

Jennifer Neville
Microsoft Research Redmond
Purdue University, Departments of Computer Science and Statistics
West Lafayette, IN 47907-2107
Phone: (765) 496-9387 · Email: prof.jen.neville@gmail.com

Research Interests

My research focuses on understanding how machine learning and AI systems behave in realistic settings, from networks and relational data to modern AI assistants and agents. I am particularly interested in how complexity, interaction, and mismatches between modeling assumptions and deployment environments shape system behavior. My work combines empirical analysis, theory, evaluation, and algorithm design to identify and address these gaps.

Research Impact

150+ peer-reviewed publications and 12k+ citations
40 students mentored, including 22 PhD graduates
\$6M+ in direct research funding and participation in collaborative initiatives exceeding \$60M
135+ invited talks, including keynotes, distinguished lectures, and plenary presentations.

Education

Ph.D., University of Massachusetts Amherst, Computer Science, 2006. Advisor: David Jensen.

Professional Experience

2020–present:	Microsoft Research, Redmond, WA <i>Partner Research Manager</i> (2024–present) <i>Senior Principal Research Manager</i> (2022–2024) <i>Senior Principal Researcher</i> (2021–2022) <i>Visiting Researcher</i> (2020) Rebuilt and repositioned an established information retrieval research group into the AI Interaction and Learning team, focused on understanding the behavior of large-scale interactive AI systems. The multidisciplinary team spans technical, behavioral, and social perspectives on AI systems. Current research develops theory, evaluation methodologies, simulations, and learning algorithms to study AI behavior in realistic settings, including complex reasoning tasks, multi-turn interactions, and long-horizon workflows. The work combines controlled experimentation with analyses of deployed systems involving more than 300 million user interactions.
2006–present:	Purdue University, West Lafayette, IN Department of Computer Science (joint appointment in Statistics) <i>Samuel Conte Professor</i> (2020–present) <i>Professor</i> (2019–2020) <i>Miller Family Chair Associate Professor</i> (2014–2019) <i>Associate Professor</i> (2013–2014) <i>Assistant Professor</i> (2006–2013) Founded and led a research program in machine learning for relational and network data. Graduated 22 PhD students and produced foundational contributions in rela-

	tional learning, graph mining, and statistical methodology. Secured over \$6M in direct research funding and participated in large-scale research initiatives totaling more than \$60M.
2013	<i>Visiting Scientist</i>, Simons Institute for Theoretical Computing, UC Berkeley

Awards and Honors

Best Paper Award, 14th International Conference on Learning Representations (ICLR), 2026
 Program Committee Chair, 37th AAAI Conference on Artificial Intelligence, 2023
 Purdue University Faculty Scholar, 2020
 Program Committee Chair, 18th SIAM International Conference on Data Mining, 2019
 ACM Doctoral Dissertation Award Committee, 2019-2021
 AAAI Executive Council, Elected Councilor, 2015-2018
 Program Committee Chair, 9th ACM International Conference on Web Search and Data Mining, 2016
 NSF Career Award, 2012
 ICDM Best Research Paper Award Runner-Up, 2009
 IEEE Intelligent Systems Top Ten to Watch, 2008
 DARPA Computer Science Study Panel Member, 2007

Selected Keynotes and Distinguished Lectures

Selected from more than 135 invited talks, including keynote addresses, distinguished lectures, and plenary presentations.

6th International Joint Conference on Learning & Reasoning (IJCLR), 2026
 Midwest Machine Learning Symposium, 2026
 University of Southern California Symposium on the Future of Computing: A 25-Year Vision, 2025
 University of California San Diego Distinguished Lecture, Computer Science Department, 2020
 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2019
Future of Artificial Intelligence Plenary Panelist, 33rd AAAI Conference on Artificial Intelligence, 2019
 Virginia Tech University Distinguished Lecture, Computer Science Department, 2019
Societal Impact of Data Science and Artificial Intelligence Plenary Panelist, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018
 27th International Conference on Inductive Logic Programming, 2017
 3rd ACM International Conference on the Theory of Information Retrieval, 2017
 Max Planck Institute of Informatics Distinguished Lecture, 2016

Publications¹

Conference Papers

1. LLMs Corrupt Your Documents When You Delegate.
 P Laban, T Schnabel, J Neville.
Proceedings of the 3rd Conference on Language Modeling (COLM), 2026.
2. Reasoning about Reasoning: BAPO Bounds on Chain-of-Thought Token Complexity in LLMs.
 K Tomlinson, T Schnabel, A Swaminathan, J Neville.
Proceedings of the 43rd International Conference on Machine Learning (ICML), 2026.

¹My authorship policy is to list student authors ahead of me, regardless of contribution. Graduate students are denoted with †; undergraduates are denoted with ‡.

3. Wildfeedback: Aligning LLMs with in-situ user interactions and feedback.
T Shi[†], Z Wang[†], L Yang, YC Lin[†], Z He[†], M Wan, P Zhou, S Jauhar, S Chen, S Xia, H Zhang, J Zhao, X Xu, X Song, J Neville.
Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL), 2026.
4. LLMs get lost in multi-turn conversation.
P Laban, H Hayashi, Y Zhou, J Neville. **Best Paper Award**
Proceedings of the 14th International Conference on Learning Representations (ICLR), 2026.
5. Flipping the dialogue: Training and evaluating user language models.
T Naous[†], P Laban, W Xu, J Neville.
Proceedings of the 14th International Conference on Learning Representations (ICLR), 2026.
6. Lost in transmission: When and why LLMs fail to reason globally.
T Schnabel, K Tomlinson, A Swaminathan, J Neville.
Advances in Neural Information Processing Systems (NeurIPS), 2025.
7. Group Preference Alignment: Customizing LLM Responses from In-Situ Conversations.
I Mondal[†], JW Stokes, SK Jauhar, L Yang, M Wan, X Xu, X Song, J. Boyd-Graber, J. Neville.
Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2025
8. GenTool: Enhancing Tool Generalization in Language Models through Zero-to-One and Weak-to-Strong Simulation.
J He[†], J Neville, M Wan, L Yang, H Liu, X Xu, X Song, JZ Pan, P Zhou.
Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025.
9. Node Similarities under Random Projections: Limits and Pathological Cases.
T. Tadic, C. Becker, J. Neville.
Proceedings of the 13th International Conference on Learning Representations (ICLR), 2025.
10. Symbolic prompt program search: A structure-aware approach to efficient compile-time prompt optimization.
T. Schnabel and J. Neville.
Findings of the Association for Computational Linguistics (EMNLP), 2024.
11. TnT-LLM: Text mining at scale with large language models.
M. Wan, T. Safavi, S. Jauhar, Y. Kim, S. Counts, J. Neville, S. Suri, C. Shah, R. White, L. Yang, R. Andersen, G. Buscher, D. Joshi, N. Rangan.
Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2024.
12. Automatic Pair Construction for Contrastive Post-training.
C. Xu[†], C. Rosset, E. Chau, L. Corro, S. Mahajan, J. McAuley, J. Neville, A. Awadallah, N. Rao.
Findings of the Association for Computational Linguistics (NAACL), 2024.
13. On Overcoming Miscalibrated Conversational Priors in LLM-based Chatbots.
C. Herlihy[†], J. Neville, T. Schnabel, A. Swaminathan.
Proceedings of the 40th Conference on Uncertainty in Artificial Intelligence (UAI), 2024.
14. MS MARCO Web Search: a Large-scale Information-rich Web Dataset with Millions of Real Click Labels
Q. Chen, X. Geng, C. Rosset, C. Buractaon, J. Lu, T. Shen, K. Zhou, C. Xiong, Y. Gong, P. Bennett, N. Craswell, X. Xie, F. Yang, B. Tower, N. Rao, A. Dong, W. Jiang, Z. Liu, M. Li, C. Liu, Z. Li, R. Majumder, J. Neville, A. Oakley, K. Risvik, H. Simhadri, M. Varma, Y. Wang, L. Yang, M. Yang, C. Zhang.
Companion Proceedings of the ACM on Web Conference, 2024.

15. Interpretable User Satisfaction Estimation for Conversational Systems with Large Language Models.
Y. Lin[†], J. Neville, J. Stokes, L. Yang, T. Safavi, M. Wan, S. Counts, S. Suri, R. Andersen, X. Xu, D. Gupta, S. Jauhar, X. Song, G. Buscher, S. Tiwary, B. Hecht, and J. Teevan.
Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL), 2024.
16. DYANE: DYnamic Attributed Node rolEs Generative Model.
G. Zeno[†] and J. Neville.
Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM), 2023.
17. Hindsight Learning for MDPs with Exogenous Inputs.
S. Sinclair[†], F. Frujeri, C. Cheng, L. Marshall, H. Barbalho, J. Li, J. Neville, I. Menache, A. Swaminathan.
Proceedings of the 40th International Conference on Machine Learning (ICML), 2023.
18. Stationary Algorithmic Balancing Over Dynamic Email Re-Ranking Problem
J. Liu[†] and J. Neville.
Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023.
19. Workplace Recommendation with Temporal Network Objectives
K. Tomlinson, J. Neville, L. Yang, M. Wang, C. Lu
Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023.
20. Expressive and Efficient Representation Learning for Ranking Links in Temporal Graphs
S. Suresh[†] M. Shrivastava, A. Mukherjee, J. Neville, and P. Li.
Proceedings of the 32nd ACM Web Conference, 2023.
21. Adversarial Graph Augmentation to Improve Graph Contrastive Learning
S. Suresh[†] P. Li, C. Hao, and J. Neville.
Advances in Neural Information Processing Systems (NeurIPS), 2021.
22. A Collective Learning Framework to Boost GNN Expressiveness for Node Classification.
M. Hang[†] J. Neville, and B. Ribeiro.
Proceedings of the 38th International Conference on Machine Learning (ICML), 2021.
23. Breaking the Limit of Graph Neural Networks by Improving the Assortativity of Graphs with Local Mixing Patterns
S. Suresh[†] V. Budde, J. Neville, P. Li, J. Ma.
Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2021.
24. Towards Decentralized Social Reinforcement Learning via Ego-Network Extrapolation
M. Goindani[†] and J. Neville.
Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2021.
25. Dymond: Dynamic motif-nodes network generative model
G. Zeno[†] T. La Fond, and J. Neville.
Proceedings of the 30th International Web Conference, 2021.
26. A Hybrid Model for Learning Embeddings and Logical Rules Simultaneously from Knowledge Graphs
S. Suresh[†] and J. Neville.
Proceedings of the 20th IEEE International Conference on Data Mining (ICDM), 6 pages, 2020.
27. Online Bayesian Sparse Learning with Spike and Slab Priors
S. Fang, S. Zhe, K. Lee, K. Zhang, and J. Neville.
Proceedings of the 20th IEEE International Conference on Data Mining (ICDM), 8 pages, 2020.

28. MERL: Multi-View Edge Representation Learning in Social Networks
Y. Lai[†], and J. Neville.
Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2020.
29. Role Equivalence Attention for Label Propagation in Graph Neural Networks
H. Park[†] and J. Neville
Proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 12 pages, 2020.
30. Cluster-Based Social Reinforcement Learning
M. Goindani[†] and J. Neville.
Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 3 pages, 2020.
31. ReadNet: A Hierarchical Transformer Framework for Web Article Readability Analysis
C. Meng[†], M. Chen, J. Mao, and J. Neville
Proceedings of the 42nd European Conference on Information Retrieval (ECIR), 15 pages, 2020.
32. Exploiting Interaction Links for Node Classification with Deep Graph Neural Networks
H. Park[†] and J. Neville
Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI), 8 pages, 2019.
33. Learning How to Intervene in True News Diffusion to Combat Fake News Spread.
M. Goindani[†] and J. Neville.
Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2019.
34. HATS: A Hierarchical Sequence-Attention Framework for Inductive Set-of-Sets Embeddings.
C. Meng[†], J. Yang[†], B. Ribeiro, and J. Neville.
Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2019.
35. A Stein-Papangelou Goodness-of-Fit Test for Point Processes.
J. Yang[†], V. Rao and J. Neville.
Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTAT), 10 pages, 2019.
36. TransConv: Relationship Embedding in Social Networks.
Y. Lai[†], J. Neville, and D. Goldwasser.
Proceedings of the 33rd Conference on Artificial Intelligence (AAAI), 9 pages, 2019.
37. Multi-level hypothesis testing for populations of heterogeneous networks
G. Gomes[†], J. Neville, and V. Rao.
Proceedings of the 18th IEEE International Conference on Data Mining (ICDM), 6 pages, 2018.
38. The Indian Buffet Hawkes Process to Model Evolving Latent Influences.
X. Tan[†], V. Rao and J. Neville.
Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2018.
39. Exploring Student Check-In Behavior for Improved Point-of-Interest Prediction.
M. Hang[†], I. Pytlarz, and J. Neville.
Proceedings of the 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2018.
40. Goodness-of-fit Testing for Discrete Distributions via Stein Discrepancy.
J. Yang[†], Q. Liu, V. Rao and J. Neville.
Proceedings of the 35th International Conference on Machine Learning (ICML), 9 pages, 2018.
41. Nested CRP with Hawkes-Gaussian Processes.
X. Tan[†], V. Rao and J. Neville.
Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTAT), 10 pages, 2018.

42. Subgraph Pattern Neural Networks for High-Order Graph Evolution Prediction.
C. Meng[†], C. Mouli[†], B. Ribeiro, and J. Neville.
Proceedings of the 32nd Conference on Artificial Intelligence (AAAI), 9 pages, 2018.
43. Decoupling Homophily and Reciprocity with Latent Space Network Models
J. Yang[†], V. Rao and J. Neville
Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2017.
44. Unified Representation and Lifted Sampling for Generative Models of Social Networks
P. Robles[†], S. Moreno, and J. Neville
Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 8 pages, 2017.
45. Should We Be Confident in Peer Effects Estimated From Partial Crawls of Social Networks?
J. Yang[†], B. Ribeiro and J. Neville
Proceedings of the 11th International AAAI Conference on Weblogs and Social Media (ICWSM), 4 pages, 2017.
46. Deep Collective Inference
J. Moore[†] and J. Neville
Proceedings of the 31st Conference on Artificial Intelligence (AAAI), 9 pages, 2017.
47. Sampling of Attributed Networks from Hierarchical Generative Models
P. Robles[†], S. Moreno, and J. Neville
Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2016.
48. Efficient Graphlet Counting for Large Networks
N. Ahmed[†], J. Neville, R. Rossi[†], and N. Duffield.
Proceedings of the 15th IEEE International Conference on Data Mining (ICDM), 10 pages, 2015.
49. Overcoming Relational Learning Biases to Accurately Predict Preferences in Large Scale Networks
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the 24th International World Wide Web Conference (WWW), 11 pages, 2015.
50. Incorporating Assortativity and Degree Dependence into Scalable Network Models
S. Musmann[‡], J. Moore[‡], J. Pfeiffer III[†], and J. Neville
Proceedings of the 29th Conference on Artificial Intelligence (AAAI), 9 pages, 2015.
51. Composite Likelihood Data Augmentation for Within-Network Statistical Relational Learning
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the 14th IEEE International Conference on Data Mining (ICDM), 10 pages, 2014.
52. A Scalable Method for Accurate Sampling from Kronecker Models
S. Moreno, J. Pfeiffer III[†], and J. Neville
Proceedings of the 14th IEEE International Conference on Data Mining (ICDM), 10 pages, 2014.
53. Active Exploration in Networks: Using Probabilistic Relationships for Learning and Inference
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2014.
54. Graph Sample and Hold: A Framework for Big-Graph Analytics
N. Ahmed[†], N. Duffield, J. Neville, and R. Kompella
Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2014.
55. Attributed Graph Models: Modeling network structure with correlated attributes
J. Pfeiffer III[†], S. Moreno[†], T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the 23rd International World Wide Web Conference (WWW), 11 pages, 2014.

56. Network Hypothesis Testing Using Mixed Kronecker Product Graph Models
S. Moreno[†] and J. Neville
Proceedings of the 13th IEEE International Conference on Data Mining (ICDM), 6 pages, 2013.
57. Learning Mixed Kronecker Product Graph Models with Simulated Method of Moments
S. Moreno[†], J. Neville and S. Kirshner
Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 9 pages, 2013.
58. Collective Inference for Network Data with Copula Latent Markov Networks
R. Xiang[†] and J. Neville
Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM), 10 pages, 2013.
59. Modeling Dynamic Behavior in Large Evolving Graphs
R. Rossi[†], B. Gallagher, J. Neville and K. Henderson
Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM), 10 pages, 2013.
60. An Analysis of How Ensembles of Collective Classifiers Improve Predictions in Graphs
H. Eldardiry[†] and J. Neville
Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2012.
61. Fast Generation of Large Scale Social Networks While Incorporating Transitive Closures
J. Pfeiffer III[†], T. La Fond[†], S. Moreno[†], and J. Neville
Proceedings of the 4th ASE/IEEE International Conference on Social Computing (SocialCom), 12 pages, 2012.
62. The Impact of Communication Structure and Interpersonal Dependencies on Distributed Teams
T. La Fond[†], D. Roberts[‡], J. Neville, J. Tyler, and S. Connaughton
Proceedings of the 4th ASE/IEEE International Conference on Social Computing (SocialCom), 8 pages, 2012.
63. Network Sampling Designs for Relational Classification
N. Ahmed[†], J. Neville, and R. Kompella
Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM), 4 pages, 2012.
64. Time-Evolving Relational Classification and Ensemble Methods
R. Rossi[†] and J. Neville
Proceedings of the 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 12 pages, 2012.
65. Structured Comparative Analysis of Systems Logs to Diagnose Performance Problems
K. Nagaraj[†], C. Killian, and J. Neville.
Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2012.
66. Understanding Propagation Error and Its Effect on Collective Classification
R. Xiang[†] and J. Neville
Proceedings of the 11th IEEE International Conference on Data Mining (ICDM), 10 pages, 2011.
67. Correcting Bias in Statistical Tests for Network Classifier Evaluation
T. Wang, J. Neville, B. Gallagher, and T. Eliassi-Rad
Proceedings of the 21st European Conference on Machine Learning (ECML), 16 pages, 2011.
68. Relational Active Learning for Joint Collective Classification Models
A. Kuwadekar[†] and J. Neville
Proceedings of the 28th International Conference on Machine Learning (ICML), 8 pages, 2011.

69. Across-Model Collective Ensemble Classification
H. Eldardiry[†] and J. Neville
Proceedings of the 25th Conference on Artificial Intelligence (AAAI), 7 pages, 2011.
70. Methods to Determine Node Centrality and Clustering in Graphs with Uncertain Structure
J. Pfeiffer III[†] and J. Neville
Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM), 4 pages, 2011.
71. Relational Learning with One Network: An Asymptotic Analysis
R. Xiang[†] and J. Neville
Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTAT), 11 pages, 2011.
72. ERACER: A Database Approach for Statistical Inference and Data Cleaning
C. Mayfield[†], J. Neville, and S. Prabhakar
Proceedings of the 2010 ACM SIGMOD Conference (SIGMOD), 12 pages, 2010.
73. Predicting Prefix Availability in the Internet
R. Khosla[†], S. Fahmy, C. Hu, and J. Neville
Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM) Mini-Conference, 5 pages, 2010.
74. Randomization tests for distinguishing social influence and homophily effects
T. LaFond[†] and J. Neville
Proceedings of the 19th International World Wide Web Conference (WWW), 10 pages, 2010.
75. Modeling Relationship Strength in Online Social Networks
R. Xiang[†], J. Neville, and M. Rogati
Proceedings of the 19th International World Wide Web Conference (WWW), 10 pages, 2010.
76. Tied Kronecker Product Graph Models to Capture Variance in Network Populations
S. Moreno[†], S. Kirshner, J. Neville and S.V.N. Vishwanathan
Proceedings of the 48th Annual Allerton Conference on Communications, Control and Computing, 8 pages, 2010.
77. Using Transactional Information to Predict Link Strength in Online Social Networks
I. Kahanda[†] and J. Neville
Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media (ICWSM), 8 pages, 2009.
78. Evaluating Statistical Tests for Within-Network Classifiers of Relational Data.
J. Neville, B. Gallagher, and T. Eliassi-Rad. **Best Paper Award Runner-Up**
Proceedings of the 9th IEEE International Conference on Data Mining (ICDM), 10 pages, 2009.
79. Temporal-Relational Classifiers for Prediction in Evolving Domains.
U. Sharan[†] and J. Neville
Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 10 pages, 2008.
80. A Shrinkage Approach for Modeling Non-Stationary Relational Autocorrelation.
P. Angin[†] and J. Neville
Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 6 pages, 2008.
81. Pseudolikelihood EM for Within-Network Relational Learning.
R. Xiang[†] and J. Neville
Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 6 pages, 2008.
82. Database support for probabilistic attributes and tuples.
S. Singh[†], C. Mayfield[†], R. Shah, S. Prabhakar, S. Hambrusch, J. Neville, R. Cheng.
The 24th International Conference on Data Engineering (ICDE), 9 pages, 2008.

83. Bias/Variance Analysis for Relational Domains.
J. Neville and D. Jensen.
The 17th International Conference on Inductive Logic Programming, *Lecture Notes in Artificial Intelligence 4894* (ILP), pages 27-28, 2007.
84. Leveraging Relational Autocorrelation with Latent Group Models.
J. Neville and D. Jensen.
Proceedings of the 5th IEEE International Conference on Data Mining (ICDM), pages 322-329, 2005.
85. Using Relational Knowledge Discovery to Prevent Securities Fraud.
J. Neville, O. Simsek, D. Jensen, J. Komoroske, K. Palmer and H. Goldberg.
Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pages 449-458, 2005.
86. Dependency Networks for Relational Data.
J. Neville and D. Jensen.
Proceedings of the 4th IEEE International Conference on Data Mining (ICDM), pages 170-177, 2004.
87. Why Collective Inference Improves Relational Classification.
D. Jensen, J. Neville and B. Gallagher.
Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pages 593-598, 2004.
88. Simple Estimators for Relational Bayesian Classifiers.
J. Neville, D. Jensen and B. Gallagher.
Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM), pages 609-612, 2003.
89. Learning Relational Probability Trees.
J. Neville, D. Jensen, L. Friedland and M. Hay.
Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pages 625-630, 2003.
90. Avoiding Bias When Aggregating Relational Data with Degree Disparity.
D. Jensen, J. Neville and M. Hay.
Proceedings of the 20th International Conference on Machine Learning (ICML), pages 274-281, 2003.
91. Autocorrelation and Linkage Cause Bias in Evaluation of Relational Learners.
D. Jensen and J. Neville.
Proceedings of the 12th International Conference on Inductive Logic Programming (ILP), pages 101-116, 2002.
92. Linkage and Autocorrelation Cause Feature Selection Bias in Relational Learning.
D. Jensen and J. Neville.
Proceedings of the 19th International Conference on Machine Learning (ICML), pages 259-266, 2002.

Journal Articles

93. Using large language models to generate, validate, and apply user intent taxonomies.
C Shah, R White, R Andersen, G Buscher, S Counts, S Das, A Montazer, S Manivannan, J Neville, N Rangan, T Safavi, S Suri, M Wan, L Wang, L Yang.
ACM Transactions on the Web 19 (3), 1-29, 2025.
94. Generating Post-Hoc Explanations for Skip-gram-based Node Embeddings by Identifying Important Nodes with Bridgeness
H. Park, and J. Neville.
Neural Networks, 2023.
95. Ensemble learning for relational data.
H. Eldardiry, J. Neville, and R. Rossi.
Journal of Machine Learning Research, 21 (1), 1780-1816.

96. Scalable and exact sampling method for probabilistic generative graph models.
S. Moreno[†], J. Pfeiffer III[†], and J. Neville.
Data Mining and Knowledge Discovery, Volume 32, Issue 6, pp 1561-1596, 2018.
97. Designing Size Consistent Statistics for Accurate Anomaly Detection in Dynamic Networks.
T. La Fond[†], J. Neville, and B. Gallagher.
ACM Transactions on Knowledge Discovery from Data, 12:14, 2018.
98. Tied Kronecker Product Graph Models to Capture Variance in Network Populations.
S. Moreno, J. Neville, and S. Kirshner
ACM Transactions on Knowledge Discovery from Data, 12:3, 2018.
99. Graphlet Decomposition: Framework, Algorithms, and Applications.
N. Ahmed[†], J. Neville, R. Rossi[†], N. Duffield, T. Willke
Knowledge and Information Systems, Volume 50, Issue 3, pp 689-722, 2017.
100. Network Sampling: From Static to Streaming Graphs.
N. Ahmed[†], J. Neville, and R. Kompella
Transactions on Knowledge Discovery and Data Mining, Vol. 8, Issue 2, 2014.
101. Transforming Graph Data for Statistical Relational Learning.
R. Rossi[†], L. McDowell, D. Aha and J. Neville
Journal of Artificial Intelligence Research, Vol. 45, 363-441, 2012.
102. Correcting Evaluation Bias of Relational Classifiers with Network Cross Validation.
J. Neville, B. Gallagher, T. Eliassi-Rad, and T. Wang
Knowledge and Information Systems, 30-1, 31-55, 2012.
103. Gender demographics trends and changes in U.S. CS departments.
D. Baumann[†], S. Hambruch, and J. Neville.
Communications of the ACM, 54:11, 38-42, 2011.
104. Guided Data Repair
M. Yakout[†], A. Elmagarmid, J. Neville, M. Ouzzani, and I. Ilyas
Proceedings of the VLDB Endowment, 2011.
105. Prediction models for long-term Internet prefix availability
R. Khosla[†], S. Fahmy, Y. C. Hu, and J. Neville
Computer Networks, 2010.
106. A Bias-Variance Decomposition for Collective Inference Models.
J. Neville and D. Jensen.
Machine Learning Journal, 73:1, pages 87-106, 2008.
107. Dependency Networks for Relational Data.
J. Neville and D. Jensen.
Journal of Machine Learning Research, 8(Mar):653–692, 2007.
108. Exploiting Relational Structure to Understand Publication Patterns in High-Energy Physics.
A. McGovern, L. Friedland, M. Hay, B. Gallagher, A. Fast, J. Neville and D. Jensen.
SIGKDD Explorations, Volume 5, Issue 2, pages 165-172, 2003.

Referred Workshop Papers

109. Response-Aware User Memory Selection for LLM Personalization.
J Fisher, J Neville, CY Park.
Pluralistic Alignment Workshop, (ICML), 2026.

110. Pearl: Personalizing large language model writing assistants with generation-calibrated retrievers.
S Mysore, Z Lu, M Wan, L Yang, B Sarrafzadeh, S Menezes, T Baghaee, EB Gonzalez, J Neville, T Safavi.
Proceedings of the 1st Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual, (ACL), 2024.
111. Retrieval-Augmented Graph Explanation with LLM-Based Bayesian Inference.
J Zhang, J Liu, D Luo, J Neville, H Wei.
AI Agent for Information Retrieval Workshop, (CIKM), 2024.
112. Creating Generalizable Downstream Graph Models with Random Projections.
A. Amirov, C. Quirk, J. Neville.
Proceedings of the Workshop on Machine Learning on Graphs, (WSDM), 2023.
113. Lightweight Compositional Embeddings for Incremental Streaming Recommendation
M. Hang[†] T. Schnabel, L. Yang, and J. Neville.
Proceedings of the Workshop on Machine Learning on Graphs, (WSDM), 2022.
114. Social Reinforcement Learning
M. Goindani[†] and J. Neville
Proceedings of the AAAI Symposium on Challenges and Opportunities for Multi-Agent Reinforcement Learning, (COMARL), 4 pages, 2020.
115. Dynamic Network Modeling from Motif-Activity
G. Zeno[†], T. La Fond, and J. Neville
Proceedings of the 4th International Workshop on Mining Actionable Insights from Social Networks, WWW, 8 pages, 2020.
116. Joint Embedding Models for Textual and Social Analysis
C. Li[†], Y. Lai[†], D. Goldwasser, and J. Neville
Proceedings of the 1st Workshop on Deep Structured Prediction, (ICML), 5 pages, 2017.
117. Stochastic Gradient Descent for Relational Logistic Regression via Partial Network Crawls
J. Yang[†], B. Ribeiro and J. Neville
Proceedings of the 7th International Workshop on Statistical Relational AI, UAI, 7 pages, 2017.
118. Deep Dynamic Relational Classifiers: Exploiting Dynamic Neighborhoods in Complex Networks
H. Park[†], J. Moore[†], and J. Neville
Proceedings of the Mining Actionable Insights from Social Networks Workshop, WSDM, 7 pages, 2017.
119. Online Spike-and-slab Inference with Stochastic Expectation Propagation
S. Zhe[†], K. Lee, K. Zhang, and J. Neville
Proceedings of the 2016 Workshop on Advances in Approximate Bayesian Inference, NIPS, 2016.
120. Generating Local Explanations of Network Anomalies via Score Decomposition
T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the ODD 4.0: Outlier Definition, Detection, and Description on Demand, KDD, 6 pages, 2016.
121. Investigating the Impact of Graph Structure and Attribute Correlation on Collective Classification Performance
G. Zeno[†] and J. Neville
Proceedings of the 13th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2016.
122. Combining Gradient Boosting Machines with Collective Inference to Predict Continuous Values
I. Alodah[†] and J. Neville
Proceedings of the 6th International Workshop on Statistical Relational AI, IJCAI, 7 pages, 2016.
123. Better Together: Combining Language and Social Interactions into a Shared Representation
Y. Lai[†], C. Li[†], D. Goldwasser, and J. Neville
Proceedings of the TextGraphs 2016, (NAACL), 5 pages, 2016.

124. Analyzing the Transferability of Collective Inference Models Across Networks
R. Niu[‡], S. Moreno and J. Neville
Proceedings of the International Workshop on Information Analysis and Data Mining Over Social Network, (ICDM), 9 pages, 2015.
125. Using Bayesian Network Representations for Effective Sampling from Generative Network Models
P. Robles[†], S. Moreno, and J. Neville
Proceedings of the 5th International Workshop on Statistical Relational AI, UAI, 6 pages, 2015.
126. Assortativity in Chung Lu Random Graph Models
S. Mussmann[‡], J. Moore[‡], J. Pfeiffer III[†], and J. Neville
Proceedings of the 8th SNA-KDD Workshop, KDD, KDD, 8 pages, 2014.
127. Anomaly Detection in Networks with Changing Trends
T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the Outlier Detection & Description under Data Diversity Workshop, KDD, 10 pages, 2014.
128. Block Kronecker Product Graph Models
S. Moreno[†], P. Robles[†], and J. Neville
Proceedings of the 11th Workshop on Mining and Learning with Graphs, KDD, 6 pages, 2013.
129. Combining Active Sampling with Parameter Estimation and Prediction in Single Networks
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the Structured Learning: Inferring Graphs from Structured and Unstructured Inputs Workshop, ICML, 6 pages, 2013.
130. Space-Efficient Sampling from Social Activity Streams
N. Ahmed[†], J. Neville, and R. Kompella
Space-Efficient Sampling from Social Activity Streams. *Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining, KDD*, 8 pages, 2012.
131. Using Latent Communication Styles to Predict Individual Characteristics
J. Bates[‡], J. Neville, and J. Tyler
Proceedings of the 3rd Workshop on Social Media Analytics, KDD, 8 pages, 2012.
132. Active Sampling of Networks
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the 10th Workshop on Mining and Learning with Graphs, ICML, 8 pages, 2012.
133. On the Mismatch Between Learning and Inference for Single Network Domains
R. Xiang[†] and J. Neville
Proceedings of the Workshop on Inferning: Interactions between Inference and Learning, ICML, 6 pages, 2012.
134. Role-Dynamics: Fast Mining of Large Dynamic Networks
R. Rossi[†], Brian Gallagher, J. Neville, and Keith Henderson
Proceedings of the 1st International Workshop on Large Scale Network Analysis, WWW, 9 pages, 2012.
135. Understanding Propagation Error and Its Effect on Collective Classification
R. Xiang[†] and J. Neville
Proceedings of the 9th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2011.
136. Modeling the Variance of Network Populations with Mixed Kronecker Product Graph Models
S. Moreno[†], J. Neville, S. Kirshner and S.V.N. Vishwanathan. **Most Promising Paper Award**
Proceedings of the Workshop on Analyzing Networks and Learning with Graphs, 24th Annual Conference on Neural Information Processing Systems, 8 pages, 2010.
137. Reconsidering the Foundations of Network Sampling
N. Ahmed[†], J. Neville, and R. Kompella
Proceedings of the 2nd Workshop on Information in Networks, 5 pages, 2010.

138. Time-Based Sampling of Social Network Activity Graphs
N. Ahmed[†], F. Berchmans[†], J. Neville, and R. Kompella
Proceedings of the 8th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2010.
139. Multi-Network Fusion for Collective Inference
H. Eldardiry[†] and J. Neville
Proceedings of the 8th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2010.
140. Probabilistic Paths and Centrality in Time
J. Pfeiffer III[†] and J. Neville
Proceedings of the 4th SNA-KDD Workshop, KDD, 8 pages, 2010.
141. Combining Semi-supervised Learning and Relational Resampling for Active Learning in Network Domains
A. Kuwadekar[‡] and J. Neville. **Best Paper Award**
Proceedings of the Budgeted Learning Workshop, ICML, 8 pages, 2010.
142. Modeling the Evolution of Discussion Topics and Communication to Improve Relational Classification
R. Rossi[†] and J. Neville
Proceedings of the 1st Workshop on Social Media Analytics, KDD, 8 pages, 2010.
143. Ranking for Data Repairs
M. Yakout[†], A. Elmagarmid, and J. Neville
Proceedings of the 4th International Workshop on Ranking in Databases, ICDE 2010, 6 pages, 2010.
144. Modeling Relationship Strength in Online Social Networks.
R. Xiang[†] and J. Neville
Proceedings of the Workshop on Analyzing Networks and Learning with Graphs, 23rd Annual Conference on Neural Information Processing Systems, 8 pages, 2009.
145. An Investigation of the Distributional Characteristics of Generative Graph Models.
S. Moreno[†] and J. Neville.
Proceedings of the 1st Workshop on Information in Networks, 5 pages, 2009.
146. A Shrinkage Approach for Modeling Non-Stationary Relational Autocorrelation.
P. Angin[†] and J. Neville.
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 6 pages, 2008.
147. A Resampling Technique for Relational Data Graphs.
H. Eldardiry[†] and J. Neville.
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 6 pages, 2008.
148. Pseudolikelihood EM for Within-Network Relational Learning.
R. Xiang[†] and J. Neville
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 8 pages, 2008.
149. Exploiting Time-Varying Relationships in Statistical Relational Models.
U. Sharan[†] and J. Neville.
Proceedings of the 1st Social Network Analysis KDD Workshop, 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 7 pages, 2007.
150. Bias/Variance Analysis for Network Data.
J. Neville and D. Jensen.
Proceedings of the Workshop on Statistical Relational Learning, 23rd International Conference on Machine Learning, 8 pages, 2006.

151. Structure Learning for Statistical Relational Models.
J. Neville.
Proceedings of the 20th National Conference on Artificial Intelligence (Doctoral Consortium), pages 1656-1657, 2005.
152. Autocorrelation and Relational Learning: Challenges and Opportunities.
J. Neville, O. Simsek and D. Jensen.
Proceedings of the Workshop on Statistical Relational Learning, 21st International Conference on Machine Learning, 8 pages, 2004.
153. Collective Classification with Relational Dependency Networks. J. Neville and D. Jensen.
Proceedings of the 2nd Multi-Relational Data Mining Workshop, 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 77-91, 2003.
154. Statistical Relational Learning: Four Claims and a Survey.
J. Neville, M. Rattigan and D. Jensen.
Proceedings of the Workshop on Learning Statistical Models from Relational Data, 18th International Joint Conference on Artificial Intelligence, 5 pages, 2003.
155. Clustering Relational Data Using Attribute and Link Information.
J. Neville, M. Adler and D. Jensen.
Proceedings of the Text Mining and Link Analysis Workshop, 18th International Joint Conference on Artificial Intelligence, 6 pages, 2003.
156. Schemas and Models.
D. Jensen and J. Neville.
Proceedings of the Multi-Relational Data Mining Workshop, 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 15 pages, 2002.
157. Supporting Relational Knowledge Discovery: Lessons in Architecture and Algorithm Design.
J. Neville and D. Jensen.
Proceedings of the Data Mining Lessons Learned Workshop, 19th International Conference on Machine Learning, pages 57-64, 2002.
158. Correlation and Sampling in Relational Data Mining.
D. Jensen and J. Neville.
Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics, 14 pages, 2001.
159. Iterative Classification in Relational Data.
J. Neville and D. Jensen.
Proceedings of the Workshop on Learning Statistical Models from Relational Data, 17th National Conference on Artificial Intelligence, pages 42-49, 2000.

Book Chapters

160. Relational Dependency Networks.
J. Neville and D. Jensen.
Introduction to Statistical Relational Learning, L. Getoor and B. Taskar, editors, pages 239-268, 2007.

Invited Papers

161. Data Mining in Social Networks.
D. Jensen and J. Neville.
National Academy of Sciences Symposium on Dynamic Social Network Analysis, 13 pages, 2002.

Professional Service

Program Chair: AAAI 2023, SDM 2019, WSDM 2016

Elected Councilor, AAAI Executive Council (2015–2018)

ACM Doctoral Dissertation Award Committee (2019–2021)

Associate Chair or Area Chair for NeurIPS, ICML, KDD, AAAI, IJCAI, UAI, ICDM, and WSDM

Editorial boards: Machine Learning Journal; Data Mining and Knowledge Discovery

National Science Foundation: NSF Expeditions and IIS review panels

Treasurer, International Machine Learning Society (2009–2013)

Selected Sponsored Research²

Secured over \$6M in direct research funding and participated in collaborative research initiatives totaling more than \$60M. Funding sources include NSF, DARPA, IARPA, ARO, industry partners, and large multi-institution research centers.

Learning from Theorem Proving Search via Graph Representations

NSF/CISE/FMitF, co-Primary Investigator
\$250,000 (33% of total), 08/01/19 - 07/31/22

I²DS: Intelligent Interaction Defense System

DARPA/I2O, Primary Investigator
\$744,302 (33% of total), 09/01/18 - 08/31/22

Transfer Learning Within & Across Networks for Collective Classification

NSF/CISE/IIS, Primary Investigator
\$495,308 (100% of total), 07/01/16 - 06/30/19

Frontiers of Science of Information

NSF/Science & Technology Center, Senior Personnel
\$200,000 (0.8% of total), 08/01/15 - 07/31/20

Career: Machine Learning Methods and Statistical Analysis Tools for Single Network Domains

NSF/CISE/IIS, Primary Investigator
\$496,638 (100% of total), 01/01/12 - 12/31/16

Fusion and Analysis of Multi-Source Relational Data

DARPA/ISO, Primary Investigator
\$499,877 (100% of total), 06/23/08–06/22/10

Research Mentorship

Mentored and graduated 40 students, including 22 PhD students. Alumni hold research, engineering, and faculty positions at leading technology companies, national laboratories, and universities, including Google, Microsoft, Amazon, Meta, Apple, Intel, MIT Lincoln Laboratory, Lawrence Livermore National Laboratory, Virginia Tech, Notre Dame, and the University of Utah.

²Reported amounts reflect funds 100% under my responsibility.