

Chenhao Tan

Assistant Professor
University of Colorado
Department of Computer Science
ECES 118A
Boulder, CO, 80309

Email: chenhao@chenhaot.com
Homepage: <https://chenhaot.com>

Research Interests

Natural language processing and social interaction, computational social science, and human-centered machine learning.

Professional experience

Assistant Professor, 2017-current,
University of Colorado, Boulder, CO.

Research Associate, 2016-2017,
University of Washington, Seattle, WA
Supervisor: Noah Smith.

Education

Ph.D. in Computer Science, 2010-2016,
Cornell University, Ithaca, NY,
Advisor: Lillian Lee.

B.E. in Computer Science, 2006-2010,
Tsinghua University, Beijing, China,
Advisor: Jie Tang.

B.E. in Economics (second major), 2007-2010,
Tsinghua University, Beijing, China.

Honors and Awards

Selected to participate in Heidelberg Laureate Forum (2016)

OARU travel award for Heidelberg Laureate Forum (2016)

Facebook Fellowship (2015)

Telluride Residential Scholarship (2012-2016)

Yahoo! Key Scientific Challenges (2012)

KDD Student Travel Award (2011, 2010)

Best Undergraduate Thesis of Tsinghua University (Top 1 among ~160) (2010)

Excellent Undergraduate of Tsinghua University (Top 58 among 3144) (2010)

Excellent Undergraduate of Beijing (2010)

National Scholarship (Top 3 among ~160) (2009, 2008)

Grants

National Science Foundation (2018). BIGDATA: IA: Automating Analysis and Feedback to Improve Mathematics Teachers' Classroom Discourse (with Tamara Sumner, Jennifer Jacobs, James Martin, and Wayne Ward). Budget \$1,998,505.

Facebook (2018). Research gift: \$25,000.

NVIDIA GPU grant (2017).

University of Washington (2017). Postdoc research awards: Understand media bias through the lens of images (\$10,000).

Selected Media Coverage

Harvard Business Review (2018): Want Less-Biased Decisions? Use Algorithms.

NPR (2017): Change My View On Reddit Helps People Challenge Their Own Opinions.

New Scientist (2017): Phone metadata reveals where city migrants go and who they call.

You are not so smart podcasts (2016): Change My View.

Slate (2016): You Can't Win: How algorithms are changing the ways we argue.

The Washington Post (2016): How to change someone's mind, according to science.

The Washington Post (2016): How to win a Facebook argument, according to science.

New York Magazine (2016): A Subreddit Sparked a Scientific Inquiry Into How to Change Someone's Mind.

Daily Mail (2016): How to win ANY argument using science: Experts reveal which words to use and how one simple trick can help you get your own way.

The New York Times (2014): A 25-Question Twitter Quiz to Predict Retweets.

The Washington Post (2014): This is the perfect tweet, according to science.

The Physics arXiv Blog (2014): Data Mining Reveals How Wording Influences Tweet Propagation.

engadget (2014): Crafting the perfect tweet is one-third talent, two-thirds science.

The Daily Dot (2014): A data miner's guide to crafting the perfect tweet.

Publications

[Google scholar profile]

Vivian Lai and Chenhao Tan. *On Human Predictions with Explanations and Predictions of Machine Learning Models: A Case Study on Deception Detection*. In Proceedings of ACM FAT* Conference 2019.

Jason Shuo Zhang, Chenhao Tan, and Qin Lv “*This is why we play*”: *Characterizing Online Fan Communities of the NBA Teams*. In Proceedings of the Conference on Computer-Supported Cooperative Work and Social Computing (CSCW’2018).

Tal August, Nigini Oliveira, Chenhao Tan, Noah A. Smith, and Katharina Reinecke. *Framing Effect: The Choice of Slogans Used to Advertise Online Experiments Can Boost Recruitment and Lead to Sample Biases*. In Proceedings of the Conference on Computer-Supported Cooperative Work and Social Computing (CSCW’2018).

Dallas Card, Chenhao Tan, and Noah A. Smith. *Neural Models for Documents with Metadata*. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL’2018).

Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan, and Noah A. Smith. *LSTMs Exploit Linguistic Attributes of Data*. In Proceedings of ACL 2018 RepL4NLP Workshop (**Best paper award**).

Chenhao Tan. *Tracing Community Genealogy: How New Communities Emerge from the Old*. In Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM’18).

Chenhao Tan, Hao Peng and Noah A. Smith. “*You are no Jack Kennedy*”: *On Media Selection of Highlights from Presidential Debates*. In Proceedings of The Web Conference (WWW’2018).

Yang Yang, Zongtao Liu, Chenhao Tan, Fei Wu, Yueting Zhuang and Yafeng Li. *To Stay or to Leave: Churn Prediction for Urban Migrants in the Initial Period*. In Proceedings of The Web Conference (WWW’2018).

Isil Erel, Léa H. Stern, Chenhao Tan, and Michael S. Weisbach. *Could Machine Learning Help Companies Select Better Board Directors?* Harvard Business Review. April, 2018.

Elizabeth Clark, Anne Ross, Chenhao Tan, Yangfeng Ji, Noah A. Smith. *Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories*. In Proceedings of the 23rd ACM Conference on Intelligent User Interfaces (IUI’2018).

Yang Yang, Chenhao Tan, Zongtao Liu, Fei Wu, Yueting Zhuang. *Urban Dreams of Migrants: A Case Study of Migrant Integration in Shanghai*. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI’2018).

Yangfeng Ji, Chenhao Tan, Sebastian Martschat, Yejin Choi, Noah A. Smith. *Dynamic Entity Representations in Neural Language Models*. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP’2017).

Chenhao Tan, Dallas Card, Noah A. Smith. *Friendships, Rivalries, and Trysts: Characterizing Relations between Ideas in Texts*. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL’2017).

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, Lillian Lee. *Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions*. In Proceedings of the 25th International World Wide Web Conference (WWW’2016).

Isabel Kloumann, Chenhao Tan, Jon Kleinberg, Lillian Lee. *Internet Collaboration on Extremely Difficult Problems: Research versus Olympiad Questions on the Polymath Site*. In Proceedings of the 25th International World Wide Web Conference (WWW’2016).

Chenhao Tan, Adrien Friggeri, Lada Adamic. *Lost in propagation? Unfolding news cycles from the source*. In Proceedings of 10th International AAAI Conference on Weblogs and Social Media (ICWSM'2016).

Jack Hessel, Chenhao Tan, Lillian Lee. *Science, AskScience, and BadScience: On the Coexistence of Highly Related Communities*. In Proceedings of 10th International AAAI Conference on Weblogs and Social Media (ICWSM'2016).

Chenhao Tan, Lillian Lee. *All Who Wander: On the Prevalence and Characteristics of Multi-community Engagement*. In Proceedings of the 24th International World Wide Web Conference (WWW'2015).

Chenhao Tan, Lillian Lee, Bo Pang. *The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter*. In Proceedings of the Association for Computational Linguistics (ACL'2014).

Chenhao Tan, Lillian Lee. *A Corpus of Sentence-level Revisions in Academic Writing: A Step towards Understanding Statement Strength in Communication*. In Proceedings of the Association for Computational Linguistics (ACL'2014) (short paper).

Jon Kleinberg, Sendhil Mullainathan, Chenhao Tan, Thomas Zimmermann. *Inductive Testing: Theory and an Application to the Disposition Effect*. NBER Behavioral Finance Working Group, 2014.

Chenhao Tan, Ed H. Chi, David Huffaker, Gueorgi Kossinets and Alexander J. Smola. *Instant Foodie: Predicting Expert Ratings From Grassroots*. In Proceedings of 22nd ACM International Conference on Information and Knowledge Management (CIKM'2013).

Daniel M. Romero, Chenhao Tan and Johan Ugander. *On the Interplay between Social and Topical Structure*. In Proceedings of 7th International AAAI Conference on Weblogs and Social Media (ICWSM'2013).

Eunsol Choi, Chenhao Tan, Lillian Lee, Cristian Danescu-Niculescu-Mizil and Jennifer Spindel. *Hedge detection as a lens on framing in the GMO debates: A position paper*. In Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics (2012).

Chenhao Tan, Evgeniy Gabrilovich, and Bo Pang. *To Each His Own: Personalized Content Selection based on Text Comprehensibility*. In Proceedings of the Fifth ACM International Conference on Web Search and Data Mining (WSDM'2012).

Wenbin Tang, Jie Tang, Tao Lei, Chenhao Tan, Bo Gao, and Tian Li. *On optimization of expertise matching with various constraints*. In Neurocomputing (July, 2012).

Chenhao Tan, Lillian Lee, Jie Tang, Long Jiang, Ming Zhou, and Ping Li. *User-level sentiment analysis incorporating social networks*. In Proceedings of the Seventeenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'2011).

Chenhao Tan, Jie Tang, Jimeng Sun, Quan Lin, and Fengjiao Wang. *Social Action Tracking via Noise Tolerant Time-varying Factor Graphs*. In Proceedings of the Sixteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'2010).

Bin Lu, Chenhao Tan, Claire Cardie and Benjamin K. Tsou. *Joint Bilingual Sentiment Classification with Unlabeled Parallel Corpora*. In Proceedings of the Association for Computational Linguistics (ACL'2011) (oral paper).

Shaomei Wu, Chenhao Tan, Jon Kleinberg, and Michael Macy. *Does Bad News Go Away Faster?* In Proceedings of Fifth International AAAI Conference on Weblogs and Social Media (ICWSM'2011) (short paper).

Fei Wang, *Chenhao Tan*, Arnd Christian Konig, and Ping Li. *Efficient Document Clustering via Online Nonnegative Matrix Factorizations*. In Proceedings of the SIAM International Conference on Data Mining (SDM'2011).

Wenbin Tang, Jie Tang, and *Chenhao Tan*. *Expertise Matching via Constrained-based Optimization*. In Proceedings of 2010 IEEE/WIC/ACM International Conferences (WIC'2010).

Invited talks

The Future of Brand Engagement: Community, Trust and Digital Discourse, Reddit, 2018.

Characterizing the Ecosystem of Ideas, Summer Institute in Computational Social Science, University of Colorado Boulder, 2018.

Towards Understanding and Generating Language for Effective Social Interaction, IST Seminar, Caltech, 2018.

Winning Arguments: Interaction Dynamics and Persuasion Strategies in Good-Faith online discussions, the 24th Annual Northwest Dispute Resolution Conference, 2018.

Phrasing Matters: An Interactive Session on Message Virality Based on Twitter, the 24th Annual Northwest Dispute Resolution Conference, 2018.

Characterizing the Ecosystems of Ideas in Texts, Department of Computer Science, Tsinghua University, 2018.

Wandering Individuals and Emerging Communities in a World of Communities, Department of Sociology, Tsinghua University, 2018.

It Takes Two to Tango: Understanding the Effects of Wording through Online Interactions, Fudan University, 2018.

It Takes Two to Tango: Understanding the Effects of Wording through Online Interactions, Toutiao AI lab, 2018.

It Takes Two to Tango: Understanding the Effects of Wording through Online Interactions, Harbin Institute of Technology, 2018.

It Takes Two to Tango: Understanding the Effects of Wording through Online Interactions, Zhejiang University, 2017.

A Short Talk on Research, Zhejiang University, 2017.

Wandering Individuals and Emerging Communities in a World of Communities, INFO/HCC Seminar at University of Colorado Boulder, 2017.

Characterizing the Ecosystems of Ideas in Texts, Computational Semantics Group at University of Colorado Boulder, 2017.

Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions, Colorado Data Science Team at University of Colorado Boulder, 2017.

Characterizing the Ecosystems of Ideas in Texts, UW-MS NLP Symposium, 2017.

Online social interactions: a lens on humans and a world for humans, CSSS Seminar at University of Washington, 2016.

Online social interactions: a lens on humans and a world for humans, MIT EECS, 2016.

Online social interactions: a lens on humans and a world for humans, CMU Computer Science Department, 2016.

Online social interactions: a lens on humans and a world for humans, CMU Heinz College, 2016.

Online social interactions: a lens on humans and a world for humans, Max Planck Institute for Software Systems, 2016.

Online social interactions: a lens on humans and a world for humans, University of Michigan, 2016.

Online social interactions: a lens on humans and a world for humans, University of Colorado Boulder, 2016.

Online social interactions: a lens on humans and a world for humans, Rice University, 2016.

Online social interactions: a lens on humans and a world for humans, Northeastern University, 2016.

It takes two to tango: Understanding the effects of language via "natural experiments", AI seminar at Cornell University, 2014

The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter, Tsinghua University, 2014

More than Accuracy: Interpretability, Machine learning reading group at Cornell University, 2013

User-level sentiment analysis incorporating social networks, Tsinghua University, 2011

User-level sentiment analysis incorporating social networks, Microsoft Research Asia NLP group seminar, 2011

Industrial Experience

Data Scientist Intern, 2015.6-2015.8,
Facebook, Menlo Park, CA,
Understanding news propagation.

Research Intern, 2012.5-2012.8,
Google, Mountain View, CA
Harmonizing place scores & instant foodie.

Research Intern, 2011.6-2011.8,
Yahoo! Research, Santa Clara, CA
Personalization based on text comprehensibility.

Research Intern, 2010.4-2010.7,
Microsoft Research Asia, Beijing, China,
Sentiment analysis incorporating social networks.

Teaching Experience

Instructor, 2018.1-current

University of Colorado, Boulder, CO,

Human-centered Machine Learning, CSCI 7000, Spring 2017.

Instructor, 2017.8-2017.12

University of Colorado, Boulder, CO,

Machine Learning, CSCI 5622, Fall 2017.

Teaching Assistant, 2013.9-2013.12,

Cornell University, Ithaca, NY,

Information Retrieval, CS4300/INFO4300, Fall 2013.

Teaching Assistant, 2011.9-2011.12,

Cornell University, Ithaca, NY,

Machine Learning, CS4780/5780, Fall 2011.

PhD Students

Vivian Lai.

Shuo Zhang (co-advised with Qin Lv).

Nora Connor (committee).

Forough Poursabzi (committee).

Master Advising

Vivian Lai (class of 2018, University of Colorado Boulder).

Linzi Xing (class of 2018, University of Colorado Boulder).

Hayeong Song (class of 2018, University of Colorado Boulder).

Anupama (class of 2019, University of Colorado Boulder).

Saumya Sinha (class of 2019, University of Colorado Boulder)

Brian Lubars (class of 2019, University of Colorado Boulder).

Shantanu Karnwal (class of 2019, University of Colorado Boulder).

Hunter Wapman (class of 2019, University of Colorado Boulder).

Kumar Srinivasan (class of 2019, University of Colorado Boulder).

Undergraduate Advising

Eunsol Choi (class of 2012, Cornell, honorable mention for the CRA Outstanding Undergraduate Research Award in 2011), now a PhD student at the University of Washington.

Kelvin Luu (class of 2015, Cornell), now a PhD student at the University of Washington.

Cora Schneck (class of 2018, University of Colorado Boulder).

Tyler Scott (class of 2018, University of Colorado Boulder).

Xiaochuang Han (class of 2019, Georgia Tech).

David Atkinson (class of 2019, University of Colorado Boulder).

Professional Service

Area Chair:

2019: NAACL

PC Member:

2019: WWW, FAT*

2018: WWW, NAACL, ACL, KDD, NIPS, ICWSM, BigNet Workshop

2017: WWW, WSDM, KDD, ACL, EMNLP, CONLL

2016: WWW, KDD, WSDM, IJCAI, NAACL, ICWSM, SDM, CIKM

2015: WWW, KDD, WSDM

2014: WWW, EMNLP, ICWSM, DYAD

2013: IJCAI, IJCNLP, NAACL Student Research Workshop

2012: ADMA

2011: ADMA

Session chair: WWW (2015)

Reviewer: ICWSM (2019), TACL (2017, 2016), Transactions on Knowledge Discovery from Data (2015, 2012), CHI (2017), CSCW (2016), WWW (2011, 2010), ICDM (2011), KDD (2010), CIKM (2010)