Jonathan K. Kummerfeld

jkummerf@umich.edu

Current   University of Michigan   September 2016 – Present
Postdoctoral Research Fellow in Computer Science Engineering.
CROMA Lab (2018–Present), Sapphire Project (2016–2018)

Education   PhD in Computer Science   August 2010–August 12th, 2016
"Algorithms for Identifying Syntactic Errors and Parsing with Graph Structured Output"
University of California, Berkeley
Adviser: Dan Klein

Bachelor of Science (Advanced) (Honours)   March 2006–May 29th, 2010
Class I with University Medal
"Adaptive Supertagging for Faster Parsing"
University of Sydney
Adviser: James R. Curran

Publications: Exploring the Value of Personalized Word Embeddings
Conference   C. Welch, J. K. Kummerfeld, V. Perez-Rosas, R. Mihalcea
Proceedings   **CoLing**: *International Conference on Computational Linguistics*, 2020

Inconsistencies in Crowdsourced Slot-Filling Annotations: A Typology and Identification
Methods
S. Larson, A. Cheung, A. Mahendran, K. Leach, J. K. Kummerfeld
CoLing: *International Conference on Computational Linguistics*, 2020

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

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

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

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

Crowdsourced Detection of Emotionally Manipulative Language
J. S. Huffaker, J. K. Kummerfeld, W. S. Lasecki, M. S. Ackerman
CHI: *Conference on Human Factors in Computing Systems*, 2020

No-Press Diplomacy: Modeling Multi-Agent Gameplay
P. Paquette, Y. Lu, S. Bocco, M. O. Smith, S. Ortiz-Gagné, J. K. Kummerfeld, J. Pineau,
S. Singh, A. Courville
NeurIPS: *Advances in Neural Information Processing Systems*, 2019

An Evaluation for Intent Classification and Out-of-Scope Prediction

S. Larson, A. Mahendran, J. J. Peper, C. Clarke, A. Lee, P. Hill, <u>J. K. Kummerfeld</u>,
K. Leach, M. A. Laurenzano, L. Tang and J. Mars
EMNLP: *Empirical Methods in Natural Language Processing*, 2019

A Large-Scale Corpus for Conversation Disentanglement
<u>J. K. Kummerfeld</u>, S. R. Gouravajhala, J. Peper, V. Athreya, C. Gunasekara, J. Ganhotra,
S. S. Patel, L. Polymenakos, W. S. Lasecki
ACL: *Association for Computational Linguistics*, 2019

SLATE: A Super-Lightweight Annotation Tool for Experts
<u>J. K. Kummerfeld</u>
ACL Demo: *Association for Computational Linguistics: Demonstrations*, 2019

Outlier Detection for Improved Data Quality and Diversity in Dialog Systems
S. Larson, A. Mahendran, A. Lee, <u>J. K. Kummerfeld</u>, P. Hill, M. Laurenzano, J. Hauswald,
L. Tang, J. Mars
NAACL: *North American Chapter of the Association for Computational Linguistics*, 2019

Look Who's Talking: Inferring Speaker Attributes from Personal Longitudinal Dialog
C. Welch, V. Perez-Rosas, <u>J. K. Kummerfeld</u>, R. Mihalcea
CICLing: *Computational Linguistics and Intelligent Text Processing*, 2019
<u>Best Student Paper Award</u>

Improving Text-to-SQL Evaluation Methodology
C. Finegan-Dollak*, <u>J. K. Kummerfeld</u>*, L. Zhang, K. R. D. Ramanathan, S. Sadasivam,
R. Zhang, D. Radev
ACL: *Association for Computational Linguistics*, 2018

Factors Influencing the Surprising Instability of Word Embeddings
L. Wendlandt, <u>J. K. Kummerfeld</u>, R. Mihalcea
NAACL: *North American Chapter of the Association for Computational Linguistics*, 2018

Effective Crowdsourcing for a New Type of Summarization Task
Y. Jiang, C. Finegan-Dollak, <u>J. K. Kummerfeld</u>, W. Lasecki
NAACL: *North American Chapter of the Association for Computational Linguistics*, 2018

Data Collection for Dialogue System: A Startup Perspective
Y. Kang, Y. Zhang, <u>J. K. Kummerfeld</u>, P. Hill, J. Hauswald, M. A. Laurenzano, L. Tang,
J. Mars
NAACL: *North American Chapter of the Association for Computational Linguistics*, 2018

World Knowledge for Semantic Parsing with Abstract Meaning Representation
C. Welch, <u>J. K. Kummerfeld</u>, S. Feng, R. Mihalcea
LREC: *Language Resources and Evaluation Conference*, 2018

Identifying Products in Online Cybercrime Marketplaces: A Dataset for Fine-grained
Domain Adaptation
G. Durrett, <u>J. K. Kummerfeld</u>, T. Berg-Kirkpatrick, R. S. Portnoff, S. Afroz, D. McCoy,
K. Levchenko, V. Paxson
EMNLP: *Empirical Methods in Natural Language Processing*, 2017

Understanding Task Design Trade-offs in Crowdsourced Paraphrase Collection
Y. Jiang, <u>J. K. Kummerfeld</u>, W. S. Laseck
ACL: *Association for Computational Linguistics*, 2017

Tools for Automated Analysis of Cybercriminal Markets
R. S. Portnoff, S. Afroz, G. Durrett, <u>J. K. Kummerfeld</u>, T. Berg-Kirkpatrick, D. McCoy,
K. Levchenko, V. Paxson

WWW: *International World Wide Web Conferenc*, 2017

An Empirical Analysis of Optimization for Max-Margin NLP
<u>J. K. Kummerfeld</u>, T. Berg-Kirkpatrick, D. Klein
EMNLP: *Empirical Methods in Natural Language Processing*, 2015

Error-Driven Analysis of Challenges in Coreference Resolution
<u>J. K. Kummerfeld</u>, D. Klein
EMNLP: *Empirical Methods in Natural Language Processing*, 2013

An Empirical Examination of Challenges in Chinese Parsing
<u>J. K. Kummerfeld</u>, D. Tse, J. R. Curran, D. Klein
ACL: *Association for Computational Linguistics*, 2013

Parser Showdown at the Wall Street Corral: An Empirical Investigation of Error Types
in Parser Output
<u>J. K. Kummerfeld</u>, D. Hall, J. R. Curran, D. Klein
EMNLP: *Empirical Methods in Natural Language Processing*, 2012

Robust Conversion of CCG Derivations to Phrase Structure Trees
<u>J. K. Kummerfeld</u>, D. Klein, J. R. Curran
ACL: *Association for Computational Linguistics*, 2012

Mention Detection: Heuristics for the OntoNotes annotations
<u>J. K. Kummerfeld</u>, M. Bansal, D. Burkett, D. Klein
CoNLL Shared Task: *Conference on Natural Language Learning Shared Task*, 2011

Faster Parsing by Supertagger Adaptation
<u>J. K. Kummerfeld</u>, J. Roesner, T. Dawborn, J. Haggerty, J. Curran, S. Clarke
ACL: *Association for Computational Linguistics*, 2010

Morphological Analysis Can Improve a CCG Parser for English
M. Honnibal, <u>J. K. Kummerfeld</u>, J. R. Curran
CoLing: *International Conference on Computational Linguistics*, 2010

Publications:
Journals

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

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

High-velocity Clouds in the Galactic All Sky Survey. I. Catalog
V. A. Moss, N. M. McClure-Griffiths, T. Murphy, D. J. Pisano, <u>J. K. Kummerfeld</u>, J. R. Curran
The Astrophysical Journal Supplement Series, 2013, 209:12

Spatiotemporal Hierarchy of Relaxation Events, Dynamical Heterogeneities, and Structural Reorganization in a Supercooled Liquid
R. Candelier, A. Widmer-Cooper, <u>J. K. Kummerfeld</u>, O. Dauchot, G. Biroli, P. Harrowell, D. R Reichman
PRL: *Physical Review Letters*, 2010, 105:135702

The Densest Packing of AB Binary Hard-Sphere Homogeneous Compounds across all Size Ratios
<u>J. K. Kummerfeld</u>, T. S. Hudson, P. Harrowell

The Journal of Physical Chemistry B, 2008, 112:10773–10776

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

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

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

The Eighth Dialog System Technology Challenge
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
NeurIPS Workshop: *Conversational AI*, 2019

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

Dialog System Technology Challenge 7
K. Yoshino, C. Hori, J. Perez, L. F. D'Haro, L. Polymenakos, C. Gunasekara, W. S. Lasecki, J. K. Kummerfeld, M. Galley, C. Brockett, J. Gao, B. Dolan, X. Gao, H. Alamari, T. K. Marks, D. Parikh, D. Batra
NeurIPS Workshop: *Conversational AI*, 2018

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

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

Grants, Awards, and Honors	**Bloomberg Data Science Research Grant** Co-PI on grant to develop adaptive annotation methods	2019
	Outstanding Postdoctoral Fellow Award For excellence in research, mentorship, teaching, leadership, and service	2018
	Michigan Institute for Data Science Music Initiative Co-PI on grant to develop models that map text to music (75k)	2018
	Allen Institute for AI Key Scientific Challenges Program Unrestricted funding to support research (10k)	2017
	Chair's Special Award For services to the EECS department during student admissions and recruitment	2013
	Outstanding Graduate Student Instructor Award	2011

University wide award in recognition of exceptional achievements as a teacher

General Sir John Monash Award   2009
"Australia's most prestigious postgraduate scholarships"

University Medal in Computer Science   2009
For outstanding results, awarded at the discretion of the university senate

Employment

Postdoctoral Research Fellow
University of Michigan   September 21st, 2016 – September 9th, 2021

Software Engineer – 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

Teaching, Mentoring, Invited Talks

Instructor
Artificial Intelligence, Berkeley   Summer 2014
Co-taught upper division course – gave lectures, managed TAs, and prepared exams.
Evaluation: 6.3 / 7 (top 11% of UCB EECS instructors with public results).

Teaching Assistant
Artificial Intelligence, Berkeley   Fall 2011
Held sections and office hours, prepared assignments, exams and sections.
Evaluation: 4.6 / 5.

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
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   2017, 2018
University of Macquarie, Centre for Language Technology   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   2012

Undergraduate & Masters Mentor

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

High School Outreach

Australian Informatics Olympiad	2007–2010

Gave lectures, prepared exams and autograders, prepared labs.

National Computer Science School	2007, 2008, 2010

Led a group of 16 high school 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
DSTC: Dialog System Technology Challenges, Task Organiser	2018, 2019

University Internal

Hiring Comittee 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
TACL: Transactions of the Association for Computational Linguistics	2016, 2020-2022
JAIR: Journal of Artificial Intelligence Research	2017
CSL: Computer Speech & Language	2017, 2019

Program Committee - Conferences

ACL: Association for Computational Linguistics	2014 (2nd), 2016, 2017, 2018, 2019
NAACL: North American Chapter of the ACL	2015, 2016, 2018, 2019
EMNLP: Empirical Methods in Natural Language Processing	2015, 2016, 2017, 2018,

2019, 2020

EACL: European Chapter of the ACL	2017, 2021
AACL: Asia-Pacific Chapter of the ACL	2020
CoNLL: Conference on Natural Language Learning	2016, 2017, 2018, 2019
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
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
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

Other Reviewer

Tapia Conference Poster Reviewer	2020
Tapia Doctoral Consortium Judge	2020

Mentor

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