language Processing (EMNLP)*.
- J. K. Kummerfeld**, D. Tse, J. R. Curran, D. Klein. 2013. An Empirical Examination of Challenges in Chinese Parsing. *Association for Computational Linguistics (ACL)*.
- J. K. Kummerfeld**, D. Hall, J. R. Curran, D. Klein. 2012. Parser Showdown at the Wall Street Corral: An Empirical Investigation of Error Types in Parser Output. *Empirical Methods in Natural Language Processing (EMNLP)*.
- J. K. Kummerfeld**, D. Klein, J. R. Curran. 2012. Robust Conversion of CCG Derivations to Phrase Structure Trees. *Association for Computational Linguistics (ACL)*.
- J. K. Kummerfeld**, M. Bansal, D. Burkett, D. Klein. 2011. Mention Detection: Heuristics for the OntoNotes annotations. *Conference on Natural Language Learning Shared Task (CoNLL Shared Task)*.
- J. K. Kummerfeld**, J. Roesner, T. Dawborn, J. Haggerty, J. Curran, S. Clarke. 2010. Faster Parsing by Supertagger Adaptation. *Association for Computational Linguistics (ACL)*.

M. Honnibal, **J. K. Kummerfeld**, J. R. Curran. 2010. Morphological Analysis Can Improve a CCG Parser for English. *International Conference on Computational Linguistics (CoLing)*.

Publications: L. F. D’Haro, K. Yoshino, C. Hori, T. K. Marks, L. Polymenakos, **J. K. Kummerfeld**, M. Galley, X. Gao. 2020. Overview of the seventh Dialog System Technology Challenge: DSTC7. *Computer Speech & Language (CSL)*.

Journals

J. K. Kummerfeld, D. Klein. 2017. Parsing with Traces: An $O(n^4)$ Algorithm and a Structural Representation. *Transactions of the Association for Computational Linguistics (TACL)*.

V. A. Moss, N. M. McClure-Griffiths, T. Murphy, D. J. Pisano, **J. K. Kummerfeld**, J. R. Curran. 2013. High-velocity Clouds in the Galactic All Sky Survey. I. Catalog. *The Astrophysical Journal Supplement Series*.

R. Candelier, A. Widmer-Cooper, **J. K. Kummerfeld**, O. Dauchot, G. Biroli, P. Harrowell, D. R. Reichman. 2010. Spatiotemporal Hierarchy of Relaxation Events, Dynamical Heterogeneities, and Structural Reorganization in a Supercooled Liquid. *Physical Review Letters (PRL)*.

J. K. Kummerfeld, T. S. Hudson, P. Harrowell. 2008. The Densest Packing of AB Binary Hard-Sphere Homogeneous Compounds across all Size Ratios. *The Journal of Physical Chemistry B*.

Publications: **J. K. Kummerfeld**. 2020. Qualification Labour: A Fair Wage Isn’t Enough if Workers Need to Do 5,000 Low Paid Tasks to Qualify for Your Task. *AAAI Conference on Human Computation and Crowdsourcing (HComp WIP)*.

Workshops

C. Gunasekara, **J. K. Kummerfeld**, L. Lastras, W. S. Lasecki. 2020. NOESIS II: Predicting Responses, Identifying Success, and Managing Complexity in Task-Oriented Dialogue. *Dialogue System Technology Challenges (AAAI Workshop)*.

J. S. Huffaker, **J. K. Kummerfeld**, W. S. Lasecki, M. S. Ackerman. 2019. Training Data Voids: Novel Attacks Against NLP Content Moderation. *Volunteer Work: Mapping the Future of Moderation Research (CSCW Workshop)*.

S. Kim, M. Galley, C. Gunasekara, S. Lee, A. Atkinson, B. Peng, H. Schulz, J. Gao, J. Li, M. Adada, M. Huang, L. Lastras, **J. K. Kummerfeld**, W. S. Lasecki, C. Hori, A. Cherian, T. K. Marks, A. Rastogi, X. Zang, S. Sunkara, R. Gupta. 2019. The Eighth Dialog System Technology Challenge. *Conversational AI (NeurIPS Workshop)*.

C. Gunasekara, **J. K. Kummerfeld**, L. Polymenakos, W. S. Lasecki. 2019. DSTC7 Task 1: Noetic End-to-End Response Selection. *Dialog System Technology Challenges (AAAI Workshop)*.

J. K. Kummerfeld, J. Roesner, J. R. Curran. 2009. Faster Parsing and Supertagging Model Estimation. *Australasian Language Technology Association (ALTA)*.

J. K. Kummerfeld, J. R. Curran. 2008. Classification of Verb Particle Constructions with the Google Web1T Corpus. *Australasian Language Technology Association (ALTA)*.

Grants	Bloomberg Data Science Research Grant	2019
	Co-PI on grant to develop adaptive annotation methods	
	Michigan Institute for Data Science Music Initiative	2018
	Co-PI on grant to develop models that map text to music	
	Allen Institute for AI Key Scientific Challenges Program	2017
	Unrestricted funding to support research	
Awards	Outstanding Postdoctoral Fellow Award	2018
	For excellence in research, mentorship, teaching, leadership, and service	
	Chair's Special Award	2013
	For services to the EECS department during student admissions and recruitment	
	Outstanding Graduate Student Instructor Award	2011
	University wide award in recognition of exceptional achievements as a teacher	
	General Sir John Monash Award	2009
Employment	"Australia's most prestigious postgraduate scholarships"	
	University Medal in Computer Science	2009
	For outstanding results, awarded at the discretion of the university senate	
	Postdoctoral Research Fellow	
	University of Michigan	September 21st, 2016 – Present
	Consultant / Technical Adviser	
	Naming Matters	May 2015 – Present
	Clinic	January 2017 – February 2020
	WriteLab (acquired by Chegg)	March 2014 – May 2018
	Techbridge	September 2013 – May 2014
	Intern	
	Google Research – Mountain View, NLP team	May - August 2012
	Google – Sydney, Infrastructure team	December 2007 - February 2008
	Research Assistant	
	U. Sydney, School of Information Technologies, Prof. Kay	2010
	U. Sydney, School of Physics, Dr. Murphy	2010
	U. Sydney, School of Chemistry, Prof. Harrowell	2009
	U. Sydney, School of Letters, Art and Media, Dr. Bode	2008, 2009
	Instructor of Record	
	Artificial Intelligence, Berkeley	Summer 2014
Teaching,		
Mentoring,		
Invited Talks	Co-taught upper division course: gave lectures, managed TAs, and wrote exams. Evaluation: 6.3 / 7 (Median in CS is 5.7, and only 16% have a rating above 6.3)	

Teaching Assistant

Artificial Intelligence, Berkeley	Fall 2011
Held sections and office hours, prepared assignments, exams and sections.	
Evaluation: 4.6 / 5 (Only 11% of scores are public, of those the median is 4.6)	
Artificial Intelligence (Advanced), Sydney	Fall 2009
Held sections (third year elective Computer Science course).	
Informatics (Advanced), Sydney	Fall 2009
Held sections and graded assignments (first year core Computer Science course).	

Guest Lecturer

Conversational AI, Michigan	Spring 2020
STEM Posse, Michigan	Summer 2019
The Anatomy of Natural Language Processing Systems, Michigan	Winter 2018
The Beauty and Joy of Computing, Berkeley	Summer 2014
Applied Natural Language Processing, Berkeley	Fall 2013
Proseminar in the Digital Humanities, Berkeley	Spring 2013

Talks

University of California, Berkeley, NLP Seminar	February 2018
Midwest Speech and Language Days & CL Colloquium	May 2017, May 2018
University of Macquarie, Centre for Language Tech.	June 2016, August 2017
Sydney Humans, Language, Technology MeetUp	May 2016, August 2017
University of Sydney	May 2016
Information Sciences Institute, Natural Language Seminar	March 2016
CommonCrawl, Big Open Data Hackathon	May 2014
Cambridge Computer Laboratory	August 2013
Berkeley Syntax Circle	December 2012

Undergraduate & Masters Mentoring

Note, this list only includes students I was the primary mentor for.

Joseph Peper (transfer), Research Opportunity Program	2016-2017
Thomas Searle (1st year), Research Opportunity Program	2016-2017
Hyun A. Chung (2nd year), Research Opportunity Program	2017-2018
Orion Cleaver (2nd year), Research Opportunity Program	2017-2018
Alexandra Wu (2nd year), Explore CS Research	Spring 2019
Yuyi Qu (2nd year), Explore CS Research	Spring 2019
Yuntian Zhao (2nd year), Explore CS Research	Spring 2020
Prateek Jain (4th year), Explore CS Research	Spring 2020
Olivia Garrahan (1st year), CROMA Lab	2018-2019
Sophia He (2nd year), CROMA Lab	Fall 2018
Varun Kutirakulam (2nd year), CROMA Lab	Fall 2018
Jack Bandon (3rd year), CROMA Lab	Fall 2019
Utkarsh Mehta (3rd year), CROMA Lab	Fall 2019
Kejia Yang (Masters), CROMA Lab	2018-2019
Andrew Vernier (4th year), CROMA Lab	2019-2020
Yiming Shi (4th year), CROMA Lab	2019-2020
Zihan Li (Masters), CROMA Lab	Fall 2019
Vignesh Athreya (Masters), Project Sapphire	Spring 2017
Jiyong Shim (3rd year), Berkeley AI Lab	2013-2014
Philip Su (3rd year), Intern at Naming Matters	Summer 2015

High School Mentoring / Outreach

Australian Informatics Olympiad	2007–2010
Gave lectures, prepared exams and autograded, prepared labs.	
National Computer Science School	2007, 2008, 2010
Led a group of 16 students, who learned python and built a search engine.	
Computational Linguistics Olympiads	2013, 2014, 2015, 2016
Set problems for the competition.	

Professional Service **Area Chair**

ACL: Association for Computational Linguistics	
Computational Social Science and Social Media	2020
Resources and Evaluation	2021
DSTC: Dialog System Technology Challenges, Task Organiser	2018, 2019

University Internal

DEI Committee: Inclusive Environment, Michigan	2020-2021
Hiring Committee for Research Investigator, Michigan Medical School	2017
Graduate Admissions Committee, Michigan Artificial Intelligence	2017
Graduate Admissions Committee, Berkeley Artificial Intelligence	2015
President, Berkeley Computer Science Graduate Student Association	2012–2013

Thesis Committee

Laura (Wendlandt) Burdick	Michigan, 2020
Sai R. Gouravajhala	Michigan, 2020
Janarthanan Rajendran	Michigan, 2020

Journal Reviewer

CL: Computational Linguistics	2020, 2021
TACL: Transactions of the Assoc. for Computational Linguistics	2016, 2020-2022
JAIR: Journal of Artificial Intelligence Research	2017
CSL: Computer Speech & Language	2017, 2019

Program Committee - Conferences

ACL: Assoc. for Computational Linguistics	2014 (2nd), 2016, 2017, 2018, 2019
NAACL: North American Chapter of the ACL	2015, 2016, 2018, 2019, 2021
EMNLP: Empirical Methods in NLP	2015, 2016, 2017, 2018, 2019, 2020
EACL: European Chapter of the ACL	2017, 2021
AACL: Asia-Pacific Chapter of the ACL	2020
CoNLL: Conf. on Natural Language Learning	2016, 2017, 2018, 2019, 2020
IJCNLP: International Joint Conference on Natural Language Processing	2017
AAAI: Association for the Advancement of Artificial Intelligence	2016, 2019
IJCAI: International Joint Conference on Artificial Intelligence	2019
CHI: Conference on Human Factors in Computing Systems	2019
CSCW: Computer-Supported Cooperative Work	2020, 2021
UIST: ACM User Interface Software and Technology Symposium	2019
WWW / WebConf: World Wide Web Conference	2016, 2019

Program Committee - Workshops

SRW: Student Research, ACL / NAACL	2014, 2015, 2016, 2018, 2019, 2020, 2021 (x2)
IWPT: International Conference on Parsing Technologies	2020
EthNLP: Ethics in NLP	2017
RepEval: Evaluating Vector Space Representations for NLP	2019
NLP-CSS: NLP and Computational Social Science	2019, 2020
W-NUT: Noisy User-generated Text	2018, 2019, 2020
Gen-Deep: New Forms of Generalization in Deep Learning and NLP	2018
DeepLo: Deep Learning for Low Resource NLP	2019
IntEx-SemPar: Interactive and Executable Semantic Parsing	2020
DaSH-LA: Data Science with Human-in-the-Loop: Language Advances	2021

Other Reviewer

Tapia Conference Poster Reviewer	2020
Tapia Doctoral Consortium Judge	2020

Mentor

SRW: Student Research Workshop, ACL / NAACL	2019, 2020
---	------------

References

Dan Klein

Professor, University of California, Berkeley
klein@cs.berkeley.edu

Rada Mihalcea

Professor, University of Michigan
mihalcea@umich.edu

Dragomir R. Radev

Professor, Yale University
dragomir.radev@yale.edu

Jason Eisner

Professor, Johns Hopkins University
jason@cs.jhu.edu