Xiaojin (Jerry) Zhu

Contact

jerryzhu@cs.wisc.edu

Department of Computer Sciences
University of Wisconsin-Madison
1205 University Avenue, Madison, WI 53706

Research Interests

Machine learning

Education

PhD in Language Technologies — May, 2005
Carnegie Mellon University, Pittsburgh, PA
Dissertation: Semi-Supervised Learning with Graphs
Advisors: John Lafferty, Ronald Rosenfeld

MS in Knowledge Discovery and Data Mining — December, 2002
Carnegie Mellon University, Pittsburgh, PA

MS in Language and Information Technologies — May, 2000
Carnegie Mellon University, Pittsburgh, PA

MS in Computer Science — March, 1996
Shanghai Jiao Tong University, Shanghai, China

BS in Computer Science — July, 1993
Shanghai Jiao Tong University, Shanghai, China

Professional Positions

Professor — 2016–present
Department of Computer Sciences
University of Wisconsin–Madison. Madison, WI, USA

Associate Professor — 2011–2016
Department of Computer Sciences
University of Wisconsin–Madison. Madison, WI, USA

Assistant Professor — 2005–2011
Department of Computer Sciences
University of Wisconsin–Madison. Madison, WI, USA

Research Scientist — 1996–1998
IBM China Research Laboratory. Beijing, China

Awards and Honors

2025 Vilas Distinguished Achievement Professor, University of Wisconsin-Madison

2024 Kellett Mid-Career Award, University of Wisconsin-Madison

2021 Stephen C. Kleene Professorship, University of Wisconsin-Madison

2020 Vilas Associate Award, University of Wisconsin-Madison

2017 Vilas Faculty Mid-Career Investigator Award, University of Wisconsin-Madison

2016 Sheldon & Marianne Lubar Professorship, University of Wisconsin-Madison

2015 AAAI / Computing Community Consortium "Blue Sky Ideas" Track Prize for the paper "Machine Teaching: an Inverse Problem to Machine Learning and an Approach Toward Optimal Education" by Xiaojin Zhu.

2013 Classic Paper Prize, "Semi-supervised learning using Gaussian fields and harmonic functions" originally published in 2003 by Xiaojin Zhu, Zoubin Ghahramani and John Lafferty, International Conference on Machine Learning (ICML),

2012 Best Paper on Knowledge Discovery, "Socioscope: Spatio-Temporal Signal Recovery from Social Media" by Jun-Ming Xu, Aniruddha Bhargava, Robert Nowak, and Xiaojin Zhu. The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)

2012 The SACM "COW" Student Choice Professor of the Year Award. This award goes annually to the best teaching faculty in the department as voted on by CS undergraduate and graduate students.

2011 ACM SIGSOFT Distinguished Paper Award, "Recovering the Toolchain Provenance of Binary Code" by Nathan Rosenblum, Barton P. Miller and Xiaojin Zhu. International Symposium on Software Testing and Analysis (ISSTA)

2010 National Science Foundation Faculty Early Career Development (CAREER) award

2000 Microsoft Research Graduate Fellowship

1998 IBM Research Division Award

1997 IBM First Patent Application Invention Achievement Award

1997 IBM Greater China Group Team Award

Publications

Co-Authors noted as (s) for student under my direction, (p) for post-doctoral associate under my direction, (o) for students or post-docs under the direction of others, and (a) for my thesis Advisors.

Books and Book Chapters

Xiaojin Zhu and Andrew B. Goldberg$^{(s)}$. *Introduction to Semi-Supervised Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, San Rafael, CA, 2009.

Xiaojin Zhu. Semi-supervised learning. In Claude Sammut and Geoffrey Webb, editors, *Encyclopedia of Machine Learning*. Springer, first edition, 2010.

Xiaojin Zhu, Jaz Kandola, John Lafferty$^{(a)}$, and Zoubin Ghahramani. Graph kernels by spectral transforms. In O. Chapelle, B. Schölkopf, and A. Zien, editors, *Semi-Supervised Learning*. MIT Press, 2006.

Journal Papers

Yun-Shiuan Chuang, Xiaojin Zhu, and Timothy T. Rogers. The delusional hedge algorithm as a model of human learning from diverse opinions. *Topics in Cognitive Science*, 2025.

Robert M. Nosofsky, Craig A. Sanders, Xiaojin Zhu, and Mark A. McDaniel. Model-guided search for optimal natural- science-category training exemplars: A work in progress. *Psychonomic Bulletin & Review*, pages 1–29, 2018.

Ji Liu and Xiaojin Zhu. The teaching dimension of linear learners. *Journal of Machine Learning Research*, 17(162):1–25, 2016.

Anna Kaatz, You-Geon Lee, Aaron Potvien, Wairimu Magua, Amarette Filut, Anupama Bhattacharya, Renee Leatherberry, Xiaojin Zhu, and Molly Carnes. Analysis of national institutes of health r01 application critiques, impact, and criteria scores: Does the sex of the principal investigator make a difference? *Academic medicine: journal of the Association of American Medical Colleges*, 91(8):1080–8, 2016.

Felice Resnik, Amy Bellmore, Junming Xu(s), and Xiaojin Zhu. Celebrities emerge as advocates in tweets about bullying. *Translational Issues in Psychological Science*, 2016.

Amy Bellmore, Angela Calvin, Jun-Ming Xu, and Xiaojin Zhu. The five W's of bullying on Twitter: Who, what, why, where, when. *Computers in Human Behavior*, 2014. accepted.

Charles Kalish, Xiaojin Zhu, and Timothy Rogers. Drift in children's categories: When experienced distributions conflict with prior learning. *Developmental Science*, 2014. accepted.

Angela J. Calvin(o), Amy Bellmore, Jun-Ming Xu(s), and Xiaojin Zhu. #bully: Uses of hashtags in posts about bullying on Twitter. *Journal of School Violence*, 2014. accepted.

Mark Liu(o), Mutlu Ozdogan, and Xiaojin Zhu. Crop type classification by simultaneous use of satellite images of different resolutions. *Geoscience and Remote Sensing, IEEE Transactions on*, 52(6):3637–3649, June 2014.

Bryan R. Gibson(s), Timothy T. Rogers, and Xiaojin Zhu. Human semi-supervised learning. *Topics in Cognitive Science*, 5(1):132–172, 2013.

Jun-Ming Xu(s), Xiaojin Zhu, and Timothy T. Rogers. Metric learning for estimating psychological similarities. *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 2011.

Charles W. Kalish, Timothy T. Rogers, Jonathan Lang, and Xiaojin Zhu. Can semi-supervised learning explain incorrect beliefs about categories? *Cognition*, 2011.

Arthur Glenberg, Jonathan Willford(o), Bryan Gibson(s), Andrew Goldberg(s), and Xiaojin Zhu. Improving reading to improve math. *Scientific Studies in Reading*, 2011.

Arthur Glenberg, Andrew Goldberg(s), and Xiaojin Zhu. Improving early reading comprehension using embodied CAI. *Instructional Science*, 39:27–39, 2011.

Ronald Rosenfeld(a), Stanley Chen, and Xiaojin Zhu. Whole-sentence exponential language models: a vehicle for linguistic-statistical integration. *Computers Speech and Language*, 15(1):55–73, 2001.

Refereed Conference Papers

Hyeong Kyu Choi(o), Xiaojin Zhu, and Sharon Li. When identity skews debate: Anonymization for bias-reduced multi-agent reasoning. In *Annual Meeting of the Association for Computational Linguistics (ACL Main)*, 2026. Oral presentation.

Jason Jisen Li(s), Young Wu, Yancheng Zhu(s), Jin yi Cai, and Xiaojin Zhu. Pure nash equilibria under the affine mechanism: A potential game of exaggeration. In *The 17th Conference on Game Theory and AI for Security (GameSec)*, 2026.

Brandon Han(s), Young Wu, Shiyun Cheng(s), and Xiaojin Zhu. The game changer problem: Controlling equilibria with discrete rewards. In *The 17th Conference on Game Theory and AI for Security (GameSec)*, 2026.

Hyeong Kyu Choi(o), Xiaojin Zhu, and Yixuan Li. Debate or vote: Which yields better decisions in multi-agent large language models? In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025. (acceptance rate 24.52% out of 21575 submissions. Spotlight).

Alex Clinton(o), Thomas Zeng(o), Yiding Chen(s), Xiaojin Zhu, and Kirthevasan Kandasamy. A Cramér–von Mises approach to incentivizing truthful data sharing. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025. (acceptance rate 24.52% out of 21575 submissions).

Alex Clinton[o], Yiding Chen[s], Xiaojin Zhu, and Kirthevasan Kandasamy. Collaborative mean estimation among heterogeneous strategic agents: Individual rationality, fairness, and truthful contribution. In *The 42nd International Conference on Machine Learning (ICML)*, 2025. (acceptance rate 3260/12107=26.9%).

Shubham Bharti[s], Stephen Wright, Adish Singla, and Xiaojin Zhu. On the complexity of teaching a family of linear behavior cloning learners. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024. (acceptance rate 25.8% out of 15671 submissions).

Jeremy McMahan[s], Young Wu[s], Yudong Chen, Xiaojin Zhu, and Qiaomin Xie. Inception: Efficiently computable misinformation attacks on Markov games. In *The first Reinforcement Learning Conference (RLC)*, 2024.

Young Wu[s], Jeremy McMahan[s], Yiding Chen[s], Yudong Chen, Xiaojin Zhu, and Qiaomin Xie. Minimally modifying a markov game to achieve any nash equilibrium and value. In *The 41st International Conference on Machine Learning (ICML)*, 2024. (acceptance rate 2609/9473=27.5%).

Jihyun Rho[o], Martina Rau, Shubham Bharti, Rosanne Luu[o], Jeremy McMahan, Andrew Wang, and Xiaojin Zhu. Various misleading visual features in misleading graphs: Do they truly deceive us? In *The Annual Conference of the Cognitive Science Society (CogSci)*, 2024. (acceptance rate (236+663)/1242=72%).

Yun-Shiuan Chuang[o], Xiaojin Zhu, and Timothy Rogers. The delusional hedge algorithm as a model of human learning from diverse opinions. In *The Annual Conference of the Cognitive Science Society (CogSci)*, 2024. (**Winner of the Computational Modeling Prize for Higher-Level Cognition for CogSci 2024**. Oral, acceptance rate 236/1242=19%).

Ara Vartanian[s], Xiaoxi Sun[s], Yun-Shiuan Chuang[o], Siddharth Suresh[o], Xiaojin Zhu, and Timothy Rogers. Learning interactions to boost human creativity with bandits and GPT-4. In *The Annual Conference of the Cognitive Science Society (CogSci)*, 2024. (acceptance rate (236+663)/1242=72%).

Jeremy McMahan[s] and Xiaojin Zhu. Anytime-constrained reinforcement learning. In *The 27th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024. (acceptance rate 546/1980=28%).

Young Wu[s], Jeremy McMahan[s], Xiaojin Zhu, and Qiaomin Xie. Data poisoning to fake a nash equilibrium in markov games. In *The Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI)*, 2024. (acceptance rate 2342/9862=24%).

Yiding Chen[s], Xuezhou Zhang[s], Qiaomin Xie, and Xiaojin Zhu. Exact policy recovery in offline rl with both heavy-tailed rewards and data corruption. In *The Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI)*, 2024. (acceptance rate 2342/9862=24%).

Jeremy McMahan[s], Young Wu[s], Xiaojin Zhu, and Qiaomin Xie. Optimal attack and defense on reinforcement learning. In *The Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI)*, 2024. (acceptance rate 2342/9862=24%).

Yiding Chen[s], Xiaojin Zhu, and Kirthevasan Kandasamy. Mechanism design for collaborative normal mean estimation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. (acceptance rate 3222/12343=26.1%).

Xuefeng Du[o], Yiyou Sun[o], Xiaojin Zhu, and Yixuan Li. Dream the impossible: Outlier imagination with diffusion models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. (acceptance rate 3222/12343=26.1%).

Leitian Tao[o], Xuefeng Du[o], Xiaojin Zhu, and Yixuan Li. Non-parametric outlier synthesis. In *The 11th International Conference on Learning Representations (ICLR)*, 2023. (acceptance rate 31.8% out of 5000 submissions).

Yiding Chen[s], Xuezhou Zhang[s], Kaiqing Zhang[o], Mengdi Wang, and Xiaojin Zhu. Byzantine-robust online and offline distributed reinforcement learning. In *The 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023. (acceptance rate 29% out of 1689 submissions).

Young Wu[s], Jeremy McMahan[s], Xiaojin Zhu, and Qiaomin Xie. Reward poisoning attacks on offline multi-agent reinforcement learning. In *The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI)*, 2023. (acceptance rate 1721/8777=19.6%).

Shubham Kumar Bharti[s], Xuezhou Zhang, Adish Singla, and Xiaojin Zhu. Provable defense against backdoor policies in reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. (acceptance rate 2665/10411=25.6%).

Yiyou Sun[o], Yifei Ming[o], Xiaojin Zhu, and Yixuan Li. Out-of-distribution detection with deep nearest neighbors. In *The 39th International Conference on Machine Learning (ICML)*, 2022. (acceptance rate 1235/5630=21.9%).

Yuzhe Ma[s], Young Wu[s], and Xiaojin Zhu. Game redesign in no-regret game playing. In *The 31st International Joint Conference on Artificial Intelligence and the 25th European Conference on Artificial Intelligence (IJCAI-ECAI 22)*, 2022. (acceptance rate 15% out of 4535 submissions).

Xuezhou Zhang[s], Yiding Chen[s], Xiaojin Zhu, and Wen Sun. Corruption-robust offline reinforcement learning. In *The 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2022. (acceptance rate 492/1685=29%).

Xuezhou Zhang[s], Yiding Chen[s], Xiaojin Zhu, and Wen Sun. Robust policy gradient against strong data corruption. In *The 38th International Conference on Machine Learning (ICML)*, 2021. (acceptance rate 1184/5513=21.5%).

Yun-Shiuan Chuang[o], Xuezhou Zhang[s], Yuzhe Ma[s], Mark Ho, Joe Austerweil, and Xiaojin Zhu. Using machine teaching to investigate human assumptions when teaching reinforcement learners. In *The 41st Annual Conference of the Cognitive Science Society (CogSci)*, 2021. (acceptance rate 484/710=68%).

Claudia Ramly[o], Ayon Sen[s], Ved P. Kale[s], Martina A. Rau, and Xiaojin Zhu. Digitally training graph viewers against misleading bar charts. In *The 41st Annual Conference of the Cognitive Science Society (CogSci)*, 2021. (acceptance rate 484/710=68%).

Xuezhou Zhang[s], Shubham Bharti[s], Yuzhe Ma[s], Adish Singla, and Xiaojin Zhu. The sample complexity of teaching by reinforcement on q-learning. In *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*, 2021. (acceptance rate 1692/7911=21%).

Yuzhe Ma[s], Jon Sharp[o], Ruizhe Wang[o], Earlence Fernandes, and Xiaojin Zhu. Adversarial attacks on kalman filter-based forward collision warning systems. In *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*, 2021. (acceptance rate 1692/7911=21%).

Ayon Sen[s], Xiaojin Zhu, Erin Marshall[o], and Robert Nowak. Popular imperceptibility measures in visual adversarial attacks are far from human perception. In *Conference on Decision and Game Theory for Security (GameSec)*, 2020.

Xuezhou Zhang[s], Yuzhe Ma[s], Adish Singla, and Xiaojin Zhu. Adaptive reward-poisoning attacks against reinforcement learning. In *The 37th International Conference on Machine Learning (ICML)*, 2020. (acceptance rate 1088/4990=21.8%).

Amin Rakhsha[(o)], Goran Radanovic[(o)], Rati Devidze[(o)], Xiaojin Zhu, and Adish Singla. Policy teaching via environment poisoning: Training-time adversarial attacks against reinforcement learning. In *The 37th International Conference on Machine Learning (ICML)*, 2020. (acceptance rate 1088/4990=21.8%).

Xuezhou Zhang[(s)], Xiaojin Zhu, and Laurent Lessard. Online data poisoning attacks. In *Learning for Dynamics and Control (L4DC)*, 2020. (oral rate 14/131=11%).

Yiding Chen[(s)] and Xiaojin Zhu. Optimal attack against autoregressive models by manipulating the environment. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, 2020. (acceptance rate 1591/7737=20.6%).

Yuzhe Ma[(s)], Xuezhou Zhang[(s)], Wen Sun, and Xiaojin Zhu. Policy poisoning in batch reinforcement learning and control. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019. (acceptance rate 1428/6743=21%).

Xuanqing Liu, Si Si, Xiaojin Zhu, Yang Li, and Cho-Jui Hsieh. A unified framework for data poisoning attack to graph-based semi-supervised learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019. (acceptance rate 1428/6743=21%).

Farnam Mansouri, Yuxin Chen, Ara Vartanian[(s)], Xiaojin Zhu, and Adish Singla. Preference-based batch and sequential teaching: Towards a unified view of models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019. (acceptance rate 1428/6743=21%).

Yuzhe Ma[(s)], Xiaojin Zhu, and Justin Hsu. Data poisoning against differentially-private learners: Attacks and defenses. In *The 28th International Joint Conference on Artificial Intelligence (IJCAI)*, 2019. (acceptance rate 850/4752=18%).

Sanjoy Dasgupta, Daniel Hsu, Stefanos Poulis, and Xiaojin Zhu. Teaching a black-box learner. In *The 36th International Conference on Machine Learning (ICML)*, 2019. (acceptance rate 773/3424=23%).

Laurent Lessard, Xuezhou Zhang[(s)], and Xiaojin Zhu. An optimal control approach to sequential machine teaching. In *The 22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019. (acceptance rate 360/1111=32%).

Kwang-Sung Jun[(s)], Lihong Li, Yuzhe Ma[(s)], and Xiaojin Zhu. Adversarial attacks on stochastic bandits. In *Advances in Neural Information Processing Systems (NIPS)*, 2018. (acceptance rate 1011/4856=21%).

Ayon Sen[(s)], Scott Alfeld[(s)], Xuezhou Zhang[(s)], Ara Vartanian[(s)], Yuzhe Ma[(s)], and Xiaojin Zhu. Training set camouflage. In *Conference on Decision and Game Theory for Security (GameSec)*, 2018.

Yuzhe Ma[(s)], Kwang-Sung Jun[(s)], Lihong Li, and Xiaojin Zhu. Data poisoning attacks in contextual bandits. In *Conference on Decision and Game Theory for Security (GameSec)*, 2018.

Ayon Sen[(s)], Purav Patel, Martina A. Rau, Blake Mason, Robert Nowak, Timothy T. Rogers, and Xiaojin Zhu. Machine beats human at sequencing visuals for perceptual-fluency practice. In *Educational Data Mining*, 2018. (long paper, acceptance rate 23/145=16%).

Ayon Sen[(s)], Purav Patel, Martina A. Rau, Blake Mason, Robert Nowak, Timothy T. Rogers, and Xiaojin Zhu. For teaching perceptual fluency, machines beat human experts. In *The 40th Annual Conference of the Cognitive Science Society (CogSci)*, 2018. (oral, acceptance rate 215/691=31%).

Yuzhe Ma[(s)], Robert Nowak, Philippe Rigollet, Xuezhou Zhang[(s)], and Xiaojin Zhu. Teacher improves learning by selecting a training subset. In *The 21st International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2018. (acceptance rate 214/645=33%).

Xuezhou Zhang[(s)], Xiaojin Zhu, and Stephen Wright. Training set debugging using trusted items. In *The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI)*, 2018. (acceptance rate 933/3800=25%).

Vraj Shah, Arun Kumar, and Xiaojin Zhu. Are key-foreign key joins safe to avoid when learning high-capacity classifiers? In *VLDB*, 2018.

Xiaojin Zhu, Ji Liu, and Manuel Lopes. No learner left behind: On the complexity of teaching multiple learners simultaneously. In *The 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 2017. (acceptance rate 660/2540=26%).

Scott Alfeld[(s)], Xiaojin Zhu, and Paul Barford. Explicit defense actions against test-set attacks. In *The Thirty-First AAAI Conference on Artificial Intelligence (AAAI)*, 2017. (acceptance rate 638/2590=25%).

Paul Bennett, David M. Chickering, Christopher Meek, and Xiaojin Zhu. Algorithms for active classifier selection:maximizing recall with precision constraints. In *The Tenth ACM International Conference on Web Search and Data Mining (WSDM)*, 2017. (acceptance rate 80/505=16%).

Tzu-Kuo Huang, Lihong Li, Ara Vartanian[(s)], Saleema Amershi, and Xiaojin Zhu. Active learning with oracle epiphany. In *Advances in Neural Information Processing Systems (NIPS)*, 2016. (acceptance rate 568/2500=23%).

Ji Liu, Xiaojin Zhu, and H. Gorune Ohannessian[(s)]. The teaching dimension of linear learners. In *The 33rd International Conference on Machine Learning (ICML)*, 2016. (acceptance rate 322/1327=24%).

Jina Suh, Xiaojin Zhu, and Saleema Amershi. The label complexity of mixed-initiative classifier training. In *The 33rd International Conference on Machine Learning (ICML)*, 2016. (acceptance rate 322/1327=24%).

Xiaojin Zhu, Ara Vartanian[(s)], Manish Bansal[(s)], Duy Nguyen[(o)], and Luke Brandl[(s)]. Stochastic multiresolution persistent homology kernel. In *The 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 2016. (acceptance rate =25%).

Kwang-Sung Jun[(s)], Kevin Jamieson[(o)], Rob Nowak, and Xiaojin Zhu. Top arm identification in multi-armed bandits with batch arm pulls. In *The 19th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016. (acceptance rate 165/537=31%).

Scott Alfeld[(s)], Xiaojin Zhu, and Paul Barford. Data poisoning attacks against autoregressive models. In *The Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, 2016. (acceptance rate 549/2132=26%).

Arun Kumar[(o)], Jeffrey Naughton, Jignesh M. Patel, and Xiaojin Zhu. To join or not to join? thinking twice about joins before feature selection. In *ACM SIGMOD*, 2016.

Newsha Ardalani[(o)], Clint Lestourgeon[(o)], Karthikeyan Sankaralingam, and Xiaojin Zhu. Cross-architecture performance prediction (XAPP) using CPU code to predict GPU performance. In *Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-48)*, 2015.

Kwang-Sung Jun[(s)], Xiaojin Zhu, Timothy Rogers, Zhuoran Yang[(s)], and Ming Yuan. Human memory search as initial-visit emitting random walk. In *Advances in Neural Information Processing Systems (NIPS)*, 2015. (acceptance rate 403/1838=22%).

Gautam Dasarathy[(o)], Robert Nowak, and Xiaojin Zhu. s^2: An efficient graph based active learning algorithm with application to nonparametric classification. In *Conference on Learning Theory (COLT)*, 2015. (acceptance rate 62/176=35%).

Bryan Gibson[(s)], Timothy Rogers, Charles Kalish, and Xiaojin Zhu. What causes category-shifting in human semi-supervised learning? In *The 32nd Annual Conference of the Cognitive Science Society (CogSci)*, 2015.

Shike Mei[(s)] and Xiaojin Zhu. The security of latent Dirichlet allocation. In *The Eighteenth International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015. (oral, acceptance rate 27/442=6%).

Xiaojin Zhu. Machine teaching: an inverse problem to machine learning and an approach toward optimal education. In *The Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015. **AAAI / Computing Community Consortium "Blue Sky Ideas" Track Prize**.

Shike Mei(s) and Xiaojin Zhu. Using machine teaching to identify optimal training-set attacks on machine learners. In *The Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015. (acceptance rate 531/1991=27%).

Kaustubh Patil(o), Xiaojin Zhu, Lukasz Kopec(o), and Bradley Love. Optimal teaching for limited-capacity human learners. In *Advances in Neural Information Processing Systems (NIPS)*, 2014. (spotlight; overall acceptance rate 414/1678 = 25%).

Shike Mei(s), Han Li(o), Jing Fan(o), Xiaojin Zhu, and Charles R. Dyer. Inferring air pollution by sniffing social media. In *The 2014 IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, 2014. (acceptance rate 18%).

C. Gokhale(o), S. Das(o), A. Doan, J. Naughton, N. Rampali, J. Shavlik, and X. Zhu. Corleone: Hands-off crowdsourcing for entity matching. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, 2014.

Jun-Ming Xu(s), Hsun-Chih Huang(o), Amy Bellmore, and Xiaojin Zhu. School bullying in twitter and weibo: a comparative study. In *The Eighth International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2014. (acceptance rate 18/44=41%).

Shike Mei, Jun Zhu, and Xiaojin Zhu. Robust RegBayes: Selectively incorporating First-Order Logic domain knowledge into Bayesian models. In *The 31st International Conference on Machine Learning (ICML)*, 2014.

Xiaojin Zhu. Machine teaching for Bayesian learners in the exponential family. In *Advances in Neural Information Processing Systems*, 2013. (acceptance rate 360/1420=25%).

Kwang-Sung Jun(s), Xiaojin Zhu, Burr Settles, and Timothy Rogers. Learning from human-generated lists. In *The 30th International Conference on Machine Learning (ICML)*, 2013. (acceptance rate 283/1204=24%).

Xiaojin Zhu. Persistent homology: An introduction and a new text representation for natural language processing. In *The 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, 2013. (acceptance rate 413/1473=28%).

Junming Xu(s), Benjamin Burchfiel(s), Xiaojin Zhu, and Amy Bellmore. An examination of regret in bullying tweets. In *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, 2013. Short paper (acceptance rate 32%).

Jun-Ming Xu(s), Aniruddha Bhargava(o), Robert Nowak, and Xiaojin Zhu. Socioscope: Spatio-temporal signal recovery from social media (extended abstract). In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, 2013. (Invited abstract based on the ECML/PKDD'12 paper).

Jun-Ming Xu(s), Aniruddha Bhargava(o), Robert Nowak, and Xiaojin Zhu. Socioscope: Spatio-temporal signal recovery from social media. In *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)*, 2012. (acceptance rate 105/443=24%) **Best Paper in Knowledge Discovery at ECML-PKDD 2012**.

Jun-Ming Xu(s), Kwang-Sung Jun(s), Xiaojin Zhu, and Amy Bellmore. Learning from bullying traces in social media. In *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, 2012. (acceptance rate 61/197=31%).

Burr Settles and Xiaojin Zhu. Behavioral factors in interactive training of text classifiers. In *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, 2012. (short paper, acceptance rate 36/104=35%).

Faisal Khan(o), Xiaojin Zhu, and Bilge Mutlu. How do humans teach: On curriculum learning and teaching dimension. In *Advances in Neural Information Processing Systems (NIPS) 25*. 2011. (acceptance rate 305/1400=22%).

Shilin Ding(o), Grace Wahba, and Xiaojin Zhu. Learning higher-order graph structure with features by structure penalty. In *Advances in Neural Information Processing Systems (NIPS) 25*. 2011. (acceptance rate 305/1400=22%).

Nathan Rosenblum(o), Xiaojin Zhu, and Barton P. Miller. Who wrote this code? identifying the authors of program binaries. In *The European Symposium on Research in Computer Security (ESORICS)*, 2011. (acceptance rate 36/155=23%).

Xiaojin Zhu, Bryan Gibson(s), and Timothy Rogers. Co-training as a human collaboration policy. In *The Twenty-Fifth Conference on Artificial Intelligence (AAAI-11)*, 2011. (acceptance rate 242/975=25%).

Andrew Goldberg(s), Xiaojin Zhu, Alex Furger(s), and Jun-Ming Xu(s). OASIS: Online active semi-supervised learning. In *The Twenty-Fifth Conference on Artificial Intelligence (AAAI-11)*, 2011. (acceptance rate 242/975=25%, selected for additional poster highlight).

Nathan Rosenblum(o), Barton P. Miller, and Xiaojin Zhu. Recovering the toolchain provenance of binary code. In *International Symposium on Software Testing and Analysis (ISSTA)*, 2011. (acceptance rate 35/121=29%) **ACM SIGSOFT Distinguished Paper Award**.

Chen Yu, Jun-Ming Xu(s), and Xiaojin Zhu. Word learning through sensorimotor child-parent interaction: A feature selection approach. In *The 33rd Annual Conference of the Cognitive Science Society (CogSci 2011)*, 2011. (oral, acceptance rate 32%).

David Andrzejewski(s), Xiaojin Zhu, Mark Craven, and Ben Recht. A framework for incorporating general domain knowledge into Latent Dirichlet Allocation using First-Order Logic. In *The Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI-11)*, 2011. (acceptance rate 227/1325=17%).

Mariyam Mirza(o), Paul Barford, Xiaojin Zhu, Suman Banerjee, and Michael Blodgett(o). Fingerprinting 802.11 rate adaptation algorithms. In *The 30th IEEE International Conference on Computer Communications (INFOCOM)*, Shanghai, China, 2011. (acceptance rate 291/1823=16%).

Bryan Gibson(s), Xiaojin Zhu, Tim Rogers, Chuck Kalish, and Joseph Harrison(o). Humans learn using manifolds, reluctantly. In *Advances in Neural Information Processing Systems (NIPS) 24*. 2010. (**Plenary oral presentation: 20/1219=2%**).

Andrew Goldberg(s), Xiaojin Zhu, Benjamin Recht, Junming Sui(s), and Robert Nowak. Transduction with matrix completion: Three birds with one stone. In *Advances in Neural Information Processing Systems (NIPS) 24*. 2010. (acceptance rate 293/1219=24%).

Xiaojin Zhu, Bryan R. Gibson(s), Kwang-Sung Jun(s), Timothy T. Rogers, Joseph Harrison(o), and Chuck Kalish. Cognitive models of test-item effects in human category learning. In *The 27th International Conference on Machine Learning (ICML)*, 2010. (acceptance rate 152/594=25.6%).

Timothy Rogers, Charles Kalish, Bryan Gibson(s), Joseph Harrison(o), and Xiaojin Zhu. Semi-supervised learning is observed in a speeded but not an unspeeded 2D categorization task. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society (CogSci)*, 2010. (poster; acceptance rate 74% out of 810 submissions).

Xiaojin Zhu, Timothy T. Rogers, and Bryan Gibson$^{(s)}$. Human Rademacher complexity. In *Advances in Neural Information Processing Systems (NIPS) 23*. 2009. (Acceptance rate 263/1105=23.8%).

David Andrzejewski$^{(s)}$, Xiaojin Zhu, and Mark Craven. Incorporating domain knowledge into topic modeling via Dirichlet forest priors. In *The 26th International Conference on Machine Learning (ICML)*, 2009. (acceptance rate 160/595=26.9%).

Andrew Goldberg$^{(s)}$, Nathanael Fillmore$^{(s)}$, David Andrzejewski$^{(s)}$, Zhiting Xu$^{(s)}$, Bryan Gibson$^{(s)}$, and Xiaojin Zhu. May all your wishes come true: A study of wishes and how to recognize them. In *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT)*, 2009. (acceptance rate 75/260=28.8%).

Andrew Goldberg$^{(s)}$, Xiaojin Zhu, Aarti Singh, Zhiting Xu$^{(s)}$, and Robert Nowak. Multi-manifold semi-supervised learning. In *Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2009. (acceptance rate 84/210=40%).

Rui Castro$^{(o)}$, Charles Kalish, Robert Nowak, Ruichen Qian$^{(s)}$, Timothy Rogers, and Xiaojin Zhu. Human active learning. In *Advances in Neural Information Processing Systems (NIPS) 22*. 2008. (Acceptance rate 250/1022=24.5%).

Aarti Singh$^{(o)}$, Robert Nowak, and Xiaojin Zhu. Unlabeled data: Now it helps, now it doesn't. In *Advances in Neural Information Processing Systems (NIPS) 22*. 2008. (**Plenary oral presentation: 28/1022=3%**).

Mariyam Mirza$^{(o)}$, Kevin Springborn, Suman Banerjee, Paul Barford, Mike Blodgett$^{(o)}$, and Xiaojin Zhu. On the accuracy of TCP throughput prediction for opportunistic wireless networks. In *Proceedings of IEEE Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '09)*, 2009.

Andrew B. Goldberg$^{(s)}$, Ming Li$^{(s)}$, and Xiaojin Zhu. Online manifold regularization: A new learning setting and empirical study. In *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)*, 2008. (acceptance rate 98/521=18.8%).

Andrew B. Goldberg$^{(s)}$, Xiaojin Zhu, Charles R. Dyer, Mohamed Eldawy$^{(o)}$, and Lijie Heng$^{(s)}$. Easy as ABC? Facilitating pictorial communication via semantically enhanced layout. In *Twelfth Conference on Computational Natural Language Learning (CoNLL)*, 2008. (acceptance rate 20/85=23.5%).

Xiaojin Zhu, Michael Coen, Shelley Prudom, Ricki Colman, and Joseph Kemnitz. Online learning in monkeys. In *Twenty-Third AAAI Conference on Artificial Intelligence (AAAI-08)*, 2008. (short paper, overall acceptance rate (23+227)/958=26%).

Nathan Rosenblum$^{(o)}$, Xiaojin Zhu, Barton Miller, and Karen Hunt. Learning to analyze binary computer code. In *Twenty-Third AAAI Conference on Artificial Intelligence (AAAI-08)*, 2008. (full paper, acceptance rate 227/937=24%; **selected for additional poster highlight, 5%**).

Xiaojin Zhu, Andrew B. Goldberg$^{(s)}$, Michael Rabbat, and Robert Nowak. Learning bigrams from unigrams. In *The 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL)*, 2008. (acceptance rate 25%).

Pedro DeRose$^{(o)}$, Xiaoyong Chai$^{(o)}$, Byron Gao$^{(o)}$, Warren Shen$^{(o)}$, AnHai Doan, Philip Bohannon$^{(o)}$, and Xiaojin Zhu. Building community Wikipedias: A machine-human partnership approach. In *IEEE International Conference on Data Engineering (ICDE)*, 2008. (acceptance rate 12.1%).

David Andrzejewski$^{(s)}$, Anne Mulhern$^{(o)}$, Ben Liblit, and Xiaojin Zhu. Statistical debugging using latent topic models. In *Proceedings of the 18th European Conference on Machine Learning (ECML)*, 2007. (acceptance rate 11.6%).

Gregory Druck(o), Chris Pal(o), Xiaojin Zhu, and Andrew McCallum. Semi-supervised classification with hybrid generative/discriminative methods. In *The Thirteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2007. (acceptance rate 20%).

Jordan Boyd-Graber(o), David Blei, and Xiaojin Zhu. A topic model for word sense disambiguation. In *Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL)*, 2007. (acceptance rate 27%).

Xiaojin Zhu, Timothy Rogers, Ruichen Qian(s), and Chuck Kalish. Humans perform semi-supervised classification too. In *Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, 2007. (full paper, acceptance rate 27%; **selected for additional poster highlight, 5%**).

Xiaojin Zhu, Andrew Goldberg(s), Mohamed Eldawy(o), Charles Dyer, and Bradley Strock(s). A Text-to-Picture synthesis system for augmenting communication. In *Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, pages 1590–1595, 2007. (acceptance rate 27%).

Xiaojin Zhu and Andrew Goldberg(s). Kernel regression with order preferences. In *Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, 2007. (acceptance rate 27%).

Mariyam Mirza(o), Joel Sommers(o), Paul Barford, and Xiaojin Zhu. A machine learning approach to TCP throughput prediction. In *The International Conference on Measurement and Modeling of Computer Systems (ACM SIGMETRICS)*, 2007. (acceptance rate 17%).

Xiaojin Zhu, Andrew Goldberg(s), Jurgen Van Gael(s), and David Andrzejewski(s). Improving diversity in ranking using absorbing random walks. In *Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, 2007. (acceptance rate 24%).

Andrew Goldberg(s), Xiaojin Zhu, and Stephen Wright. Dissimilarity in graph-based semi-supervised classification. In *Eleventh International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007. (acceptance rate 50%).

Jurgen Van Gael(s) and Xiaojin Zhu. Correlation clustering for crosslingual link detection. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2007. (acceptance rate 34%).

Xiaojin Zhu and John Lafferty(a). Harmonic mixtures: combining mixture models and graph-based methods for inductive and scalable semi-supervised learning. In *The 22nd International Conference on Machine Learning (ICML)*. ACM Press, 2005. (acceptance rate 27%).

Xiaojin Zhu, Jaz Kandola, Zoubin Ghahramani, and John Lafferty(a). Nonparametric transforms of graph kernels for semi-supervised learning. In Lawrence K. Saul, Yair Weiss, and Léon Bottou, editors, *Advances in Neural Information Processing Systems (NIPS) 17*. MIT Press, Cambridge, MA, 2005. (acceptance rate 25%).

John Lafferty(a), Xiaojin Zhu, and Yan Liu. Kernel conditional random fields: Representation and clique selection. In *The 21st International Conference on Machine Learning (ICML)*, 2004. (acceptance rate 32%).

Xiaojin Zhu, Zoubin Ghahramani, and John Lafferty(a). Semi-supervised learning using Gaussian fields and harmonic functions. In *The 20th International Conference on Machine Learning (ICML)*, 2003. (acceptance rate 32%) **ICML Classic Paper Prize**.

Stefanie Shriver, Arthur Toth, Xiaojin Zhu, Alex Rudnicky, and Roni Rosenfeld(a). A unified design for human-machine voice interaction. In *Human Factors in Computing Systems (CHI)*. ACM Press, 2001. (acceptance rate 20%).

Xiaojin Zhu and Ronald Rosenfeld[(a)]. Improving trigram language modeling with the World Wide Web. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2001. (acceptance rate 51%).

Ronald Rosenfeld[(a)], Xiaojin Zhu, Stefanie Shriver, Arthur Toth, Kevin Lenzo, and Alan Black. Towards a universal speech interface. In *International Conference on Spoken Language Processing (ICSLP)*, 2000. (acceptance rate 78%).

Xiaojin Zhu, Jie Yang, and Alex Waibel. Segmenting hands of arbitrary color. In *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000. (acceptance rate 52%).

Jie Yang, Xiaojin Zhu, Ralph Gross, John Kominek, Yue Pan, and Alex Waibel. Multimodal people ID for a multimedia meeting browser. In *The Seventh ACM International Multimedia Conference*, 1999. (acceptance rate 21%).

Xiaojin Zhu, Stanley F. Chen, and Ronald Rosenfeld[(a)]. Linguistic features for whole sentence maximum entropy language models. In *Proceedings of the 5th European Conference on Speech Communication and Technology (Eurospeech)*, 1999. (acceptance rate 66%).

Refereed Workshop Papers

Owen Levin[(s)], Zihang Meng, Vikas Singh, and Xiaojin Zhu. Fooling computer vision into inferring the wrong body mass index. In *KDD Workshop on Adversarial Learning Methods for Machine Learning and Data Mining (AdvML)*, 2019.

Xuezhou Zhang[(s)], Hrag Gorune Ohannessian[(s)], Ayon Sen[(s)], Scott Alfeld[(s)], and Xiaojin Zhu. Optimal teaching for online perceptrons. In *NIPS 2016 workshop on Constructive Machine Learning*, 2016.

Christopher Meek, Patrice Y. Simard, and Xiaojin Zhu. Analysis of a design pattern for teaching with features and labels. *CoRR*, abs/1611.05950, 2016.

Scott Alfeld[(s)], Xiaojin Zhu, and Paul Barford. Machine teaching as search. In *The International Symposium on Combinatorial Search (SoCS)*, 2016.

Yung-Hsiang Lu, Milind Kulkarni, and Xiaojin Zhu. Programming language support for analyzing non-persistent data. In *2016 IEEE International Symposium on Technologies for Homeland Security*, 2016.

Gabriel Cadamuro[(o)], Ran Gilad-Bachrach, and Xiaojin Zhu. Debugging machine learning (extended abstract). In *CHI 2016 workshop on Human Centred Machine Learning*, 2016.

Nick Bridle[(s)] and Xiaojin Zhu. p-voltages: Laplacian regularization for semi-supervised learning on high-dimensional data. In *Eleventh Workshop on Mining and Learning with Graphs (MLG2013)*, 2013.

Jun-Ming Xu[(s)], Xiaojin Zhu, and Amy Bellmore. Fast learning for sentiment analysis on bullying. In *ACM KDD Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM)*, 2012.

Xiaojin Zhu, Jun-Ming Xu[(s)], Christine M. Marsh, Megan K. Hines, and F. Joshua Dein. Machine learning for zoonotic emerging disease detection. In *ICML 2011 Workshop on Machine Learning for Global Challenges*, 2011.

Bryan R Gibson[(s)], Kwang-Sung Jun[(s)], and Xiaojin Zhu. With a little help from the computer: Hybrid human-machine systems on bandit problems. In *NIPS 2010 Workshop on Computational Social Science and the Wisdom of Crowds*, 2010.

Faisal Khan[(o)], Bilge Mutlu, and Xiaojin Zhu. Modeling social cues: Effective features for predicting listener nods. In *NIPS 2010 Workshop on Human Communication Dynamics*, 2010.

Nathan Rosenblum(o), Barton Miller, and Xiaojin Zhu. Extracting compiler provenance from program binaries. In *Proceedings of the 9th ACM SIGPLAN-SIGSOFT workshop on Program Analysis for Software Tools and Engineering (PASTE)*, 2010. (acceptance rate 41%).

David Andrzejewski(s), David G. Stork, Xiaojin Zhu, and Ron Spronk. Inferring compositional style in the neo-plastic paintings of Piet Mondrian by machine learning. In *Electronic Imaging: Computer Image Analysis in the Study of Art (SPIE 2010)*, 2010.

Andrew B. Goldberg(s), Jake Rosin, Xiaojin Zhu, and Charles R. Dyer. Toward text-to-picture synthesis. In *NIPS 2009 Symposium on Assistive Machine Learning for People with Disabilities*, 2009.

Xiaojin Zhu, Zhiting Xu(s), and Tushar Khot(s). How creative is your writing? a linguistic creativity measure from computer science and cognitive psychology perspectives. In *NAACL 2009 Workshop on Computational Approaches to Linguistic Creativity*, 2009. (acceptance rate 8/19=42%).

Andrew B. Goldberg(s) and Xiaojin Zhu. Keepin' it real: Semi-supervised learning with realistic tuning. In *NAACL 2009 Workshop on Semi-supervised Learning for NLP*, 2009. (acceptance rate 10/17=59%).

David Andrzejewski(s) and Xiaojin Zhu. Latent Dirichlet allocation with topic-in-set knowledge. In *NAACL 2009 Workshop on Semi-supervised Learning for NLP*, 2009.

Xiaojin Zhu, Andrew B. Goldberg(s), and Tushar Khot(s). Some new directions in graph-based semi-supervised learning (invited paper). In *IEEE International Conference on Multimedia and Expo (ICME), Special Session on Semi-Supervised Learning for Multimedia Analysis*, 2009.

Nathan Rosenblum(o), Xiaojin Zhu, Barton Miller, and Karen Hunt. Machine learning-assisted binary code analysis. In *NIPS 2007 workshop on Machine Learning in Adversarial Environments for Computer Security*, 2007.

SaiSuresh Krishnakumaran(s) and Xiaojin Zhu. Hunting elusive metaphors using lexical resources. In *NAACL 2007 Workshop on Computational Approaches to Figurative Language*, 2007.

Andrew Goldberg(s), Dave Andrzejewski(s), Jurgen Van Gael(s), Burr Settles(o), Xiaojin Zhu, and Mark Craven. Ranking biomedical passages for relevance and diversity: University of Wisconsin, Madison at TREC genomics 2006. In *Proceedings of the Fifteenth Text Retrieval Conference (TREC)*, 2006.

Andrew Goldberg(s) and Xiaojin Zhu. Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization. In *HLT-NAACL 2006 Workshop on Textgraphs: Graph-based Algorithms for Natural Language Processing*, New York, NY, 2006.

Maria-Florina Balcan, Avrim Blum, Patrick Pakyan Choi, John Lafferty(a), Brian Pantano, Mugizi Robert Rwebangira, and Xiaojin Zhu. Person identification in webcam images: An application of semi-supervised learning. In *ICML 2005 Workshop on Learning with Partially Classified Training Data*, 2005.

Xiaojin Zhu, John Lafferty(a), and Zoubin Ghahramani. Combining active learning and semi-supervised learning using Gaussian fields and harmonic functions. In *ICML 2003 workshop on The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining*, 2003.

Unrefereed Technical Reports

Xiaojin Zhu. An optimal control view of adversarial machine learning. *arXiv*, 1811.04422, 2018.

Evan Hernandez, Ara Vartanian, and Xiaojin Zhu. Program synthesis with visual specification. *ArXiv e-prints*, 2018. https://arxiv.org/abs/1806.00938.

Xiaojin Zhu, Adish Singla, Sandra Zilles, and Anna N. Rafferty. An Overview of Machine Teaching. *ArXiv e-prints*, January 2018. https://arxiv.org/abs/1801.05927.

Scott Alfeld(s), Paul Barford, and Xiaojin Zhu. Optimal defense actions against test-set attacks. In *ICML Workshop on Reliable Machine Learning in the Wild*, 2016.

Shalini Ghosh, Patrick Lincoln, Ashish Tiwari, and Xiaojin Zhu. Trusted machine learning for probabilistic models. In *ICML Workshop on Reliable Machine Learning in the Wild*, 2016.

Gabriel Cadamuro(o), Ran Gilad-Bachrach, and Xiaojin Zhu. Debugging machine learning models. In *ICML Workshop on Reliable Machine Learning in the Wild*, 2016.

Shike Mei(s) and Xiaojin Zhu. Some submodular data-poisoning attacks on machine learners. Technical Report Computer Science TR1822, University of Wisconsin-Madison, 2015.

Yimin Tan(s) and Xiaojin Zhu. Dragging: Density-ratio bagging. Technical Report Computer Science TR1795, University of Wisconsin-Madison, 2013.

Michael Maynord(s), Jitrapon Tiachunpun(s), Xiaojin Zhu, Charles R. Dyer, Kwang-Sung Jun(s), and Jake Rosin(s). An image-to-speech iPad app. Technical Report Computer Science TR1774, University of Wisconsin-Madison, 2012.

Jake Rosin(s), Andrew B. Goldberg(s), Xiaojin Zhu, and Charles Dyer. A Bayesian model for image sense ambiguity in pictorial communication systems. Technical Report Computer Science TR1692, University of Wisconsin-Madison, 2011.

Nathanael Fillmore(s), Andrew B. Goldberg(s), and Xiaojin Zhu. Document recovery from bag-of-word indices. Technical Report Computer Science TR1645, University of Wisconsin-Madison, 2008.

Xiaojin Zhu and Andrew Goldberg(s). Semi-supervised regression with order preferences. Technical Report 1578, Department of Computer Sciences, University of Wisconsin-Madison, 2006.

Xiaojin Zhu, David Blei, and John Lafferty(a). TagLDA: Bringing document structure knowledge into topic models. Technical Report 1553, Department of Computer Sciences, University of Wisconsin-Madison, 2006.

Xiaojin Zhu. Semi-supervised learning literature survey. Technical Report 1530, Department of Computer Sciences, University of Wisconsin, Madison, 2005.

Xiaojin Zhu, Zoubin Ghahramani, and John Lafferty(a). Time-sensitive Dirichlet process mixture models. Technical Report CMU-CALD-05-104, Carnegie Mellon University, 2005.

Xiaojin Zhu. *Semi-Supervised Learning with Graphs*. PhD thesis, Carnegie Mellon University, 2005. CMU-LTI-05-192.

Xiaojin Zhu, John Lafferty(a), and Zoubin Ghahramani. Semi-supervised learning: From Gaussian fields to Gaussian processes. Technical Report CMU-CS-03-175, Carnegie Mellon University, 2003.

Xiaojin Zhu and Zoubin Ghahramani. Learning from labeled and unlabeled data with label propagation. Technical Report CMU-CALD-02-107, Carnegie Mellon University, 2002.

Xiaojin Zhu and Zoubin Ghahramani. Towards semi-supervised classification with Markov random fields. Technical Report CMU-CALD-02-106, Carnegie Mellon University, 2002.

Xiaojin Zhu and Ronald Rosenfeld(a). Improving trigram language modeling with the World Wide Web. Technical Report CMU-CALD-00-171, Carnegie Mellon University, 2000.

Research Grants

Federal Grants

2024-2026. NSF. SLES: Foundations of Safety-Aware Learning in the Wild. CoPI.

2022-2025. NSF. Digitally Inoculating Viewers Against Visual Misinformation With a Perceptual Training. CoPI.

2021-2026. ARO MURI. Cohesive and Robust Human-Bot Cybersecurity Teams. CoPI.

2021-2022. University of Wisconsin-Milwaukee. Risk and Resilience in Urban Black American Acute Trauma Survivors. CoPI.

2020-2022. NSF. EAGER: Rule Induction Games to Explore Differences between Human and Machine Intelligence. CoPI.

2020-2025. University of Wisconsin-Milwaukee. Acute predictors of long-term post-trauma outcomes in youth victims of violence. CoPI.

2019-2024. National Institutes of Health. Computational Biases of Learning and Decision-Making in PTSD. CoPI.

2019-2024. United States Department of Agriculture. FACT: AN Innovative Cyber-Framework Integrating Public/Private Data for Evidence-Based Recommendations. CoPI.

2019-2024. National Institutes of Health. Contextualized daily prediction of lapse risk in opiod use disorder by digital phenotyping. CoPI.

2019-2023. National Institutes of Health. Neurobehavioral mechanisms of parent-child extinction learning in adolescent PTSD. CoPI.

2018-2022. NSF. FMitF: Collaborative Research: Formal Methods for Machine Learning System Design. CoPI.

2018-2021. AFOSR. PI: Rob Nowak, Co-PIs: Xiaojin Zhu et al. "Machines, Algorithms and Data Lab (MADlab): A University Center of Excellence in Efficient and Robust Machine Learning."

2017-2020. National Science Foundation CCF-1704117. PI: Aws Albarghouthi, Co-PIs: Loris D'Antoni, Shuchi Chawla, Xiaojin Zhu. "SHF: Medium: Formal Methods for Program Fairness"

2016-2019. National Science Foundation IIS-1623605. Martina Rau (PI), Robert Nowak, and Xiaojin Zhu (CoPI). "EXP: Modeling Perceptual Fluency with Visual Representations in an Intelligent Tutoring System for Undergraduate Chemistry"

2016-2019. National Science Foundation CMMI-1561512. Shiyu Zhou PI and X. Zhu coPI. $299K, "Enabling Cloud-Based Quality-Data Management Systems"

2015-2020. NSF DGE-1545481. CoPI (PI Timothy Rogers). "NRT-DESE LUCID: A project-focused cross-disciplinary graduate training program for data-enabled research in human and machine learning and teaching"

2015-2020. NIH/National Institute of Alcohol Abuse and Alcoholism (R01 AA024391). CoI (PI John Curtin). "Dynamic, real-time prediction of alcohol use lapse using mHealth technologies"

2013–2018. NIH/NIGMS. R01 GM111002-02S1 (PI Carnes). "Exploring the Science of Scientific Review–Admin Supplement"

2014-2018. NIH Big Data to Knowledge 1U54AI117924-01. CoPI (PI M. Craven). "The Center for Predictive Computational Phenotyping"

2014-2017. National Science Foundation, CCF-1423237. Zhu (PI), Reps and Liblit (coPIs). "SHF: Small: Transforming Natural Language to Programming Languages"

2014-2017. National Science Foundation, CNS-1343363. S Banerjee PI, X. Zhang and X. Zhu coPIs. $963K, "EARS: A TV Whitespace Communication System for Connected Vehicles"

2012-2016. National Science Foundation, IIS-1216758. X. Zhu PI. $500K, "III: Small: Advancing the Scientific Understanding of Bullying Through the Lens of Social Media"

2011-2013. National Science Foundation, IIS-1148012. C. Dyer PI, X. Zhu CoPI. $152K, "III: EAGER: Discovering Spontaneous Social Events"

2010-2016. National Science Foundation, IIS-0953219. X. Zhu PI. $466K, "CAREER: Using Machine Learning to Understand and Enhance Human Learning Capacity"

2009-2012. National Science Foundation, IIS-0916038. X. Zhu PI. $414K, "RI:Small:Semi-Supervised Learning for Non-Experts"
Supplemental funding: Research Experiences for Undergraduates (REU). $12K

2009-2011. Air Force Office of Scientific Research, FA9550-09-1-0313. X. Zhu PI, T. Rogers Co-PI. $437K, "A Cognitive Study of Learning with Labeled and Unlabeled Data"

2007-2010. National Science Foundation, IIS-0711887. X. Zhu PI, C. Dyer Co-PI. $400K, "RI: Text-to-Picture Synthesis"
Supplemental funding: Research Experiences for Undergraduates (REU). $12K

Industrial Grants

2020. American Family Insurance.

2018. Intuit.

2016. American Family Insurance Research Award.

2013. Google Faculty Research Award. "Corroborating Knowledge Graph with Timely Facts from Social Media"

Intramural Grants

2024-2026. University of Wisconsin-Madison Research Forward Initiative. "An AI Terrarium for Understanding Intermediated and Personal Communication."

2020-2022. University of Wisconsin-Madison UW2020: WARF Discovery Initiative competition. "Human and Machine Learning: The Search for Anomalies."

2020-2022. University of Wisconsin-Madison Vilas Associate. "Machine Teaching for Optimal Personalized Education."

2020-2021. University of Wisconsin-Madison Graduate School Research Award. X. Zhu PI. $34K, "Optimal Control for Better Machine Learning." Awarded, not utilized.

2015-2016. University of Wisconsin-Madison Graduate School Research Award. M. Alibali PI, P. Matthews, T. Rogers, X. Zhu CoPI. $117K, "Optimizing Mathematics Learning."

2014-2015. University of Wisconsin-Madison Graduate School Research Award. X. Zhu PI. $34K, "Machine Teaching."

2013-2014. University of Wisconsin-Madison Graduate School Research Award. X. Zhu PI, Chitturi, Noyce CoPI. $47K, "Surrogate Traffic Safety Measures from Social Media." Awarded, not utilized.

2012-2013. University of Wisconsin-Madison Graduate School Research Award. X. Zhu PI, A. Bellmore CoPI. $36K, "Combating Bullying with Machine Learning". Awarded, not utilized.

2011-2013. University of Wisconsin-Madison, Global Health Institute Award. L. Gilbert PI, J. Dein & B. Shaw & X. Zhu Co-PIs, $40K, "Evaluation of Alternative Strategies for Emerging Disease Detection"

2010-2011. University of Wisconsin Graduate School Research Award. X. Zhu PI, J. Murray-Branch Co-PI. $35K, "Situation-Aware Communication Board for People with Disabilities"

2008-2009. University of Wisconsin Graduate School Research Award. X. Zhu PI, T. Rogers Co-PI. $35K, "Semi-Supervised Learning in Humans and Machines"

2008. University of Wisconsin Cognitive Science Cluster Research Fellowship. R. Qian, X. Zhu (Project Sponsor). $3K, "Learning: Between Humans and Machines"

2007-2008. University of Wisconsin Graduate School Research Award. X. Zhu PI, E. Churchwell, Co-PI. $31K, "Application of artificial intelligence and human computing methods to panoramic astrophysical surveys"

2006-2007. University of Wisconsin Graduate School Research Award. M. Craven PI, X. Zhu Co-PI. $30K, "Extracting background knowledge from the scientific literature to improve the accuracy of gene regulatory network inference"

Professional Service

Conference Organizer

- Position Paper Track co-chair, International Conference on Machine Learning (ICML) 2026 (with Dale Schuurmans). Seoul, Korea.
- Position Paper Track co-chair, International Conference on Machine Learning (ICML) 2025 (with Kiri Wagstaff). Vancouver, Canada.
- Tutorial Chair, International Conference on Machine Learning (ICML) 2020 (with Alessandra Tosi). Vienna, Austria.
- Chair, The 40th Annual Meeting of the Cognitive Science Society (CogSci 2018), Madison, Wisconsin, USA.
- Chair, AISTATS 2017, Fort Lauderdale, Florida (with Aarti Singh).
- Workshop Chair, International Conference on Machine Learning (ICML) 2011 (with Katherine Heller). Bellevue, Washington, USA.

Workshop Organizer

- NSF Workshop on Safety and Trust in Artificial Intelligence (AI) Enabled Systems. With Taylor Johnson (Vanderbilt), Kate Saenko (Boston University), Pavithra Prabhakar (NSF). September 2022. Virtual.
- NIPS 2017 Workshop on Teaching Machines, Robots, and Humans (with Maya Cakmak, Anna Rafferty, Adish Singla, Sandra Zilles). Dec 2017, Long Beach, CA.
- Dagstuhl Seminar 17351 on Machine Learning and Formal Methods. August 2017, Germany.
- ICML 2014 Workshop on Topological Methods for Machine Learning. Beijing, China.
- ICML 2011 Workshop on Combining Strategies for Reducing the Cost of Learning. Seattle, WA.
- AAAI 2009 Fall Symposium on manifold learning.
- NIPS 2008 Workshop on Machine Learning Meets Human Learning. Whistler, Canada.
- HAMLET (Human, Animal, and Machine Learning: Experiment and Theory) lecture series, Departments of Computer Sciences and Psychology, University of Wisconsin-Madison, 2008-present.

Journal Editing

- Associate Editor, ACM / IMS Journal of Data Science, 2022-
- Action Editor, Machine Learning Journal, 2011-2014
- Editorial Board, AI Matters: A Newsletter of ACM SIGAI, 2014-2015
- Editorial Board, Machine Learning Journal, 2008-2011

Area Chair, Senior Program Committee

- Area Chair (Chair / AC / SPC / PC system), AAAI-27. 2027.
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2026.
- Senior Area Chair (Chair / SAC / AC / R system), International Conference on Machine Learning (ICML). 2026.
- Area Chair (Chair / AC / SPC / PC system), AAAI-26. 2026.
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2025.
- Senior Area Chair (Chair / SAC / AC / R system), International Conference on Machine Learning (ICML). 2025.
- Senior Area Chair (Chair / SAC / AC / R system), International Conference on Machine Learning (ICML). 2024.
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2023.
- Session Chair for machine learning, NSF Formal Methods in the Field (FMitF) PI Meeting, Nov. 14-15, 2022.
- Area Chair (Chair / AC / SPC / PC system), AAAI-23. 2023.
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2022.
- Senior Meta-Reviewer (Program Chair / SMR / MR / R system), International Conference on Machine Learning (ICML). 2022.
- Meta-Reviewer, the 25th International Conference on Artificial Intelligence and Statistics (AISTATS 2022).
- Senior Area Chair (Chair / SAC / AC / R system), International Conference on Machine Learning (ICML). 2021.
- Area Chair, the 24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021).
- Area Chair (Chair / AC / SPC / PC system), AAAI-21, online. 2021.
- Senior Area Chair (Chair / SAC / AC / SPC / PC system), The International Joint Conference on Artificial Intelligence (IJCAI-21). Montreal, Canada. 2021
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2020. Vancouver, Canada.
- Area Chair, the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020).
- Area Chair (Chair / AC / SPC / PC system), AAAI-20, New York, USA. 2020.
- Senior area chair (Chair / SAC / AC / PC system), Neural Information Processing Systems (NeurIPS) 2019. Vancouver, Canada.
- Area Chair (Chair / AC / SPC / PC system), The 28th International Joint Conference on Artificial Intelligence (IJCAI-19). Macao, China. 2019.
- Area Chair, International Conference on Machine Learning (ICML). Long Beach, CA. 2019.
- Senior Program Committee, the Twenty-Second International Conference on Artificial Intelligence and Statistics (AISTATS 2019).
- Area Chair (Chair / AC / SPC / PC system), AAAI-19, Honolulu, USA. 2019.
- Program Committee, CogSci 2019.

- Area Chair, International Conference on Machine Learning (ICML). Stockholm. 2018.
- Senior Program Committee, AAAI-18, New Orleans, Louisiana, USA, 2018.
- Senior Area Chair, Neural Information Processing Systems (NIPS). California. 2017.
- Area Chair, International Conference on Machine Learning (ICML). New York. 2016.
- Senior Program Committee, AAAI-16, Phoenix, Arizona, USA, 2016.
- Area Chair, Neural Information Processing Systems (NIPS). Montreal, Canada. 2015.
- Senior Program Committee, The 24th International Joint Conference on Artificial Intelligence (IJCAI). Buenos Aires, Argentina. 2015.
- Area Chair, International Conference on Machine Learning (ICML). Lille, France. 2015.
- Area Chair, The North American Association for Computational Linguistics (NAACL). Denver, Colorado. 2015.
- Area Chair, Conference on Empirical Methods in Natural Language Processing (EMNLP) Qatar. 2014.
- Area Chair, International Conference on Natural Language Processing and Chinese Computing (NLPCC). Shenzhen, China. 2014.
- Area Chair, International Conference on Machine Learning (ICML). Beijing, China. 2014.
- Senior Program Committee, The 14th SIAM International Conference on Data Mining (SDM). Philadelphia, PA, USA. 2014.
- Senior Program Committee, the Sixteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2013).
- Area Chair, International Conference on Machine Learning (ICML). Atlanta, Georgia. 2013.
- Senior Program Committee, The 13th SIAM International Conference on Data Mining (SDM). Austin, TX, USA. 2013.
- Area Chair, Neural Information Processing Systems (NIPS). Lake Tahoe, Nevada, United States. 2012.
- Area Chair, International Conference on Machine Learning (ICML). Edinburgh, Scotland. 2012.
- Area Chair, The North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT). Montreal, Canada. 2012.
- Senior Program Committee, The 4th Asian Conference on Machine Learning (ACML). Singapore. 2012.
- Area Chair, Neural Information Processing Systems (NIPS). Granada, Spain. 2011.
- Area Chair and the Best Paper Awards Committee, International Conference on Machine Learning (ICML). Bellevue, Washington, USA. 2011.
- Senior Program Committee, The 3rd Asian Conference on Machine Learning (ACML). Taoyuan, Taiwan. 2011.
- Area Chair, Neural Information Processing Systems (NIPS). Vancouver, BC, Canada. 2010.
- Area Chair, International Conference on Machine Learning (ICML). Haifa, Israel. 2010.
- Senior Program Committee, International Conference on Machine Learning (ICML). Corvallis, Oregon, USA. 2007.

Conference Program Committee

- The 41st Annual Meeting of the Cognitive Science Society (CogSci 2019), Montreal, Canada, USA.

- The 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2015.
- Conference on Empirical Methods in Natural Language Processing (EMNLP), 2011.
- Workshop on Robust Unsupervised and Semisupervised Methods in Natural Language Processing, 2011.
- AAAI Conference on Artificial Intelligence (AAAI), 2010.
- The 23rd International Conference on Computational Linguistics (COLING), 2010.
- The 9th IEEE International Conference on Development and Learning (ICDL), 2010.
- TextGraphs-5: Graph-based Methods for Natural Language Processing, 2010.
- International Conference on Machine Learning (ICML), 2009.
- International Conference on Artificial Intelligence and Statistics (AISTATS), 2009.
- Annual Meeting of the Association for Computational Linguistics (ACL-IJCNLP), 2009.
- North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT), 2009.
- Uncertainty in Artificial Intelligence (UAI), 2009.
- NIPS 2009 Workshop on Applications for Topic Models: Text and Beyond, 2009.
- The first International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion Measurement, 2009.
- IJCAI 2009 Workshop on Intelligence and Interaction, 2009.
- NAACL 2009 Workshop on Semi-supervised Learning for Natural Language Processing, 2009.
- International Conference on Machine Learning (ICML), 2008.
- AAAI Conference on Artificial Intelligence (AAAI), 2008.
- Annual Meeting of the Association for Computational Linguistics (ACL), 2008.
- Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2008.
- European Conference on Computer Vision (ECCV), 2008.
- The Fifth Midwest Computational Linguistics Colloquium (MCLC-5), 2008.
- The 6th International Workshop on Mining and Learning with Graphs (MLG), 2008.
- The Pacific Rim International Conference on Artificial Intelligence (PRICAI), 2008.
- AAAI Conference on Artificial Intelligence (AAAI), 2007.
- International Conference on Artificial Intelligence and Statistics (AISTATS), 2007.
- Conference on Empirical Methods in Natural Language Processing (EMNLP-CoNLL), 2007.
- European Conference on Machine Learning and European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), 2007.
- Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 2007.
- International Workshop on Mining and Learning on Graphs (MLG), 2007.
- HLT-NAACL workshop on Textgraphs: Graph-based Algorithms for Natural Language Processing, 2007.
- International Conference on Machine Learning (ICML), 2006.
- AAAI Conference on Artificial Intelligence (AAAI), 2006.
- Uncertainty in Artificial Intelligence (UAI), 2006.
- International Conference on Knowledge Discovery and Data Mining (KDD), 2006.

- European Conference on Machine Learning and European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD), 2006.
- ICML workshop on Learning with Nonparametric Bayesian Method, 2006.
- ECML/PKDD workshop on Mining and Learning with Graphs, 2006.
- HLT-NAACL workshop on Textgraphs: Graph-based Algorithms for Natural Language Processing, 2006.
- Uncertainty in Artificial Intelligence (UAI), 2005.
- International Conference on Artificial Intelligence and Statistics (AISTATS), 2005.
- ICML Workshop on Learning with Partially Classified Training Data, 2005.
- International Conference on Machine Learning (ICML), 2004.

Conference Reviewer

- ACL 2012; AISTATS 2010, 2012; CVPR 2009; NIPS 2007–2009

Journal Reviewer

- Journal of Machine Learning Research (JMLR), Machine Learning Journal, Journal of the American Statistical Association (JASA), IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), IEEE Transactions on Information Theory, Artificial Intelligence, Journal of Artificial Intelligence Research (JAIR), IEEE Transactions on Knowledge and Data Engineering (TKDE), IEEE Intelligent Systems, IEEE Transactions on Neural Networks, ACM Transactions on Knowledge Discovery from Data, Pattern Recognition Letters, Optimization Method and Software, Journal of Computational and Graphical Statistics, Neurocomputing, Neural Computation, Journal of Software (China), Springer

Grant Panelist and Reviewer

- NSF (CISE and SBE)
- AFOSR
- Israel Science Foundation
- The Institute for Clinical and Translational Research (ICTR), University of Wisconsin-Madison, 2010

Member of Visiting Chair Professor Group in Computer Science of Zhiyuan College, Shanghai Jiao Tong University, 2012.

Member of the United States Department of Defense Advanced Research Projects Agency (DARPA) Information Science and Technology (ISAT) advisory group, 2019.

Tutorials

IJCAI. *Tutorial on Adversarial Sequential Decision Making*. Vienna, Austria. 2022.

China Computer Federation Advanced Disciplines Lectures 46. *Statistical Learning Theory*. Beijing China, 2014.

Machine Learning Summer School (MLSS). *Graphical Models and Kernel Methods*. Beijing China, 2014.

China Computer Federation Advanced Disciplines Lectures 46. *Statistical Machine Learning for NLP*. Chongqing China, 2013.

China Computer Federation Advanced Disciplines Lectures 39. *Understanding Social Media with Machine Learning*. Beijing China, 2013.

Graphical Models. KDD 2012, Beijing, China. 2012.

All of Graphical Models. The tenth International Conference on Machine Learning and Applications (ICMLA'11), Honolulu, Hawaii. 2011.

Semi-Supervised Learning. Summer School on Theory and Practice of Computational Learning, University of Chicago, 2009.

Semi-Supervised Learning for Natural Language Processing (delivered by J. Blitzer). The 46th Association for Computational Linguistics meeting (ACL), Columbus, OH, 2008.

Semi-Supervised Learning. International Conference on Machine Learning (ICML), Corvallis, OR, 2007.

Semi-Supervised Classification: learning from labeled and unlabeled data. AERFAI Summer School on Action and Object Classification Techniques in Digital Images, University of Granada, Spain, 2006.

Talks

Systems, Information, Learning and Optimization (SILO) Seminar. *A Tutorial on Machine Teaching*. University of Wisconsin-Madison. July 9, 2025.

Simons Institute Workshop on Theoretical Aspects of Trustworthy AI. *Machine Teaching: Supervised Learning and Beyond*. UC Berkeley. April 30 2025.

Keynote at Workshop on Machine Teaching for Humans: Rethinking Example-Based Explanations. *Machine Teaching and "Nature vs. Nurture"*. University of Madeira. January 13, 2023.

Augmented Intelligence Workshop. *Creative Bandit*. Indiana University. September 8, 2022.

MADLab 2022 Summer Workshop. *Tragedy of the Data Scientists*. University of Chicago. June 14, 2022.

Systems, Information, Learning and Optimization (SILO) Seminar. *Game Redesign in No-regret Game Playing*. University of Wisconsin-Madison. Dec 1 2021.

Visiting Professor Series. *Train Text Classifiers Better Than Active Learning*. American Family Insurance. May 10 2021.

Simons Institute Workshop on Synthesis of Models and Systems. *A Brief Introduction to Theoretical Foundations of Machine Learning and Machine Teaching*. UC Berkeley. March 15 2021.

Air Force Research Laboratory and UW-Madison's Efficient and Robust Machine Learning Center of Excellence Workshop on Adversarial Robustness of Machine Learning. *Keynote: Introduction to Adversarial Machine Learning*, May 2020.

Center for Language and Speech Processing Fall Seminar Series, Johns Hopkins University. *Adversarial Machine Learning: Beyond Manipulating Pixels and Words*, Oct 2019.

1st Workshop on Adversarial Learning Methods for Machine Learning and Data Mining at KDD 2019. *Adversarial Machine Learning in Sequential Decision Making*, August 2019.

The 30th International Conference on Algorithmic Learning Theory (ALT). *How Fast can a Learner Learn under an Optimal Teacher?*, March 2019.

Applied and Computational Math Seminar, UW-Madison. *Machine Teaching: Optimal Control of Machine Learning*, Jan. 2019.

Vanderbilt University EECS. *Machine Teaching: Optimal Control of Machine Learning*, Nov. 2018.

Google. *How to Poison Linear Regression*, Oct. 2018.

ADVERSARY-AWARE LEARNING TECHNIQUES AND TRENDS IN CYBERSECURITY (ALEC), AAAI Fall Symposium. *An Optimal Control View of Adversarial Machine Learning*. October 18-19, 2018, Arlington, VA.

The University of Iowa Computer Science Colloquium. *Machine Teaching: Optimal Control of Machine Learning*, Oct. 2018.

Systems Information Learning Optimization (SILO), University of Wisconsin-Madison. *How to Poison Linear Regression*, Sept 2018.

2nd AOR/IARPA Workshop on Adversarial Machine Learning. *Toward adversarial learning as control*. University of Maryland, May 2018.

Department of Computer Science, Duke University. *Machine Teaching and its Applications*. 2018.

Tsinghua Laboratory of Brain and Intelligence Workshop on Brain and Artificial Intelligence. *Machine Teaching as a Probe for Learning Mechanism in Humans*. Beijing, China. Dec. 27, 2017

Tsinghua University. *Debugging the machine learning pipeline: machine teaching for and against adversarial attacks*. Beijing, China. Dec. 28, 2017.

The Interpretable Machine Learning Symposium at NIPS 2017. Long Beach, CA. *Debugging the Machine Learning Pipeline*

The Workshop on Teaching Machines, Robots, and Humans at NIPS 2017. Long Beach, CA. *Introduction to Machine Teaching*

Dagstuhl Seminar on Machine Learning and Formal Methods. August 2017, Germany. *A Challenge in Machine Teaching*.

Simons Institute Workshop on Interactive Learning 2017, Berkely, CA. *Machine Teaching in Interactive Learning*.

ICML 2016 Workshop on Reliable Machine Learning in the Wild, New York. *Machine Teaching and Security*.

The Center for Information and Systems Engineering, Boston University. 2016.

Department of Statistics, University of Indiana. 2016.

Machine Learning Lunch, Department of Computer Science and Engineering, University of Washington. *Machine Teaching*. 2016.

College of Computer Science, Northeastern University. 2016.

NIPS 2015 Workshop on Machine Learning from and for Adaptive User Technologies, Montreal, Canada. *Machine Teaching for Personalized Education, Security, Interactive Machine Learning*. 2015.

ICML 2015 Workshop on Machine Learning for Education, Lille, France. *Machine Teaching*. 2015.

Microsoft Research. *Machine Teaching*. 2015.

UC Berkeley Workshop on Active Learning. *Active Learning on Graphs*, 2015.

Systems Information Learning Optimization (SILO), University of Wisconsin-Madison. *An Approach to Bridge Topology and Machine Learning*, 2014.

Signatures workshop. *Topological Kernels*, 2014.

Simons Institute Workshop on Spectral Algorithms: From Theory to Practice. *Some Applications in Human Behavior Modeling*, 2014

Keynote at ICML 2014 Workshop on Learning, Security and Privacy, Beijing China. *Optimal Training Set Attacks on Machine Learning*. 2014.

Institute of Computer Science & Technology, Peking University. *Corpus Attacks on Natural Language Processing and Machine Learning*. 2014

Department of Computer Science, Tsinghua University. *Maximally Influencing Learning by Machine Teaching*. 2014

Department of Computer and Information Science, The University of Oregon. *Machine Teaching: Frenemy of Machine Learning*. 2014.

School of Electrical Engineering & Computer Science Colloquium, Oregon State University. *Machine Teaching: Frenemy of Machine Learning*. 2014.

STATISTICS COLLOQUIUM, Department of Statistics, The University of Chicago. *Machine Teaching: Frenemy of Machine Learning*. 2014.

Distinguished Speaker Series, Department of Computer Science, University of Virginia. *How to Make Machines Learn: Passive, Active, and Teaching*. 2014.

Duke University Machine Learning Seminar Series. *Machine Teaching: Frenemy of Machine Learning*. 2014.

Keynote at The Second Conference on Natural Language Processing & Chinese Computing, Chongqing, China. *Some Mathematical Models to Turn Social Media into Knowledge*. 2013.

Chongqing University, China. *How can a Machine Learn: Passive, Active, and Teaching*. 2013.

Systems Information Learning Optimization (SILO) Seminar, University of Wisconsin–Madison. *Optimal Teaching: The Inverse Problem of Machine Learning*, 2013

Microsoft Research Faculty Summit. *Three Assertions about Interactive Machine Learning*. Panelist, Interaction for Machine Learning panel, 2013.

Workshop on Integrating Approaches to Computational Cognition, National Science Foundation. *Capacity, Learning, Teaching*. 2013.

Department of Computer Science, Tsinghua University. *Persistent Homology Tutorial*. 2013

Department of Computer Science, Tsinghua University. *Social Media and Bullying Research*. 2013

Department of Computer Science, Tsinghua University. *A Brief Review of Semi-Supervised Learning*. 2013

China Telecom Group Shanghai Co. *Using Twitter to study bullying*. 2013

Department of Computer Science and Engineering, Shanghai Jiao Tong University. *A computational teaching theory for Bayesian learners*. 2012

Department of Computer Science and Engineering, Shanghai Jiao Tong University. *Persistent homology and an application in natural language processing*. 2012

Frontiers of Information Science and Technology Workshop, *Mining Social Media for Spatiotemporal Signal with Inhomogeneous Poisson Point Processes*, Shanghai, 2012.

NIPS Workshop on Algebraic Topology and Machine Learning, *Persistent Homology for Natural Language Processing*, 2012.

NIPS Workshop on Personalizing education with machine learning, *A Computational Teaching Theory for Bayesian Learners*, 2012.

Computer Science Department, Tsinghua University. *The nature of non-iid training data produced by humans for machine learning*, 2012.

Invited talk at Optimal Teaching Workshop, UCSD Temporal Dynamics of Learning Center. *Can Machine Learning Rationalize Simple Human Teaching Behaviors?*, 2012.

Machine Learning Special Seminar, Carnegie Mellon University. *Can Machine Learning Rationalize Simple Human Teaching Behaviors?*. 2012.

Online presentation to USGS Mobile Applications Development group. *Harvesting information from Twitter*. 2012.

Talk at the Eye Research Institute Seminar Series, University of Wisconsin-Madison. *Global Compound Eyes*. 2012.

Invited Talk, Purdue University, IN. *Machine learning theory by the people, for the people, of the people*. 2011.

Talk at Wisconsin Institute for Discovery, University of Wisconsin-Madison. *Man vs. Data: Domain Knowledge + Latent Dirichlet Allocation*. 2011.

Invited Talk, Duke University, NC. *Adding Human Guidance to Latent Dirichlet Allocation*. 2011.

Invited Talk, University of Massachusetts Amherst, MA. *Adding Knowledge to Latent Topic Models*. 2011.

Invited Talk, Microsoft Research, Redmond, WA. 2011. *Adding Knowledge to Latent Topic Models*.

Invited Talk, Center for Information and Systems Engineering, Boston University, Boston, MA, 2010. *Is Machine Learning the Wrong Name?*

Invited Talk, Cognitive Science, Indiana University, Bloomington, IN, 2010. *Some machine learning models for human learning*.

Talk at NIPS 2009 workshop on Analysis and Design of Algorithms for Interactive Machine Learning. Whistler, BC, Canada, 2009. *Human Machine Co-Learning*.

Talk at NIPS 2009 workshop on Bounded-rational analyses of human cognition: Bayesian models, approximate inference, and the brain. Whistler, BC, Canada, 2009. *A Tiny Bit More Rational Model of Categorization: The Influence of Test Items as Semi-Supervised Learning*

Invited Talk, Computer Science Seminars, Department of Computer & Information Science, Indiana University-Purdue University Indianapolis, Indianapolis, IN, 2009. *Computers Discover Wishes and Creativity in Text*.

Invited Talk, Merck & Co. Rahway, NJ, 2009. *Semi-Supervised Learning*.

Invited Talk, "Math, Algorithms, Learning, Brains, Engineering, Computing" (MALBEC) seminar series, University of Wisconsin Department of Mathematics, 2009. *HAMLET*.

Invited Talk, IBM Thomas J. Watson Research Center. Yorktown Heights, NY, 2009. *HAMLET*.

Invited Talk, "Computation and Informatics in Biology and Medicine" (CIBM) seminar series, University of Wisconsin-Madison, 2009. *HAMLET*

Talk at NIPS 2008 workshop on Machine Learning Meets Human Learning. Whistler, BC, Canada, 2008. *Human Semi-Supervised Learning and Human Active Learning*.

Invited talk, Hot Topics Workshop: Multi-Manifold Data Modeling and Applications. The Institute for Mathematics and its Applications (IMA), University of Minnesota, MN, 2008. *Semi-Supervised Learning by Multi-Manifold Separation*.

Invited talk, Language Technologies Institute Seminar. Carnegie Mellon University, Pittsburgh, PA, 2008. *Text-to-Picture Synthesis*.

Invited talk, Workshop on natural language processing. University of Washington and the Information Sciences Institute at the University of Southern California, Seattle, WA, 2008.

Invited talk, The 40th Interface Symposium (annual conference on the interface of computing science and statistics), Durham, NC, 2008.

Presentation and demo, University of Wisconsin Cognitive Science Conference Hertz Foundation poster session, Madison, WI, 2008.

Invited talk, Computer Science and Engineering Department Colloquium, Michigan State University, Lansing, MI, 2007. *Semi-Supervised Learning in Computers and Humans*.

Invited talk, Department of Statistics, University of Michigan, Ann Arbor, MI, 2007. *Semi-Supervised Learning in Computers and Humans*.

Invited alumnus talk, Language Technology Institute Retreat, Carnegie Mellon University, Pittsburgh, PA, 2007.

Invited talk, Psychology Department, University of Wisconsin, Madison, WI, 2007.

Invited participant, BIRS workshop of mathematical programming in machine learning and data mining, Banff, Canada, 2007.

Invited talk, Joint Statistical Meetings (JSM), Seattle, WA, 2006.

Invited talk, Electrical and Computer Engineering Department, University of Wisconsin, Madison, WI, 2006.

Invited talk, Computer Science and Engineering Department, Washington University in St. Louis, MO, 2006.

Invited talk, Statistics Department, University of Wisconsin, Madison, WI , 2006.

Invited talk, University of Cambridge, UK, 2004.

Invited talk, Gatsby Computational Neuroscience Unit, University College London, UK, 2004.

Invited talk, Microsoft Research Cambridge, UK, 2004.

Invited talk, NSF Aladdin Workshop on Graph Partitioning in Vision and Machine Learning, Pittsburgh, PA, 2003.

Media Coverage

Computer Sciences Faculty Jerry Zhu Guides UW-Madison's RISE-AI Initiative. By Sara Nelson. 2025.

Four Computer Sciences professors earn cross-campus interdisciplinary Research Forward 2024 Awards. 2024.

Professor Jerry Zhu honored with Kellett Mid-Career Award. 2024.

Jerry Zhu selected for DARPA ISAT study group. 2019.

A new approach for steganography among machine learning agents. By Ingrid Fadelli, Tech Xplore. Jan 4, 2019.

Squashing the bugs in machine learning: Researchers make computer-trained models more trustworthy. By Jennifer Smith, March 8, 2018.

UW-Madison researchers tackle bias in algorithms.

- Wisconsin State Journal. UW software aims to find and fix biased computer programs. By STEVEN VERBURG, Jul 10, 2017.
- University of Wisconsin-Madison News. UW-Madison researchers tackle bias in algorithms. By Jennifer Smith, July 3, 2017

University of Wisconsin-Madison News. Learning like humans, machines extend the reach of research. By Chris Barncard, August 16, 2016

The Chronicle of Higher Education. 'Machine Teaching' Is Seen as Way to Develop Personalized Curricula. August 12, 2015 by Mary Ellen McIntire.

- University of Wisconsin-Madison News, Machine teaching holds the power to illuminate human learning, by Jennifer A. Smith, August 10, 2015
- Inside Higher ED. In the Mind of a Student. By Jacqueline Thomsen, September 25, 2015
- Science 2.0. Machine Learning? No, Machine Teaching. August 13, 2015
- The R&D Magazine. The Flipside of Machine Learning. By Greg Watry, August 13, 2015

CBS News. Scientists using social media to track air pollution in China. By Michael Casey, Nov. 21, 2014.

- The Badger Herald, UW researchers use social media to estimate air quality, by Emily Neinfeldt, Nov. 20, 2014
- University of Wisconsin-Madison News, Social media for social good: Researchers estimate air pollution from online posts, by Jennifer Smith, Nov. 17, 2014

University of Wisconsin-Madison News, Learning machines scour Twitter in service of bullying research, by Chris Barncard, August 1, 2012

- Huffington Post, Bullying On Twitter: Researchers Find 15,000 Bully-Related Tweets Sent Daily (STUDY), by Britney Fitzgerald, August 2, 2012
- Gigaom.com, Big data as a tool for detecting (and punishing?) bullies, by Derrick Harris, August 2, 2012
- RedOrbit.com, Twitter Helps Researchers Study Factors Of Bullying, by Connie K. Ho, August 2, 2012
- Milwaukee Journal Sentinel, UW-Madison culling tweets about bullying, by Arthur Thomas, August 4, 2012
- The Capital Times, Campus Connection: UW-Madison researchers scour Twitter for bullying language, by Todd Finkelmeyer, August 7, 2012
- NBC Latino, Bullying prevention summit shines light on bystanders, role of technology, by Adrian Carrasquillo, August 7, 2012
- Time.com, Using Twitter to Crack Down on Bullying, by Kayla Webley, August 17, 2012
- WKOW, UW researchers create method to track bullying on Twitter, by Dani Maxwell, August 17, 2012
- Allvoices.com, Researchers leverage Twitter to identify cyberbullying, by Steve Kinney, August 21, 2012
- OnWisconcin, "Bullies Exposed / Social media reveals bad behavior offline", by Chris Barncard, p.11, Winter 2012

Carnegie Mellon University School of Computer Science, Alumni Snapshots in The Link magazine, p.26, Spring 2011.

NewScientist, "Picture This", August 18-24, p.22, 2007.

Teaching

Computer Sciences Department, University of Wisconsin, Madison, WI 2005–present
Professor
CS 540 – Introduction to Artificial Intelligence. 2005–present

CS 639 – Topics in Game Theory and Learning. Fall 2024, 2025
CS 639 – Topics in Sequential Decision Making and Learning. Spring 2021, 2025
CS 731 – Advanced Artificial Intelligence. Spring 2011
CS 760 – Machine Learning. 2019–present
CS 761 – Mathematical Foundations of Machine Learning. 2012–present
CS 769 – Advanced Natural Language Processing. Spring 2008, 2009, 2010
CS 838 – Topics in Advanced Natural Language Processing. Spring 2006, 2007
CS 839 – Topics in Sequential Decision Making. Spring 2024
CS 839 – Topics in Theory of Reinforcement Learning. Spring 2025
CS 861 – Theoretical Foundations of Machine Learning. 2018–present

Tsinghua University, Beijing, China 2014
7.5-hour seminar course instructor, "Topics in Statistical Learning Theory"

Tsinghua University, Beijing, China 2013
5-day seminar course instructor, "Exploring Topology for Machine Learning"

Zhiyuan College, Shanghai Jiao Tong University, Shanghai, China 2012
Instructor (co-taught with Eric Xing), "Machine Learning"

The Symposium on Educational Advances in AI (EAAI-11), San Francisco, CA August 10, 2011
Panelist, "Teaching Challenges in the Classroom"

Center of Automatic Learning and Discovery, Carnegie Mellon University June 16, 2004
Instructor, Learning from Labeled and Unlabeled Data, CALD Summer School.

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA Fall 2000
Teaching Assistant, 15-681 – Machine Learning

Student Advising

Current Graduate Students

Ara Vartanian, 2015–present, University of Wisconsin Computer Sciences Department

Former Students (Degree Year, First Employment)

PhDs

Shubham Bharti, PhD 2025. Google

Jeremy McMahan, PhD 2025.

Yiding Chen, PhD 2023. Postdoc, Cornell University

Young Wu, PhD 2023. Teaching Faculty, Department of Computer Sciences, University of Wisconsin-Madison

Yuzhe Ma, PhD 2021. Microsoft

Xuezhou Zhang, PhD 2021. Assistant Professor, Computing & Data Sciences, Boston University

Ayon Sen, PhD 2020. Facebook

Scott Alfeld, PhD 2017 (co-advised with Paul Barford). Assistant Professor, Amherst College

Bryan Gibson, PhD 2015. Voxgov

Kwang-Sung Jun, PhD 2015. Assistant Professor, Computer Science Department, University of Arizona.

Junming Xu (Sui), PhD 2015. Google.

David Andrzejewski (co-advised with Mark Craven), PhD 2010, Postdoc, Lawrence Livermore National Laboratory

Andrew Goldberg, PhD 2010, Senior Scientist, Arcode

Masters and Visitors

Gorune Hrag Ohannessian, MS 2016.

Shike Mei, MS 2015. Facebook.

Yimin Tan, MS 2014, Software Engineer, Twitter

Nick Bridle, MS 2013, Software Engineer, Google

Zhiting Xu, MS 2010, Google

Lijie Heng, MS 2008, Software Engineer, Oracle

Ming Li, Visiting Student 2008, Assistant Professor, Nanjing University

Jake Rosin (co-advised with Prof. Dyer), MS 2012

Jurgen Van Gael, MS 2007, Graduate student, University of Cambridge

Bachelors

Yancheng Zhu (CS, Wisconsin). 2023–2025

Joy Shiyun Cheng (CS, Wisconsin). 2023–2025

Jeffrey Kent Weisinger (CS, Wisconsin). 2023–2025

Luke Brandl (CS, Wisconsin). Graduate School in Computer Science at University of Michigan after graduation, 2014-2015.

Ben Burchfiel (CS, Wisconsin). NSF REU. Graduate School in Computer Science at Duke University after graduation, 2012–2013.

Alex Furger (Math, Wisconsin). NSF REU, winner of $2000 Computer Sciences Departmental Summer Fellowship under my supervision in 2010, and recipient of $2000 College of Letters & Science David Durra Scholarship in 2011. 2009-2011. Graduate School in Princeton ORFE after graduation.

Evan Hernandez (CS+Math, Wisconsin). DeWitt Scholarship for excellence in computer science and was one of six UW-Madison students selected to present his senior thesis to the Wisconsin State Legislature at the UW System's annual Research in the Rotunda event. Software engineer for Google after graduation, 2018.

Prasad Kawthekar (CS, Wisconsin), 2016. Graduate School in Computer Science at Stanford University.

Valerie Lo (CS, Wisconsin), 2007

Molly Maloney (Art, Wisconsin), NSF REU, 2009

Michael Maynord (CS, Wisconsin), NSF REU 2011–2012. Graduate School in Computer Science at the University of Maryland after graduation.

Rachael McCormick (CS and Psychology, Wisconsin). Winner of Maverick Software Scholarship in 2010, and winner of $3000 Summer Senior Honors Thesis Grant under my supervision in 2010. 2011

Mia Mueller (Fine Art, Wisconsin). NSF REU. 2009-2010.

Conor Perreault (CS, Wisconsin). Sophomore Research Fellowship. 2020

Ruichen Qian (Economics, Wisconsin), Undergraduate Research Scholars Program, 2007-2010. Morgan Stanley after graduation.

Bradley R. Strock (CS, Wisconsin), 2007

Xiaoxi Sun (CS, Wisconsin), 2022

Steve Yazicioglu (ECE and CS, Wisconsin). Senior honors thesis advisee. 2009-2010. Microsoft after graduation.

Dake Zhang (Mathematics and Economics, Wisconsin). 2013. Graduate Student Computational Finance Program at Carnegie Mellon University.

Thesis Examination Committee Member

Soumya Ray, PhD.'05, *Learning from data with Complex Interactions and Ambiguous Labels*, University of Wisconsin Computer Sciences Dept.

Mankyu Sung, PhD.'05, *Scalable, Controllable, Efficient and Convincing Crowd Simulation*, University of Wisconsin Computer Sciences Dept.

Guodong Guo, PhD.'06, *Face, Expression, and Iris Recognition Using Learning-based Approaches*, University of Wisconsin Computer Sciences Dept.

Shaohua Fan, PhD.'06, *Sequential Monte Carlo Methods for Physically Based Rendering*, University of Wisconsin Computer Sciences Dept.

Pedro Bizarro, PhD.'06, *Adaptive Query Processing: Dealing with Incomplete and Uncertain Statistics*, University of Wisconsin Computer Sciences Dept.

Ye Chen, PhD.'07, *A Bayesian Network Model of Knowledge-Based Authentication*, University of Wisconsin Operations and Information Management Dept.

Michael Wallick, PhD.'07, *Automatic Organization of Large Collections of Photographs*, University of Wisconsin Computer Sciences Dept.

Jesse Davis, PhD.'07, *View Learning: A Statistical Relational Approach to Mining Biomedical Databases*, University of Wisconsin Computer Sciences Dept.

Edward W. Wild, PhD.'08, *Optimization-Based Machine Learning and Data Mining*, University of Wisconsin Computer Sciences Dept.

Aarti Singh, PhD.'08. *Nonparametric Set Estimation Problems in Statistical Inference and Learning*, University of Wisconsin Department of Electrical and Computer Engineering

Hector Corrada Bravo, PhD.'08. *Graph-Based Data Analysis*, University of Wisconsin Computer Sciences Dept.

Yong Lu, PhD.'08. *A Computational Framework for the Analysis of Multi-Species Microarray Data*, Carnegie Mellon University, Computer Science Dept.

Mugizi Robert Rwebangira, PhD.'08. *Techniques for Exploiting Unlabeled Data*, Carnegie Mellon University, Computer Science Dept.

Burr Settles, PhD. '08. *Curious Machines: Active Learning with Structured Instances*, University of Wisconsin Computer Sciences Dept.

Lisa Torrey, PhD. '09. *Relational Transfer in Reinforcement Learning*, University of Wisconsin Computer Sciences Dept.

Louis Oliphant, PhD. '09. *Adaptively Finding and Combining First-Order Rules for Large, Skewed Data Sets*, University of Wisconsin Computer Sciences Dept.

Yu-Chi Lai, PhD.'10. *Photorealistic Animation Rendering with Population Monte Carlo Energy Redistribution*, University of Wisconsin Computer Sciences Dept.

Andrew Goldberg, PhD.'10. *New Directions in Semi-Supervised Learning*, University of Wisconsin Computer Sciences Dept., Advisor.

Gregory Cipriano, PhD. '10. *Molecular Surface Abstraction*, University of Wisconsin Computer Sciences Dept.

David Andrzejewski, PhD. '10. *Incorporating Domain Knowledge In Latent Topic Models*, University of Wisconsin Computer Sciences Dept., Advisor (co-advised with Mark Craven).

Arup Dutta, PhD. '11. *Artificial Neural Network Approach to Crash Modeling and Prediction*, University of Wisconsin Department of Civil and Environmental Engineering

Mark Wayne Liu, MS. '11. *Developing Methods to Merge Information from Multiple Sensors for Improved Crop Identification*, Environmental Studies, University of Wisconsin.

SangKyun Lee, PhD. '11. *Optimization Methods for Regularized Convex Formulations in Machine Learning*, University of Wisconsin Computer Sciences Dept.

Gregory Druck, PhD. '11. *Generalized Expectation Criteria for Lightly Supervised Learning*, University of Massachusetts at Amherst

Nathan Rosenblum, PhD. '11. *The Provenance Hierarchy of Computer Programs*, University of Wisconsin Computer Sciences Dept.

Yi Zhang, PhD. '12. *Learning with Limited Supervision by Input and Output Coding*, Carnegie Mellon University, Machine Learning Department, School of Computer Science.

Mariyam Mirza, PhD. 2012. *A Machine Learning Based Approach to Problems in Computer Network Measurement and Performance Analysis*, University of Wisconsin Computer Sciences Dept.

Shilin Ding, PhD. 2012. *Learning Graph Structure with Parametric and Nonparametric Models*, University of Wisconsin Statistics Dept.

Frank Lin, PhD 2012. *Scalable Methods for Graph-Based Unsupervised and Semi-Supervised Learning*, Carnegie Mellon University, Language Technology Institute, School of Computer Science.

Houssam Nassif, PhD. 2012. *Relational Differential Prediction*, University of Wisconsin Computer Sciences Dept.

Piramanayagam Arumuga Nainar, PhD 2012. *Applications of Static Analysis and Program Structure in Statistical Debugging*, University of Wisconsin Computer Sciences Dept.

Jeremy C. Weiss, PhD 2014. *Statistical Timeline Analysis for Electronic Health Records*, University of Wisconsin Computer Sciences Dept.

Anthony Douglas McDonald, PhD 2014. *Improving Driver Drowsiness Detection through Temporal, Contextual, and Hierarchical Modeling* , University of Wisconsin, Department of Industrial and Systems Engineering.

Jie Liu, PhD 2014. *Statistical Methods for Genome-wide Association Studies and Personalized Medicine*, University of Wisconsin Computer Sciences Dept.

Josh LaRocque, PhD 2014. *Exploring neural representations in short-term memory and visual awareness*, Neuroscience, University of Wisconsin Neuroscience.

Gautam Dasarathy, PhD 2014. *Data Efficient and Robust Algorithms for Reconstructing Large Graphs*, University of Wisconsin Department of Electrical and Computer Engineering

Ji Liu, PhD 2014. *Linearly Convergent Stochastic Algorithms for Optimization and Linear Systems*, University of Wisconsin Computer Sciences Dept.

Kendrick Boyd, PhD 2014. *Mitigating the Risks of Thresholdless Metrics in Machine Learning Evaluation*, University of Wisconsin Computer Sciences Dept.

Aditya Thakur, PhD 2014. *Symbolic Abstraction: Algorithms and Applications*, University of Wisconsin Computer Sciences Dept.

Shenqi Zhu, PhD 2014. *A Context Modeling Approach for Image Labeling Problems in Computer Vision*, University of Wisconsin Computer Sciences Dept.

Deborah Chasman, PhD 2014. *Improving the Interpretability of Integer Linear Programming Methods for Biological Subnetwork Inference*, University of Wisconsin Computer Sciences Dept.

Deepti Pachauri, PhD 2015. *Group-theoretic Algorithms for Matching Problems with Applications to Computer Vision*, University of Wisconsin Computer Sciences Dept.

Young-Bum Kim, PhD 2015. *Natural Language Technologies for Low-Resource Languages*, University of Wisconsin Computer Sciences Dept.

Tai Qin, PhD 2015. *Statistical Justifications for Computationally Tractable Network Data Analysis*, University of Wisconsin Statistics Dept.

Junming (Xu) Sui, PhD 2015. *Understanding and Fighting Bullying with Machine Learning*. Department of Computer Sciences, University of Wisconsin. Advisor.

Allison Sauppé, PhD 2015. *Designing Effective Communication Strategies For Human-Robot Collaboration*, University of Wisconsin Computer Sciences Dept.

Chaitanya Gokhale, PhD 2015. *Corleone: Hands-Off Crowdsourcing for Entity Matching*, University of Wisconsin Computer Sciences Dept.

Jia Xu, PhD 2015. *Visual Parsing with Weak Supervision*, University of Wisconsin Computer Sciences Dept.

Kwang-Sung Jun, PhD 2015. *Some Machine Learning Methods from Sequential Input*, University of Wisconsin Computer Sciences Dept. Advisor.

Bryan Gibson, PhD 2015. *Using Machine Learning to Understand and Influence Human Categorization Behavior*, University of Wisconsin Computer Sciences Dept.

Chien-Ming Huang, PhD 2015. *Human-Robot Joint Action: Coordinating Attention, Communication, and Actions*, University of Wisconsin Computer Sciences Dept.

Kevin Jamieson, PhD 2015. *The Analysis of Adaptive Data Collection Methods for Machine Learning*, University of Wisconsin Department of Electrical and Computer Engineering.

Xuezhi Wang, PhD 2016. *Active Transfer Learning*, Carnegie Mellon University, Computer Science Department, School of Computer Science.

Alex Hanna, PhD 2016. *Automated Coding of Protest Event Data: Development and Applications*. Sociology, University of Wisconsin–Madison.

Arun Kumar, PhD 2016. *Learning over Joins*, University of Wisconsin Computer Sciences Dept.

Newsha Ardalani, PhD 2016. *Cross-Architecture Performance Prediction Using Machine Learning*, University of Wisconsin Computer Sciences Dept.

Eric Alexander, PhD 2016. *Enabling Exploration and Hypothesis Formation within Topic Models*, University of Wisconsin Computer Sciences Dept.

Fei Du, PhD 2017. *Knowledge Integration in Geospatial Predictive Modeling*, University of Wisconsin Department of Geography

Hyunwoo J Kim, PhD 2017. *Statistical learning models for manifold-valued measurements with applications to computer vision and neuroimaging*, University of Wisconsin Computer Sciences Dept.

Han Chen, PhD 2017. *TENSOR-RELATED METHODS IN HIGH DIMENSIONAL STATISTICS*, University of Wisconsin Department of Statistics.

Luwan Zhang, PhD 2017. *Topics on Euclidean Distance Matrix and Unsupervised Ensemble Learning*, University of Wisconsin Department of Statistics.

Song Wang, PhD 2017. *Spectral Methods for Community Detection*, University of Wisconsin Department of Statistics.

Maxwell Collins, PhD 2017. *Scalable Optimization Methods with Side Information in Image Understanding*, University of Wisconsin Computer Sciences Dept.

Aniruddha Bhargava, PhD 2017. *Bandits and Preference Learning*, University of Wisconsin Department of Electrical and Computer Engineering.

Vamsi Ithapu, PhD 2018. *Exploiting Structure for Designing Clinical Trials: Testing, Learning and Inference Algorithms*, University of Wisconsin Computer Sciences Dept.

(Zhaobin Kuang) Charles Kwong, PhD 2018. *Learning with High Causal Fidelity from Longitudinal Event Data*, University of Wisconsin Computer Sciences Dept.

Beilun Wang, PhD 2018. *Fast and Scalable Joint Estimators for Learning Sparse Gaussian Graphical Models from Heterogeneous Data with Additional Knowledge*, University of Virginia, Computer Science Department.

Kyubin Lee, PhD 2018. *Learning and Interpreting Models that Map Viral Genotypes to Host Disease Phenotypes*, University of Wisconsin Computer Sciences Dept.

Xiaozhu Meng, PhD 2018. *Fine-Grained Binary Code Authorship Analysis: Identification and Evasion*, University of Wisconsin Computer Sciences Dept.

Sid Kiblawi, PhD 2019. *Augmenting Subnetwork Inference with Information Extracted from the Scientific Literature*, University of Wisconsin Computer Sciences Dept.

Ching-Pei Lee, PhD 2019. *Algorithms for Large-Scale Regularized Optimization*, University of Wisconsin Computer Sciences Dept.

Chao Wang, PhD 2019. *Data Driven Modeling, Monitoring and Control for Smart and Connected Systems*, Department of Industrial and Systems Engineering, University of Wisconsin-Madison

Sathya Ravi, PhD 2019. *The Wonderful World of Constraints in Learning and Vision*, University of Wisconsin Computer Sciences Dept.

Alireza Fotuhi Siahpirani, PhD 2019. *Computational methods for integrative inference of genome-scale gene regulatory networks*, University of Wisconsin Computer Sciences Dept.

Seong Jae Hwang, PhD 2019. *Latent Representation Learning for Understanding Relationships in Computer Vision and Neuroimaging*, University of Wisconsin Computer Sciences Dept.

Aubrey Barnard, PhD 2019. *Causal Discovery of Adverse Drug Events in Observational Data*, University of Wisconsin Computer Sciences Dept.

Wei Zhang, PhD 2020. *Knowledge Discovery from Event Sequences: A Point Process Perspective*, University of Wisconsin Computer Sciences Dept.

Yanghui Kang, PhD 2020. *Towards Operational Monitoring of the Agroecosystems with Satellite Remote Sensing: A Case Study in the Midwest U.S.*, Department of Geography, University of Wisconsin-Madison

Ayon Sen, PhD 2020. *Learning with the Help of a Teacher*, University of Wisconsin Computer Sciences Dept.

Moayad Ahmed Alnammi, PhD 2021. *Iterative Batched Screening and Active Learning in Drug Discovery*, University of Wisconsin Computer Sciences Dept.

Taylor Keding, PhD 2021. *Early-Life Exposure to Violence and Fronto-Amygdala Circuit Maturation: Developmental Markers of Psychiatric Risk*, University of Wisconsin Psychiatry Dept.

Mary Phuong, PhD 2021. *Underspecification in Deep Learning*, Institute of Science and Technology Austria.

Dianjing Liu, PhD 2021. *The Application of Machine Learning for Designing and Controlling Electromagnetic Fields*, Department of Electrical and Computer Engineering, University of Wisconsin-Madison

Xuezhou Zhang, PhD 2021. *Adversarial Learning in Sequential Decision Making*, University of Wisconsin Computer Sciences Dept. Advisor.

Alexander Brooks, PhD 2021. *Novice Language Adaption in Social Media Forums*, University of Wisconsin Computer Sciences Dept.

Yuzhe Ma, PhD 2021. *Adversarial Attacks in Sequential Decision Making and Control*, University of Wisconsin Computer Sciences Dept. Advisor.

Xiaomin Zhang, PhD 2021. *Statistical Learning along with Clustering*, University of Wisconsin Computer Sciences Dept.

Muni Sreenivas Pydi, PhD 2022. *Adversarial Robustness in Machine Learning: An Optimal Transport Perspective*, University of Wisconsin Department of Electrical and Computer Engineering.

Vincent V. Frigo, PhD 2022. *An Examination of Non-Normative Belief Updating Behavior in Humans (Why Is It So Hard to Change Minds?)*, University of Wisconsin Department of Psychology.

Shashank Rajput, PhD 2023. *LARGE-SCALE SGD ALGORITHMS AND THE EXPRESSIVE POWER OF MODERN NEURAL NETWORKS*, University of Wisconsin Computer Sciences Dept.

Ankit Pensia, PhD 2023. *Efficient Statistical Inference Under Sampling and Computational Constraints*, University of Wisconsin Computer Sciences Dept.

Jakwang Kim, PhD 2023. *Adversarial robustness in classification via the lens of optimal transport: theory and numerics*, University of Wisconsin Department of Statistics.

Yinglun Zhu, PhD 2023, *Interactive Machine Learning: From Theory to Scale*, University of Wisconsin Computer Sciences Dept.

David Merrell, PhD 2023. *Probabilistic Machine Learning with Omics Data and Biological Prior Knowledge*, University of Wisconsin Computer Sciences Dept.

Yiyou Sun, PhD 2023. *Towards a Reliable Open-world Machine Learning System: Algorithm and Theory*, Department of Computer Sciences, University of Wisconsin-Madison

Yiwu Zhong, PhD 2023. *Learning Visual Knowledge from Natural Language Supervision*, Department of Computer Sciences, University of Wisconsin-Madison

Young Wu, PhD 2023. *Attacks and Defense on Normal-Form Games and Markov Games*, University of Wisconsin Computer Sciences Dept. Advisor.

Kartik Sreenivasan, PhD 2023. *Towards Understanding the Challenges in Scaling Frontier Machine Learning Models*, University of Wisconsin Department of Electrical and Computer Engineering

Yiding Chen, PhD 2023. *Robust Decision-Making under Data Corruption*, Department of Computer Sciences, University of Wisconsin-Madison

Sean Yun-Shiuan Chuang, PhD 2024. *Simulating Human Opinion Dynamics Using AI agents and Large Language Models*, University of Wisconsin Department of Psychology.

Zhenmei Shi, PhD 2024. *UNDERSTANDING TRAINING AND ADAPTATION IN FEATURE LEARNING: FROM TWO-LAYER NETWORKS TO FOUNDATION MODELS*, Department of Computer Sciences, University of Wisconsin-Madison

Ziqian Lin, PhD 2025. *Understanding In-Context Learning via Synthetic and Controllable Datasets*, Department of Computer Sciences, University of Wisconsin-Madison

Harit Vishwakarma, PhD 2025. *Tackling the Data Bottleneck and Inference Challenges in ML*, Department of Computer Sciences, University of Wisconsin-Madison

Sean Xuefeng Du, PhD 2025. *Foundations of Unknown-aware Machine Learning*, Department of Computer Sciences, University of Wisconsin-Madison

Nikos Zarifis, PhD 2025. *Robustly Learning Multi-Index Models*, Department of Computer Sciences, University of Wisconsin-Madison

Jihyun Rho, PhD 2025. *Supporting Students' Resistance to Misleading Data Visualizations through Educational Approaches*, Department of Educational Psychology, University of Wisconsin-Madison

Jifan Zhang, PhD 2025. *Make Active Learning Terrific Again: Methods and Applications for Data Imbalance, Large Language Models, and Human Preferences*, Department of Computer Sciences, University of Wisconsin-Madison

Matthew Dutson, PhD 2025. *Augmenting Frame-Based Vision With Temporal Context*, Department of Computer Sciences, University of Wisconsin-Madison

Sourav Pal, PhD 2025. *Efficient Learn the Dynamics Modules are All You Need*, Department of Computer Sciences, University of Wisconsin-Madison

Nicholas Roberts, PhD 2026. *The Science of Foundation Models for Science*, Department of Computer Sciences, University of Wisconsin-Madison

Nicholas Corrado, PhD 2026. *Data Collection for Sample-Efficient Reinforcement Learning*, Department of Computer Sciences, University of Wisconsin-Madison

Alex Bocchieri, PhD 2026. *High-Energy Radiation Imaging with a Single-Photon Camera*, Department of Computer Sciences, University of Wisconsin-Madison

Thesis Proposal (Preliminary Examination) Committee Member

Su Zhang, ABD'05. *Network Traffic Characterization*, University of Wisconsin Computer Sciences Dept.

Trevor Walker, ABD'07. *Relational Methods Incorporating Domain Knowledge for Transfer in Reinforcement Learning*, University of Wisconsin Computer Sciences Dept.

Jian Liu, ABD'08. *Mapping Soil Variation with Satellite-based Observations of Surface Dynamics*, University of Wisconsin Department of Geography

Feng Liu, ABD'08. *Synthesizing Novel Multimedia from Images and Videos*, University of Wisconsin Computer Sciences Dept.

Nathanael Fillmore, ABD'13. *Generative models for transcriptome assembly and analysis*, University of Wisconsin Computer Sciences Dept.

Alexander Cobian, ABD'14. *Rich Representations of EHR Data: Inferring Latent Variables and Temporal Intervals to Improve Risk Assessment*, University of Wisconsin Computer Sciences Dept.

Yuan Wang, ABD'15. *Topological Data Analysis in Electroencephalographic Studies*, University of Wisconsin Department of Statistics.

Bui Thi Mai (Mary) Phuong, ABD'18. *Strongly Supervised Learning*, Institute of Science and Technology Austria.

Samuel Drews, ABD'19. *Fairness, Correctness, and Automation*, University of Wisconsin Computer Sciences Dept.

Ara Vartanian, ABD'19. *Machine Teaching from the Perspective of Optimal Control*, University of Wisconsin Computer Sciences Dept. Advisor.

Liam Johnston, ABD'20. *Controlling Gradient Size through the Coadjoint*, University of Wisconsin Department of Statistics.

Changhun Jo, ABD'20. *Poisoning attacks on fairness*, University of Wisconsin Department of Mathematics.

Ankit Pensia, ABD'21. *Efficient and Robust Algorithms for Estimation*, University of Wisconsin Computer Sciences Dept.

Bhumi Kumar, ABD 2022. *A Law of Iterated Logarithm for Multi-Agent Reinforcement Learning*, University of Wisconsin Department of Electrical and Computer Engineering

Yiwu Zhong, ABD 2022. *Learning Visual Knowledge from Natural Language Supervision*, University of Wisconsin Computer Sciences Dept.

Jinmeng Rao, ABD 2022. *TRAJECTORY PRIVACY PROTECTION WITH GEOSPATIAL ARTIFICIAL INTELLIGENCE*, Department of Geography, University of Wisconsin-Madison

Jihyun Rho, ABD 2023. *Understanding and Enhancing Cognitive Processes in Misleading Data Visualizations*, University of Wisconsin Department of Psychology.

Matthew Dutson, ABD 2023. *Change-Driven Computer Vision*, Department of Computer Sciences, University of Wisconsin-Madison

Jifan Zhang, ABD 2023. *Deep Active Learning in the Real World: Beyond Benchmark Overfitting*, Department of Computer Sciences, University of Wisconsin-Madison

Jeremy McMahan, ABD 2023. *Safe Multi-agent Reinforcement Learning in Polynomial Time*, Department of Computer Sciences, University of Wisconsin-Madison

Harit Vishwakarma, ABD 2024. *Towards Efficient and Reliable ML, from Data Annotation to Deployment*, Department of Computer Sciences, University of Wisconsin-Madison

Ziqian Lin, ABD 2024. *Dual Operating Modes of In-context Learning*, Department of Computer Sciences, University of Wisconsin-Madison

Nicholas Corrado, ABD 2024. *Data Collection in Reinforcement Learning*, Department of Computer Sciences, University of Wisconsin-Madison

Thanasis Pittas, ABD 2024. *Robust statistical inference under computational constraints*, Department of Computer Sciences, University of Wisconsin-Madison

Alex Bocchieri, ABD 2024. *Imaging gamma-ray interactions in a scintillator with a single photon (SPAD) camera*, Department of Computer Sciences, University of Wisconsin-Madison

Joon Suk Huh, ABD 2024. *Mechanism Design Study*, Department of Computer Sciences, University of Wisconsin-Madison

Shubham Bharti, ABD 2024. *Optimal Teaching under Constraints*, Department of Computer Sciences, University of Wisconsin-Madison

Matthew Zurek, ABD 2025. *Statistical Foundations of Average-Reward RL*, Department of Computer Sciences, University of Wisconsin-Madison

Nicholas Roberts, ABD 2025. *The Science of LLM Scaling Laws*, Department of Computer Sciences, University of Wisconsin-Madison

Chenyang Jiang, ABD 2025. *Robust Reinforcement Learning under Diffusion Models for Data with Jumps*, University of Wisconsin Department of Statistics.

Outreach

Planning your Career in the Age of AI: A Panel Discussion. Hosted by the Data Science Institute, UW-Madison. March 17, 2026

Panel on AI Meets Society (AIMS) Symposium. Hosted by the Wisconsin AI Safety Initiative (WAISI) and UW–Madison's School of Computer, Data & Information Sciences (CDIS). Feb 21, 2026.

Invited talk at Soundwaves. *Reinventing the Wheels: A fairy tale about how AI learned to write*. Wisconsin Science Festival. October 18, 2024.

Panelist, *A Conversation on AI and Society*, Wisconsin AI Safety Initiative (Student Organization at UW-Madison) Oct 2, 2024.

Panelist, *Advocacy in the AM on AI*, Greater Madison Chamber of Commerce. June 12, 2024.

Keynote talk at Teaching Writing in the Age of Chat GPT Symposium. *Of Parrots and Monkeys: A Behind the Scenes Tour Through AI*. About 70 high school teachers and educators from Wisconsin. July 25, 2023.

Invited talk at Exploring Artificial Intelligence @ UW-Madison: A summer 2023 webinar series. *From Breakthroughs to Empowerment: UW-Madison's AI Contributions and Accessibility*. About 200 participants from campus. July 14, 2023.

Panel on The Rise of AI and the Impact on College Campuses, Honors Program College of Letters and Science and Chadbourne Residential College, UW-Madison. 2023

Ask Me Anything for Computer Science undergraduate students at UW-Madison, 2020.

Seminar for AI@UW student organization at UW-Madison, 2020.

Guest lecture on AI for the Pre-College Enrichment Opportunity Program for Learning Excellence (PEOPLE), 2013

WARF Discovery Challenge judge, 2013

The North American Computational Linguistics Olympiad (NACLO). 2011.

Presentation to Wisconsin high school students "Meet Computer Science Professors" field trip to University of Wisconsin–Madison, 2010.

The North American Computational Linguistics Olympiad (NACLO). 2010.

The North American Computational Linguistics Olympiad (NACLO). Sponsored in part by NSF, the contest reaches out to high school students with challenging computational linguistic problems. 2009.

Service

University Service

2026. Physical Sciences Research Committee

Since 2022. Founder, Machine Learning Lunch Meeting series. Weekly university professor speakers from Computer Sciences, ECE, BMI, Physics, etc. Open to all. Typical audience: graduate and undergraduate students, faculty and staff.

Since 2023. RISE-AI Thought Leaders Team, UW-Madison

2024. CIO and Vice Provost for IT search committee

2024. Faculty search committee for Computational Approaches to Language Cluster Hire

2024. Faculty search committee for Musci+AI

2019. External review committee for the Language Science program

2019. Faculty search committee for Cluster Hiring Initiative: Opening Doors Through Language, 2019

2019, 2021. Faculty search committee for Cluster Hiring Initiative: Foundations of Data Science

2019. Faculty search committee for SMPH-Machine Learning in Biostatistics and Medical Informatics

2015. Wisconsin Institute for Discovery Optimization (WID/OPT) affiliate

2014-2017. Executive Leadership Committee of the McPherson Eye Research Institute

2009-2012. University of Wisconsin Eye Research Institute Member

2007. Undergraduate Research Scholars Advisor, University of Wisconsin-Madison

Department Service

Since 2012, 2018-2020, 2021 (co-chair), 2022 (chair), 2023, 2024 (co-chair). Faculty Recruiting Committee, Computer Sciences Department, University of Wisconsin-Madison

Since 2019. Faculty mentor

Since 2019. Tenure-track faculty annual review committees

Since 2019. Faculty post-tenure review committees

2019–present. Department of Biostatistics & Medical Informatics Liaison

2024, 2025. Awards Committee

2024. Full Professor Promotion Committee

2024. Named Professor Search Committee

2020-2021. Space Committee, Computer Sciences Department, University of Wisconsin-Madison

2013, 2014, 2018, 2019, 2020. Budget Committee, Computer Sciences Department, University of Wisconsin-Madison

2010-2011. Computer Sciences Department Distinguished Lecture Series, University of Wisconsin-Madison

2009-2012. Curriculum Committee, Computer Sciences Department, University of Wisconsin-Madison

2008, 2015. Graduate Advising Committee, Computer Sciences Department, University of Wisconsin-Madison

2008. University of Wisconsin Cognitive Science Cluster Faculty Search Committee

2005-2007, 2013-2017. Co-chair. Graduate Student Admissions Committee, Computer Sciences Department, University of Wisconsin-Madison

September 4, 2026