Ankit Agrawal

Current Position

Research Professor

Department of Electrical and Computer Engineering
McCormick School of Engineering and Applied Science
Northwestern University

Office: M384, TECH
Technological Institute
2145 Sheridan Road
Evanston, Illinois 60208-3118

Ph: 847-491-8163
Fax: 847-491-4455
Email: ankit-agrawal@northwestern.edu

Website: http://users.eecs.northwestern.edu/~ankitag
CUCIS: http://cucis.ece.northwestern.edu/members/ankitag
Google Scholar: http://scholar.google.com/citations?user=5e7JslgAAAAJ
DBLP: https://dblp.org/pid/55/5176-1.html
Scopus: https://www.scopus.com/authid/detail.uri?authorId=24604939500
ORCID: https://orcid.org/0000-0002-5519-0302
NCBI: https://www.ncbi.nlm.nih.gov/myncbi/ankit.agrawal.3/bibliography/public/
ScholarGPS: https://scholargps.com/scholars/45723753753118/ankit-agrawal
Loop: https://loop.frontiersin.org/people/2137579
LinkedIn: https://www.linkedin.com/in/ankit-agrawal-iitr-isu-nu/

Research Interests

Artificial Intelligence, High Performance Data Mining, Materials Informatics, Healthcare Informatics, Social Media Analytics, Bioinformatics

Research Output Summary

- 200+ peer-reviewed publications (100+ as first/last author)
- 15000+ citations (h-index: 50+)
- 20+ software
- 75+ invited/keynote talks
- 20+ research projects as PI/Co-PI
- 1 patent granted, 2 pending

Education

- Ph.D in Computer Science (with Minor in Bioinformatics and Computational Biology), Iowa State University, USA. 2006 - 2009.
  - GPA: 4.0/4.0
  - Dissertation: Sequence-Specific Sequence Comparison Using Pairwise Statistical Significance
  - Advisor: Dr. Xiaoqiu Huang
- B.Tech. in Computer Science and Engineering, Indian Institute of Technology, Roorkee, INDIA. 2002 - 2006.
  - GPA: 9.43/10.0

Appointments

- *Research Professor*, Northwestern University, 2020 - present
- *Honorary Professor*, Amity University, India, 2020 - present
- *Research Associate Professor*, Northwestern University, 2013 - 2020
- *Research Assistant Professor*, Northwestern University, 2010 - 2013
- *Postdoctoral Fellow*, Northwestern University, 2009 - 2010

- *Research Assistant*, Iowa State University, 2007 - 2009
- *Graduate Assistant*, Iowa State University, 2006 - 2007

Research Grants

1. **Senior Personnel**, "Center for Hierarchical Materials Design (CHiMaD): Phase III", *National Institute of Standards and Technology (NIST)*, $700,000, 2025-2025. (PI: Peter Voorhees) [70NANB24H136]
2. **PI**, "AI-Driven Nanocombinatorics for Accelerated Structural Characterization: Automated High-Throughput Nanoparticle Library Screening and Analytics (Renewal)", *Center for Nanocombinatorics, Northwestern University*, $100,000, 2024-2025.
3. **Co-PI**, "EAGER: XAISE: Explainable Artificial Intelligence for Science and Engineering", *National Science Foundation (NSF)*, $300,000, 2023-2025. (PI: Alok Choudhary) [OAC-2331329]
4. **PI**, "AI-Driven Nanocombinatorics for Accelerated Structural Characterization: Automated High-Throughput Nanoparticle Library Screening and Analytics (Renewal)", *Center for Nanocombinatorics, Northwestern University*, $100,000, 2023-2024.
5. **Co-PI**, "Nitrogen Activation at Catalyst Surfaces to Catalyze Net-Zero: An AI-Driven Approach", *Center for Nanocombinatorics, Northwestern University*, $100,000, 2023-2024.
6. **PI**, "AI-Driven Nanocombinatorics for Accelerated Structural Characterization: Automated High-Throughput Nanoparticle Library Screening and Analytics", *Center for Nanocombinatorics, Northwestern University*, $100,000, 2021-2022.
7. **PI**, "AI-Driven Nanocombinatorics for Functional Characterization and Optimization: Predictive Modeling and Active Learning of Catalysis in Megalibraries", *Center for Nanocombinatorics, Northwestern University*, $90,000, 2021-2022.
8. **PI**, "Collaborative Research: AI-Driven Multi-Scale Design of Materials under Processing Constraints", *National Science Foundation (NSF)*, $379,022 (Total $651,462), 2021-2025. (Lead PI: Pinar Acar) [CMMI-2053929]
9. **Co-PI**, "RAPIDS2: A SciDAC Institute for Computer Science, Data, and Artificial Intelligence", *Department of Energy (DOE)*, $650,000 (Total $28,750,000), 2020-2025. (PI: Wei-keng Liao; Lead PI: Rob Ross) [DE-SC0021399]
10. **PI**, "Center for Hierarchical Materials Design (CHiMaD): Phase II" (Agrawal Subproject), *National Institute of Standards and Technology (NIST)*, $953,221 (Total $25,000,000), 2019-2023. (Lead PI: Peter Voorhees) [70NANB19H005]
11. **PI**, "Data-driven Analytics for Understanding Materials Properties", *Toyota Motor Corporation*, $300,000, 2019-2020.
12. **PI**, "Digital Innovation Design (DID)", *Defense Logistics Agency via Steel Founders Society of America*, $200,000, 2019-2020. (Original PI: Greg Olson)
13. **Co-PI**, "PROTEUS: Machine Learning Driven Resilience for Extreme-scale Systems", *Department of Energy (DOE)*, $1,248,115, 2018-2023. (PI: Alok Choudhary) [DE-SC0019358]
14. **PI**, "The investigation of machine learning for material development", *Toyota Motor Corporation*, $200,000, 2017-2018.
15. **Senior Personnel**, "BD Spokes: SPOKE: MIDWEST: Collaborative: Integrative Materials Design (IMaD): Leverage, Innovate, & Disseminate", *National Science Foundation (NSF)*, $123,847 (Total $989,700), 2017-2022. (PI: Peter Voorhees; Lead PI: Ian Foster) [IIS-1636909]
16. **Co-PI**, "Scalable, In-situ Clustering and Data Analysis for Extreme Scale Scientific Computing", *Department of Energy (DOE)*, $1,219,899, 2015-2021. (PI: Alok Choudhary) [DE-SC0014330]
17. **Co-PI**, "SHF:Medium:Collaborative Research: Scalable Algorithms for Spatio-temporal Data Analysis", *National Science Foundation (NSF)*, $709,342 (Total $934,342), 2014-2019. (PI: Alok Choudhary) [CCF-1409601]

18. **PI**, "Advanced Materials Center for Excellence: Center for Hierarchical Materials Design (CHiMaD)" (Agrawal Subproject), *National Institute of Standards and Technology (NIST)*, $505,358 (Total $25,000,000), 2014-2018. (Lead PI: Peter Voorhees) [70NANB14H012]
19. **PI**, "Social Media mining of caregiver experiences: Opportunity for preventing caregiver burnout", *Northwestern Data Science Initiative*, $25,000, 2017-2017.
20. **PI**, "Data-driven analytics for understanding processing-structure-property-performance relationships in steel alloys", *Northwestern Data Science Initiative*, $45,000, 2016-2017.
21. **Co-PI**, "Scaling up the screening of molecular networks in the rational design of optically active materials", *Northwestern Data Science Initiative*, $9,000 (Total $45,000), 2016-2017. (PI: Kevin Kohlstedt)
22. **PI**, "Analyzing caregiving experience on Twitter", *Feinberg School of Medicine*, $14,246, 2015-2016.
23. **Co-PI**, "SIMPLEX: Data-driven Discovery of Novel Thermoelectric Materials", *Defense Advanced Research Projects Agency (DARPA)*, $601,250 (Total $1,559,999), 2015-2018. (PI: Alok Choudhary; Lead PI: Greg Olson) [N66001-15-C-4036]
24. **Co-PI**, "EAGER: Scalable Big Data Analytics", *National Science Foundation (NSF)*, $300,000, 2013-2016. (PI: Alok Choudhary) [IIS-1343639]
25. **Co-PI**, "MURI: MANAGING THE MOSAIC OF MICROSTRUCTURE: Image analysis, data structures, mathematical theory of microstructure, and hardware for the structure-property relationship", *Air Force Office of Scientific Research (AFOSR), Department of Defense (DOD)*, $750,000 (Total $5,658,616), 2012-2018. (PI: Alok Choudhary; Lead PI: Marc De Graef) [FA9550-12-1-0458]
26. **Co-PI**, "Scalable Data Management, Analysis, and Visualization (SDAV) Institute", *Department of Energy (DOE)*, $750,000 (Total $25,000,000), 2012-2019. (PI: Alok Choudhary; Lead PI: Arie Shoshani) [DE-SC0007456]
27. **Senior Researcher**, "Expeditions in Computing: Understanding Climate Change: A Data Driven Approach", *National Science Foundation (NSF)*, $900,000 (Total $10,000,000), 2010-2016. (PI: Alok Choudhary; Lead PI: Vipin Kumar) [CCF-1029166]
28. **Co-PI**, "EAGER: Discovering Knowledge from Scientific Research Networks", *National Science Foundation (NSF)*, $256,000, 2011-2014. (PI: Alok Choudhary) [ACI-1144061]
29. **Research Participant**, "Scalable and Power Efficient Data Analytics for Hybrid Exascale Systems", *Department of Energy (DOE)*, $705,000, 2010-2014. (PI: Alok Choudhary) [DE-SC0005340]

Honors and Awards

- Featured in Stanford/Elsevier's list of top 2% scientists worldwide, 2024, 2023, 2022, 2021, 2020
- Named a Top Scholar by ScholarGPS for being in top 0.5% of scholars worldwide in the fields of machine learning, deep learning, and informatics, 2024
- Hind Rattan Award, 2023 (translated as "Jewel of India", one of the highest Indian diasporic awards conferred to NRIs on the eve of Republic Day, for outstanding services, achievements, and contributions for keeping the "Flag of India High")
- Honorary Professor, Amity School of Engineering and Technology (ASET), Amity University, 2020-present
- Best Paper Finalist, HiPC 2017
- Northwestern Data Science Initiative Award, 2017
- Northwestern Data Science Initiative Award, 2016
- Best Paper Award at IEEE Cluster 2016
- Gordon Betty Moore Foundation (GBMF) Data-Driven-Discovery Competition Finalist, 2014
- Outstanding Paper Award awarded by Tata Consultancy Services for joint work published in IMMI, 2014

- Best Paper Award at KDD Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine), 2013
- Design Contest Winner at 25th International Conference on VLSI Design, 2012
- Research Excellence Award for Summer 2009 at ISU (to recognize up to 10% graduating students for outstanding research accomplishments)
- Graduate and Professional Student Senate Peer Research Award for Spring 2009 at ISU (based on research conducted at ISU)
- Best student paper award at International Conference on Information Technology, ICIT 2008
- Institute Silver Medal for obtaining the highest C.G.P.A. (9.43/10) amongst the students graduating with B.Tech. (Computer Science and Engineering) at IIT Roorkee, 2006
- Institute Silver Medal for Best Project in B.Tech. (Computer Science and Engineering) at IIT Roorkee, 2006
- Offered branch change at the end of B.Tech. first year at IIT Roorkee for being in top 1% of students, 2003

Teaching

- Co-instructor: CE/CS 395/495/396/496: AI for Science and Business, Spring 2025.
- Co-instructor: CE/CS 395/495: AI for Science, Spring 2024.
- Co-instructor: TMS Online Course on Artificial Intelligence in Materials Science and Engineering, April 03-05, 2023, November 02-04, 2021.
- Co-instructor: PSED 510-1: Predictive Science & Engineering Design Cluster Seminar, Fall 2024, Fall 2020, Fall 2017, Northwestern University.
- Co-instructor: CE/EECS 510: Social Media Mining, Spring 2024, Spring 2022, Spring 2021, Spring 2020, Spring 2019, Spring 2018, Spring 2017, Spring 2016, Spring2015, Spring 2014, Northwestern University.
- Invited Speaker: BiGmax Summer School on Harnessing Big Data in Materials Science from Theory to Experiment, September 15, 2021, Max Plank Institute, Germany (Virtual).
- Co-instructor: EECS 100: Electrons, Photons, and Bits: Adventures in Electrical and Computer Engineering, Spring 2019, Spring 2018, Spring 2017, Spring 2016, Spring 2015, Spring 2011, Northwestern University.
- Co-instructor: MSIT 423: Data Mining and Business Intelligence, Spring 2011, Northwestern University.
- Teaching Assistant: ComS/BCB 567: Bioinformatics I - Fundamentals of Genome Informatics, Fall 2008, Iowa State University.
- Teaching Assistant: ComS 309: Software Engineering, Fall 2006, Iowa State University.

Patents

Granted

1. Ankit Agrawal and Alok Choudhary, "System and method for predicting fatigue strength of alloys." U.S. Patent No. 10,830,747. 10 Nov. 2020.

Pending

1. Alexandra Lauren Day, Carolin Barbara Wahl, Roberto M. dos Reis, Alok Nidhi Choudhary, Chad A. Mirkin, Vinayak P. Dravid and Ankit Agrawal, "Rapid image segmentation pipeline for scanning transmission electron microscopy" US Application No.: 19/207,977, 2025.
2. Alexandra Lauren Day, Carolin Barbara Wahl, Roberto M. dos Reis, Wei-keng Liao, Chad A. Mirkin, Vinayak P. Dravid, Alok Nidhi Choudhary, and Ankit Agrawal, "Nanoparticle Image Processing Pipeline for Structure Characterization" US Application No.: 63/710,371, 2024 (Provisional).

Publications

Book Chapters

1. **A. Agrawal** and A. Choudhary, "Artificial Intelligence for Accelerating Materials Discovery," in Artificial Intelligence for Science: A Deep Learning Revolution, A. Choudhary, T. Hey, and G. Fox, Eds. World Scientific, 2023, pp. 431–443.

2. **A. Agrawal**, K. Gopalakrishnan, and A. Choudhary, "Materials Image Informatics Using Deep Learning," in Handbook on Big Data and Machine Learning in the Physical Sciences, vol. 1: Big Data Methods in Experimental Materials Discovery, World Scientific, 2020, pp. 205–230.
3. D. Han, **A. Agrawal**, W. Liao, and A. Choudhary, "A Fast DBSCAN Algorithm with Spark Implementation," in Big Data in Engineering Applications, S. S. R. et al., Ed. Springer Nature Singapore, 2018, pp. 173–192.
4. **A. Agrawal** and A. Choudhary, "Health Services Data: Big Data Analytics for Deriving Predictive Healthcare Insights," in Data and Measures in Health Services Research, B. Sobolev, A. Levy, and S. Goring, Eds. Springer US, 2016, pp. 1–17.
5. **A. Agrawal**, M. Patwary, W. Hendrix, W. Liao, and A. Choudhary, "High performance big data clustering," in Advances in Parallel Computing, Volume 23: Cloud Computing and Big Data, L. Grandinetti, Ed. IOS Press, 2013, pp. 192–211.
6. **A. Agrawal**, A. Choudhary, and X. Huang, "Sequence-Specific Sequence Comparison Using Pairwise Statistical Significance," in Software Tools and Algorithms for Biological Systems, vol. 696, H. R. Arabnia, Ed. Springer, 2011, pp. 297–306.

Journal Publications

7. A. L. Day, C. B. Wahl, R. D. Reis, W. Liao, Y. Li, M. N. T. Kilic, C. A. Mirkin, V. P. Dravid, A. Choudhary, and **A. Agrawal**, "Automated image segmentation for accelerated nanoparticle characterization," Scientific Reports, vol. 15, no. 1, pp. 1–13, 2025.
8. S. Keshavarz, Y. Mao, A. C. E. Reid, and **A. Agrawal**, "Advancing material simulations: Physics-Informed Neural Networks and Object-Oriented Crystal Plasticity Finite Element Methods," International Journal of Plasticity, vol. 185, p. 104221, 2025.
9. Y. Li, V. Gupta, M. N. T. Kilic, K. Choudhary, D. Wines, W. Liao, A. Choudhary, and **A. Agrawal**, "Hybrid-LLM-GNN: integrating large language models and graph neural networks for enhanced materials property prediction," Digital Discovery, vol. 4, pp. 376–383, 2025.
10. Y. Mao, S. Keshavarz, M. N. T. Kilic, K. Wang, Y. Li, A. C. E. Reid, W. Liao, A. Choudhary, and **A. Agrawal**, "A deep learning-based crystal plasticity finite element model," Scripta Materialia, vol. 254, p. 116315, 2025.
11. K. Choudhary, D. Wines, K. Li, K. F. Garrity, V. Gupta, A. H. Romero, J. T. Krogel, K. Saritas, A. Fuhr, P. Ganesh, P. R. C. Kent, K. Yan, Y. Lin, S. Ji, B. Blaiszik, P. Reiser, P. Friederich, **A. Agrawal**, P. Tiwary, E. Beyerle, P. Minch, T. D. Rhone, I. Takeuchi, R. B. Wexler, A. Mannodi-Kanakkithodi, E. Ertekin, A. Mishra, N. Mathew, M. Wood, A. D. Rohskopf, J. Hattrick-Simpers, S.-H. Wang, L. E. K. Achenie, H. Xin, M. Williams, A. J. Biacchi, and F. Tavazza, "JARVIS-Leaderboard: A large scale benchmark of materials design methods," npj Computational Materials, vol. 10, no. 1, p. 93, 2024.
12. A. L. Day, C. B. Wahl, V. Gupta, R. D. Reis, W. Liao, C. A. Mirkin, V. P. Dravid, A. Choudhary, and **A. Agrawal**, "Machine Learning-Enabled Image Classification for Automated Electron Microscopy," Microscopy and Microanalysis, p. ozae042, 2024.
13. A. L. Day, C. B. Wahl, R. dos Reis, W. Liao, Y. Li, M. N. T. Kilic, C. A. Mirkin, V. P. Dravid, A. Choudhary, and **A. Agrawal**, "Rapid Image Segmentation Pipeline to Support Multimodal STEM Acquisition," Microscopy and Microanalysis, vol. 30, no. Supplement 1, pp. 442–443, 2024.
14. V. Gupta, K. Choudhary, B. DeCost, F. Tavazza, C. Campbell, W. Liao, A. Choudhary, and **A. Agrawal**, "Structure-aware graph neural network based deep transfer learning framework for enhanced predictive analytics on diverse materials datasets," npj Computational Materials, vol. 10, p. 1, 2024.

15. V. Gupta, Y. Li, A. Peltekian, M. N. T. Kilic, W. Liao, A. Choudhary, and **A. Agrawal**, "Simultaneously improving accuracy and computational cost under parametric constraints in materials property prediction tasks," Journal of Cheminformatics, vol. 16, no. 1, p. 17, 2024.
16. C. Lee, V. Hewes, G. Cerati, K. Wang, A. Aurisano, **A. Agrawal**, A. Choudhary, and W. Liao, "Addressing GPU Memory Limitations for Graph Neural Networks in High-Energy Physics Applications," Frontiers in High Performance Computing, vol. 2, p. 1458674, 2024.
17. K. Wang, V. Gupta, C. S. Lee, Y. Mao, M. N. T. Kilic, Y. Li, Z. Huang, W. Liao, A. Choudhary, and **A. Agrawal**, "XElemNet: Towards explainable AI for deep neural networks in materials science," Scientific Reports, vol. 14, p. 25178, 2024.
18. A. Yan, M. N. T. Kilic, **A. Agrawal**, R. dos Reis, and V. Dravid, "Neural Network Models Towards Space Group Determination Using Dynamically Simulated EBSD and TKD Patterns," Microscopy and Microanalysis, vol. 30, no. Supplement 1, pp. 388–390, 2024.
19. V. Gupta, K. Choudhary, Y. Mao, K. Wang, F. Tavazza, C. Campbell, W. Liao, A. Choudhary, and **A. Agrawal**, "MPpredictor: An Artificial Intelligence-Driven Web Tool for Composition-Based Material Property Prediction," Journal of Chemical Information and Modeling, vol. 63, no. 7, pp. 1865–1871, 2023.
20. K. Gumpula, N. Koloskov, D. Grzenda, V. Hewes, A. Aurisano, G. Cerati, A. Day, J. Kowalkowski, C. Lee, K. Wang, W. Liao, M. Spiropulu, **A. Agrawal**, J. Vlimant, L. Gray, T. Klijnsma, P. Calafiura, S. Conlon, S. Farrell, X. Ju, and D. Murnane, "Graph Neural Network for Object Reconstruction in Liquid Argon Time Projection Chambers," Journal of Physics: Conference Series, vol. 2438, p. 012091, 2023.
21. V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Evolution of artificial intelligence for application in contemporary materials science," MRS Communications, pp. 1–10, 2023.
22. V. Gupta, A. Peltekian, W. Liao, A. Choudhary, and **A. Agrawal**, "Improving deep learning model performance under parametric constraints for materials informatics applications," Scientific Reports, vol. 13, no. 1, p. 9128, 2023.
23. Y. Mao, M. Hasan, A. Paul, V. Gupta, K. Choudhary, F. Tavazza, W. Liao, A. Choudhary, P. Acar, and **A. Agrawal**, "An AI-driven microstructure optimization framework for elastic properties of titanium beyond cubic crystal systems," npj Computational Materials, vol. 9, p. 111, 2023.
24. Y. Mao, H. Lin, C. X. Yu, R. Frye, D. Beckett, K. Anderson, L. Jacquemetton, F. Carter, Z. Gao, W. Liao, A. N. Choudhary, K. Ehmann, and **A. Agrawal**, "A deep learning framework for layer-wise porosity prediction in metal powder bed fusion using thermal signatures," Journal of Intelligent Manufacturing, vol. 24, pp. 315–329, 2023.
25. C. Wahl, A. Day, V. Gupta, R. R. dos Reis, W. Liao, C. Mirkin, A. Choudhary, V. P. Dravid, and **A. Agrawal**, "Machine Learning Enabled Image Classification for Automated Data Acquisition in the Electron Microscope," Microscopy and Microanalysis, vol. 29, pp. 1909–1910, 2023.
26. K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, **A. Agrawal**, S. J. L. Billinge, E. Holm, S. P. Ong, and C. Wolverton, "Recent advances and applications of deep learning methods in materials science," npj Computational Materials, vol. 8, p. 59, 2022.
27. M. Hasan, Y. Mao, K. Choudhary, F. Tavazza, A. Choudhary, **A. Agrawal**, and P. Acar, "Data-Driven Multi-Scale Modeling and Optimization for Elastic Properties of Cubic Microstructures," Integrating Materials and Manufacturing Innovation, vol. 11, pp. 230–240, 2022.
28. D. Jha, V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Moving closer to experimental level materials property prediction using AI," Scientific Reports, vol. 12, p. 11953, 2022.

29. S. Lee, K.-yuan Hou, K. Wang, S. Sehrish, M. Paterno, J. Kowalkowski, Q. Koziol, R. B. Ross, **A. Agrawal**, A. Choudhary, and W. Liao, "A case study on parallel HDF5 dataset concatenation for high energy physics data analysis," Parallel Computing, vol. 110, p. 102877, 2022.

30. S. Lee, Q. Kang, R. Al-Bahrani, **A. Agrawal**, A. Choudhary, and W. Liao, "Improving scalability of parallel CNN training by adaptively adjusting parameter update frequency," Journal of Parallel and Distributed Computing, vol. 159, pp. 10–23, 2022.

31. Y. Mao, Z. Yang, D. Jha, A. Paul, W. Liao, A. Choudhary, and **A. Agrawal**, "Generative Adversarial Networks and Mixture Density Networks-Based Inverse Modeling for Microstructural Materials Design," Integrating Materials and Manufacturing Innovation, vol. 11, pp. 637–647, 2022.

32. V. Gupta, K. Choudhary, F. Tavazza, C. Campbell, W. Liao, A. Choudhary, and **A. Agrawal**, "Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data," Nature Communications, vol. 12, no. 6595, 2021.

33. D. Jha, V. Gupta, L. Ward, Z. Yang, C. Wolverton, I. Foster, W. Liao, A. N. Choudhary, and **A. Agrawal**, "Enabling Deeper Learning on Big Data for Materials Informatics Applications," Scientific Reports, vol. 11, no. 4244, 2021.

34. A. Paul, W. Liao, A. Choudhary, and **A. Agrawal**, "Harnessing Psycho-lingual and Crowd-Sourced Dictionaries for Predicting Taboos in Written Emotional Disclosure in Anonymous Confession Boards," Journal of Healthcare Informatics Research, 2021.

35. K. Choudhary, K. F. Garrity, A. C. E. Reid, B. DeCost, A. J. Biacchi, A. R. H. Walker, Z. Trautt, J. Hattrick-Simpers, A. G. Kusne, A. Centrone, A. Davydov, J. Jiang, R. Pachter, G. Cheon, E. Reed, **A. Agrawal**, X. Qian, V. Sharma, H. Zhuang, S. V. Kalinin, B. G. Sumpter, G. Pilania, P. Acar, S. Mandal, K. Haule, D. Vanderbilt, K. Rabe, and F. Tavazza, "The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design," npj Computational Materials, vol. 6, no. 173, 2020.

36. Q. Kang, S. Lee, K. Hou, R. Ross, **A. Agrawal**, A. Choudhary, and W. Liao, "Improving MPI Collective I/O for High Volume Non-contiguous Requests With Intra-node Aggregation," IEEE Transactions on Parallel and Distributed Systems, vol. 31, no. 11, pp. 2682–2695, 2020.

37. Z. Yang, S. Papanikolaou, A. C. E. Reid, W. Liao, A. N. Choudhary, C. Campbell, and **A. Agrawal**, "Learning to Predict Crystal Plasticity at the Nanoscale: Deep Residual Networks and Size Effects in Uniaxial Compression Discrete Dislocation Simulations," Scientific Reports, vol. 10, no. 8262, 2020.

38. **A. Agrawal** and A. Choudhary, "Deep materials informatics: Applications of deep learning in materials science," MRS Communications, vol. 9, no. 3, pp. 779–792, 2019.

39. D. Jha, K. Choudhary, F. Tavazza, W. Liao, A. Choudhary, C. Campbell, and **A. Agrawal**, "Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning," Nature Communications, vol. 10, no. 5316, 2019.

40. Q. Kang, J. L. Träff, R. Al-Bahrani, **A. Agrawal**, A. Choudhary, and W. Liao, "Scalable Algorithms for MPI Intergroup Allgather and Allgatherv," Parallel Computing, vol. 85, pp. 220–230, 2019.

41. A. Paul, P. Acar, W. Liao, A. N. Choudhary, V. Sundararaghavan, and **A. Agrawal**, "Microstructure optimization with constrained design objectives using machine learning-based feedback-aware data-generation," Computational Materials Science, vol. 160, pp. 334–351, 2019.

42. A. Paul, A. Furmanchuk, W. Liao, A. Choudhary, and **A. Agrawal**, "Property Prediction of Organic Donor Molecules for Photovoltaic Applications using Extremely Randomized Trees," Molecular Informatics, vol. 38, p. 1900038, 2019.

43. Z. Yang, Y. C. Yabansu, D. Jha, W. Liao, A. N. Choudhary, S. R. Kalidindi, and **A. Agrawal**, "Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches," Acta Materialia, vol. 166, pp. 335–345, 2019.

44. **A. Agrawal** and A. Choudhary, "An online tool for predicting fatigue strength of steel alloys based on ensemble data mining," International Journal of Fatigue, vol. 113, pp. 389–400, 2018.

45. M. K. Danilovich, J. Tsay, R. Al-Bahrani, A. Choudhary, and **A. Agrawal**, "#Alzheimer's and Dementia: Expressions of Memory Loss on Twitter," Topics in Geriatric Rehabilitation, vol. 34, pp. 48–53, 2018.

46. A. Furmanchuk, J. E. Saal, J. W. Doak, G. B. Olson, A. Choudhary, and **A. Agrawal**, "Prediction of seebeck coefficient for compounds without restriction to fixed stoichiometry: A machine learning approach," Journal of Computational Chemistry, vol. 39, no. 4, pp. 191–202, 2018.

47. K. Gopalakrishnan, H. Gholami, A. Vidyadharan, A. Choudhary, and **A. Agrawal**, "Crack Damage Detection in Unmanned Aerial Vehicle Images of Civil Infrastructure Using Pre-trained Deep Learning Model," International Journal for Traffic and Transport Engineering, vol. 8, p. 1, 2018.

48. D. Jha, S. Singh, R. Al-Bahrani, W. Liao, A. N. Choudhary, M. D. Graef, and **A. Agrawal**, "Extracting Grain Orientations from EBSD Patterns of Polycrystalline Materials using Convolutional Neural Networks," Microscopy and Microanalysis, vol. 24, no. 5, pp. 497–502, 2018.

49. D. Jha, L. Ward, A. Paul, W. Liao, A. Choudhary, C. Wolverton, and **A. Agrawal**, "ElemNet: Deep Learning the Chemistry of Materials From Only Elemental Composition," Scientific Reports, vol. 8, no. 17593, 2018.

50. K. Kim, L. Ward, J. He, A. Krishna, **A. Agrawal**, P. Voorhees, and C. Wolverton, "Machine-learning-accelerated high-throughput materials screening: Discovery of novel quaternary Heusler compounds," Physical Review Materials, vol. 2, no. 123801, 2018.

51. M. Mozaffar, A. Paul, R. Al-Bahrani, S. Wolff, A. Choudhary, **A. Agrawal**, K. Ehmann, and J. Cao, "Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks," Manufacturing Letters, vol. 18, pp. 35–39, 2018.

52. A. Paul, P. Acar, R. Liu, W. Liao, A. Choudhary, V. Sundararaghavan, and **A. Agrawal**, "Data Sampling Schemes for Microstructure Design with Vibrational Tuning Constraints," American Institute of Aeronautics and Astronautics (AIAA) Journal, vol. 56, no. 3, pp. 1239–1250, 2018.

53. Z. Yang, X. Li, L. C. Brinson, A. Choudhary, W. Chen, and **A. Agrawal**, "Microstructural Materials Design via Deep Adversarial Learning Methodology," Journal of Mechanical Design, vol. 140, no. 11, p. 10, 2018.

54. Z. Yang, Y. C. Yabansu, R. Al-Bahrani, W. Liao, A. N. Choudhary, S. R. Kalidindi, and **A. Agrawal**, "Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets," Computational Materials Science, vol. 151, pp. 278–287, 2018.

55. R. Al-Bahrani, **A. Agrawal**, and A. Choudhary, "Survivability prediction of colon cancer patients using neural networks," Health Informatics Journal, p. 1460458217720395, 2017.

56. Y. Cheng, **A. Agrawal**, H. Liu, and A. Choudhary, "Legislative Prediction with Dual Uncertainty Minimization from Heterogeneous Information," Statistical Analysis and Data Mining: The ASA Data Science Journal, vol. 10, no. 2, pp. 110–120, 2017.

57. K. Gopalakrishnan, A. Choudhary, and **A. Agrawal**, "Big Data in Building Information Modeling Research: Survey and Exploratory Text Mining," MOJ Civil Eng, vol. 3, no. 6, p. 00087, 2017.

58. K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and **A. Agrawal**, "Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection," Construction and Building Materials, vol. 157, pp. 322–330, 2017.
59. A. G. Gagorik, B. Savoie, N. Jackson, **A. Agrawal**, A. Choudhary, M. A. Ratner, G. C. Schatz, and K. L. Kohlstedt, "Improved Scaling of Molecular Network Calculations: The Emergence of Molecular Domains," The Journal of Physical Chemistry Letters, vol. 8, no. 2, pp. 415–421, 2017.
60. R. Liu, Y. C. Yabansu, Z. Yang, A. N. Choudhary, S. R. Kalidindi, and **A. Agrawal**, "Context Aware Machine Learning Approaches for Modeling Elastic Localization in Three-Dimensional Composite Microstructures," Integrating Materials and Manufacturing Innovation, vol. 6, no. 2, pp. 160–171, 2017.
61. L. Ward, R. Liu, A. Krishna, V. I. Hegde, **A. Agrawal**, A. Choudhary, and C. Wolverton, "Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations," Physical Review B, vol. 96, no. 2, p. 024104, 2017.
62. Y. Xie, Z. Chen, D. Palsetia, G. Trajcevski, **A. Agrawal**, and A. Choudhary, "Silverback+: Scalable Association Mining Via Fast List Intersection For Columnar Social Data," Knowledge and Information Systems (KAIS), vol. 50, no. 3, pp. 969–997, 2017.
63. **A. Agrawal** and A. Choudhary, "Perspective: Materials informatics and big data: Realization of the 'fourth paradigm' of science in materials science," APL Materials, vol. 4, no. 053208, pp. 1–10, 2016.
64. A. Furmanchuk, **A. Agrawal**, and A. Choudhary, "Predictive analytics for crystalline materials: Bulk modulus," RSC Advances, vol. 6, no. 97, pp. 95246–95251, 2016.
65. E. Rangel, W. Hendrix, **A. Agrawal**, W. Liao, and A. Choudhary, "AGORAS: A Fast Algorithm for Estimating Medoids in Large Datasets," Procedia Computer Science, vol. 80, pp. 1159–1169, 2016.
66. L. Ward, **A. Agrawal**, A. Choudhary, and C. Wolverton, "A General-Purpose Machine Learning Framework for Predicting Properties of Inorganic Materials," npj Computational Materials, vol. 2, no. 16028, 2016.
67. R. Liu, A. Kumar, Z. Chen, **A. Agrawal**, V. Sundararaghavan, and A. Choudhary, "A Predictive Machine Learning Approach for Microstructure Optimization and Materials Design," Scientific Reports, vol. 5, no. 11551, 2015.
68. R. Liu, Y. C. Yabansu, **A. Agrawal**, S. R. Kalidindi, and A. N. Choudhary, "Machine learning approaches for elastic localization linkages in high-contrast composite materials," Integrating Materials and Manufacturing Innovation, vol. 4, no. 13, pp. 1–17, 2015.
69. **A. Agrawal**, P. D. Deshpande, A. Cecen, G. P. Basavarsu, A. N. Choudhary, and S. R. Kalidindi, "Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters," Integrating Materials and Manufacturing Innovation, vol. 3, no. 8, pp. 1–19, 2014.
70. A. R. Ganguly, E. Kodra, **A. Agrawal**, A. Banerjee, S. Boriah, S. Chatterjee, S. Chatterjee, A. Choudhary, D. Das, J. Faghmous, P. Ganguli, S. Ghosh, K. Hayhoe, C. Hays, W. Hendrix, Q. Fu, J. Kawale, D. Kumar, V. Kumar, W. Liao, S. Liess, R. Mawalagedara, V. Mithal, R. Oglesby, K. Salvi, P. K. Snyder, K. Steinhaeuser, D. Wang, and D. Wuebbles, "Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques," Nonlinear Processes in Geophysics, vol. 21, pp. 777–795, 2014.
71. B. Meredig, **A. Agrawal**, S. Kirklin, J. E. Saal, J. W. Doak, A. Thompson, K. Zhang, A. Choudhary, and C. Wolverton, "Combinatorial screening for new materials in unconstrained composition space with machine learning," Physical Review B, vol. 89, no. 094104, pp. 1–7, 2014. **BM and AA are co-first authors**.

72. D. Palsetia, M. M. A. Patwary, **A. Agrawal**, and A. Choudhary, "Excavating Social Circles via User-Interests," Social Network Analysis and Mining, vol. 4, no. 1, pp. 1–12, 2014.
73. S. W. Son, Z. Chen, W. Hendrix, **A. Agrawal**, W. Liao, and A. Choudhary, "Data Compression for the Exascale Computing Era - Survey," Supercomputing Frontiers and Innovations, vol. 1, no. 2, pp. 76–88, 2014.
74. K. Gopalakrishnan, **A. Agrawal**, H. Ceylan, S. Kim, and A. Choudhary, "Knowledge discovery and data mining in pavement inverse analysis," Transport, vol. 28, no. 1, pp. 1–10, 2013.
75. J. S. Mathias, **A. Agrawal**, J. Feinglass, A. J. Cooper, D. W. Baker, and A. Choudhary, "Development of a 5 year life expectancy index in older adults using predictive mining of electronic health record data," Journal of the American Medical Informatics Association, vol. 20, pp. e118–e124, 2013. **JSM and AA are co-first authors**.
76. Y. Xie, Z. Chen, K. Zhang, Y. Cheng, D. K. Honbo, **A. Agrawal**, and A. Choudhary, "MuSES: a Multilingual Sentiment Elicitation System for Social Media Data," IEEE Intelligent Systems, vol. 99, pp. 1541–1672, 2013.
77. **A. Agrawal**, S. Misra, R. Narayanan, L. Polepeddi, and A. Choudhary, "Lung cancer survival prediction using ensemble data mining on SEER data," Scientific Programming, vol. 20, no. 1, pp. 29–42, 2012.
78. Y. Zhang, S. Misra, **A. Agrawal**, M. M. A. Patwary, W. Liao, Z. Qin, and A. Choudhary, "Accelerating pairwise statistical significance estimation for local alignment by harvesting GPU's power," BMC Bioinformatics, vol. 13, no. Suppl 5, p. S3, 2012.
79. Y. Zhang, M. M. A. Patwary, S. Misra, **A. Agrawal**, W. Liao, Z. Qin, and A. Choudhary, "Par-psse: Software for pairwise statistical significance estimation in parallel for local sequence alignment," International Journal of Digital Content Technology and its Applications (JDCTA), vol. 6, no. 5, pp. 200–208, 2012.
80. Y. Zhang, F. Zhou, J. Gou, H. Xiao, Z. Qin, and **A. Agrawal**, "Accelerating Pairwise Statistical Significance Estimation Using NUMA Machine," Journal of Computational Information Systems (JCIS), vol. 8, no. 9, pp. 3887–3894, 2012.
81. **A. Agrawal** and A. Choudhary, "Association Rule Mining Based HotSpot Analysis on SEER Lung Cancer Data," International Journal of Knowledge Discovery in Bioinformatics (IJKDB), vol. 2, no. 2, pp. 34–54, 2011.
82. **A. Agrawal** and X. Huang, "Pairwise Statistical Significance of Local Sequence Alignment Using Sequence-Specific and Position-Specific Substitution Matrices," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 8, no. 1, pp. 194–205, 2011.
83. **A. Agrawal**, S. Misra, D. Honbo, and A. Choudhary, "Parallel pairwise statistical significance estimation of local sequence alignment using Message Passing Interface library," Concurrency and Computation: Practice and Experience, vol. 23, no. 17, pp. 2269–2279, 2011.
84. S. Misra, **A. Agrawal**, W. Liao, and A. Choudhary, "Anatomy of a hash-based long read sequence mapping algorithm for next generation DNA sequencing," Bioinformatics, vol. 27, no. 2, pp. 189–195, 2011.
85. **A. Agrawal**, A. Mittal, R. Jain, and R. Takkar, "Fuzzy-adaptive-thresholding-based exon prediction," International Journal of Computational Biology and Drug Design, vol. 3, no. 4, pp. 311–333, 2010.
86. **A. Agrawal** and X. Huang, "Pairwise Statistical Significance of Local Sequence Alignment Using Multiple Parameter Sets and Empirical Justification of Parameter Set Change Penalty," BMC Bioinformatics, vol. 10, no. Suppl 3, p. S1, 2009.
87. **A. Agrawal** and X. Huang, "PSIBLAST_PairwiseStatSig: Reordering PSI-BLAST hits using pairwise statistical significance," Bioinformatics, vol. 25, no. 8, pp. 1082–1083, 2009.

88. **A. Agrawal**, V. P. Brendel, and X. Huang, "Pairwise Statistical Significance and Empirical Determination of Effective Gap Opening Penalties for Protein Local Sequence Alignment," International Journal of Computational Biology and Drug Design, vol. 1, no. 4, pp. 347–367, 2008.
89. **A. Agrawal** and A. Mittal, "A dynamic time-lagged correlation based method to learn multi-time delay gene networks," World Academy of Science, Engineering and Technology, vol. 9, pp. 723–730, 2007.
90. **A. Agrawal**, A. Mittal, and S. Gupta, "A Restrictive Mining Algorithm to Learn Multi Time Delay Gene Networks," Bioinformatics Trends - A Journal of Bioinformatics and Its Applications, vol. 2, no. 3, 4, pp. 125–140, 2007.

Peer-Reviewed Conference Publications

91. A. L. Day, C. B. Wahl, R. dos Reis, W. Liao, V. P. Dravid, A. Choudhary, and **A. Agrawal**, "Automated Nanoparticle Image Processing Pipeline for AI-Driven Materials Characterization," in Proceedings of 33rd ACM International Conference on Information and Knowledge Management (CIKM), 2024, pp. 4462–4469.
92. V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Combining Transfer Learning and Representation Learning to Improve Predictive Analytics on Small Materials Data," in Proceedings of 23rd IEEE International Conference on Machine Learning and Applications (ICMLA), 2024, pp. 981–984.
93. M. N. T. Kilic, V. Gupta, Y. Mao, K. Wang, A. Peltekian, A. Choudhary, W. Liao, and **A. Agrawal**, "Enhancing Deep Neural Network Classification Performance through Novel Weight Initialization: t-SNE Supported Walsh Matrix Approach," in Proceedings of 23rd IEEE International Conference on Machine Learning and Applications (ICMLA), 2024, pp. 1021–1024.
94. M. N. T. Kilic, Y. Mao, V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Tackling the Nonlinearity Problem in Inverse Modeling: Mixture Density Network-Backed Quantized AutoEncoder," in Proceedings of 23rd IEEE International Conference on Machine Learning and Applications (ICMLA), 2024, pp. 997–1000.
95. K. Wang, Y. Mao, M. Hasan, M. M. Billah, M. N. T. Kilic, V. Gupta, W. Liao, A. Choudhary, P. Acar, and **A. Agrawal**, "Deep Learning Based Inverse Modeling for Materials Design: From Microstructure and Property to Processing," in Proceedings of 23rd IEEE International Conference on Machine Learning and Applications (ICMLA), 2024, pp. 236–241.
96. V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Pre-Activation based Representation Learning to Enhance Predictive Analytics on Small Materials Data," in Proceedings of International Joint Conference on Neural Networks (IJCNN), 2023, pp. 1–8.
97. V. Gupta, Y. Lyu, D. Suarez, Y. Mao, W. Liao, A. Choudhary, W. K. Liu, G. Cusatis, and **A. Agrawal**, "Physics-based Data-Augmented Deep Learning for Enhanced Autogenous Shrinkage Prediction on Experimental Dataset," in Proceedings of the 2023 Fifteenth International Conference on Contemporary Computing (IC3), 2023, pp. 188–197.
98. Z. Huang, K. Hou, **A. Agrawal**, A. Choudhary, R. Ross, and W. Liao, "I/O in WRF: A Case Study in Modern Parallel I/O Techniques," in Proceedings of 35th International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC'23), 2023, pp. 1–13.
99. C. Lee, V. Hewes, G. Cerati, J. Kowalkowski, A. Aurisano, **A. Agrawal**, A. Choudhary, and W. Liao, "A Case Study of Data Management Challenges Presented in Large-Scale Machine Learning Workflows," in Proceedings of 23rd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid), 2023, pp. 71–81.

100. Y. Mao, M. Hasan, C. Lee, M. N. T. Kilic, V. Gupta, W. Liao, A. Choudhary, P. Acar, and **A. Agrawal**, "A deep learning framework for time-series processing-microstructure-property prediction," in Proceedings of 22nd IEEE International Conference on Machine Learning and Applications (ICMLA), 2023, pp. 890–893.

101. Y. Mao, S. Keshavarz, V. Gupta, A. Reid, W. Liao, A. Choudhary, and **A. Agrawal**, "AI for Learning Deformation Behavior of a Material: Predicting Stress-Strain Curves 4000x Faster Than Simulations," in Proceedings of International Joint Conference on Neural Networks (IJCNN), 2023, pp. 1–8.

102. V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "BRNet: Branched Residual Network for Fast and Accurate Predictive Modeling of Materials Properties," in Proceedings of the 2022 SIAM International Conference on Data Mining (SDM), 2022, pp. 343–351.

103. V. Gupta, W. Liao, A. Choudhary, and **A. Agrawal**, "Which Deep Learning Framework Should I Use: A Comparative Study for Deep Regression Modeling," in Proceedings of 2022 International Conference on Computational Science and Computational Intelligence (CSCI), 2022, pp. 72–77.

104. Y. Mao, V. Gupta, K. Wang, W. Liao, A. Choudhary, and **A. Agrawal**, "To Shuffle or Not to Shuffle: Mini-Batch Shuffling Strategies for Multi-class Imbalanced Data Classification," in Proceedings of 2022 International Conference on Computational Science and Computational Intelligence (CSCI), 2022, pp. 298–301.

105. A. Sardeshmukh, S. Reddy, B. P. Gautham, and **A. Agrawal**, "Machine Learning for Materials Science (MLMS)," in Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2022, pp. 4902–4903.

106. K. Wang, S. Lee, J. Balewski, A. Sim, P. Nugent, **A. Agrawal**, A. Choudhary, K. Wu, and W. Liao, "Using Multi-Resolution Data to Accelerate Neural Network Training in Scientific Applications," in Proceedings of 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid), 2022, pp. 404–413.

107. R. Al-Bahrani, D. Jha, Q. Kang, S. Lee, Z. Yang, W. Liao, **A. Agrawal**, and A. Choudhary, "SIGRNN: Synthetic minority Instances Generation in imbalanced datasets using a Recurrent Neural Network," in Proceedings of International Conference on Pattern Recognition Applications and Methods (ICPRAM), 2021, pp. 349–356.

108. J. Hewes, A. Aurisano, G. Cerati, J. Kowalkowski, C. Lee, W. Liao, A. Day, **A. Agrawal**, M. Spiropulu, J.-R. Vlimant, L. Gray, T. Klijnsma, P. Calafiura, S. Conlon, S. Farrell, X. Ju, and D. Murnane, "Graph Neural Network for Object Reconstruction in Liquid Argon Time Projection Chambers," in Proceedings of 25th International Conference on Computing in High Energy and Nuclear Physics (CHEP), 2021, p. 03054.

109. K. Hou, Q. Kang, S. Lee, **A. Agrawal**, A. Choudhary, and W. Liao, "Supporting Data Compression in PnetCDF," in Proceedings of 2021 IEEE International Conference on Big Data (BigData), 2021, pp. 86–97.

110. D. Jha, K. V. L. V. Narayanachari, R. Zhang, D. T. Keane, W. Liao, A. Choudhary, Y.-W. Chung, M. J. Bedzyk, and **A. Agrawal**, "Enhancing Phase Mapping for High-throughput X-ray Diffraction Experiments using Fuzzy Clustering," in Proceedings of International Conference on Pattern Recognition Applications and Methods (ICPRAM), 2021, pp. 507–514.

111. S. Lee, Q. Kang, K. Wang, J. Balewski, A. Sim, **A. Agrawal**, A. Choudhary, P. Nugent, K. Wu, and W. Liao, "Asynchronous I/O Strategy for Large-Scale Deep Learning Applications," in Proceedings of 28th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC'21), 2021, pp. 322–331.

112. S. Lee, Q. Kang, **A. Agrawal**, A. Choudhary, and W. Liao, "Communication-Efficient Local Stochastic Gradient Descent for Scalable Deep Learning," in Proceedings of 2020 IEEE International Conference on Big Data (BigData), 2020, pp. 718–727.

113. Q. Kang, R. Ross, R. Latham, S. Lee, **A. Agrawal**, A. Choudhary, and W. Liao, "Improving all-to-many personalized communication in two-phase I/O," in Proceedings of 32nd International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC'20), 2020, pp. 118–130.

114. Q. Kang, A. Sim, P. Nugent, S. Lee, W. Liao, **A. Agrawal**, A. Choudhary, and K. Wu, "Predicting Resource Requirement in Intermediate Palomar Transient Factory Workflow," in Proceedings of 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid), 2020, pp. 619–628.

115. **A. Agrawal**, A. Saboo, W. Xiong, G. Olson, and A. Choudhary, "Martensite Start Temperature Predictor for Steels Using Ensemble Data Mining," in Proceedings of 6th IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2019, pp. 521–530.

116. D. Jha, A. G. Kusne, R. Al-Bahrani, N. Nguyen, W. Liao, A. Choudhary, and **A. Agrawal**, "Peak Area Detection Network for Directly Learning Phase Regions from Original X-ray Diffraction Patterns," in Proceedings of International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8.

117. D. Jha, L. Ward, Z. Yang, C. Wolverton, I. Foster, W. Liao, A. Choudhary, and **A. Agrawal**, "IRNet: A General Purpose Deep Residual Regression Framework for Materials Discovery," in Proceedings of 25th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD), 2019, pp. 2385–2393.

118. S. Lee, Q. Kang, S. Madireddy, P. Balaprakash, **A. Agrawal**, A. Choudhary, R. Archibald, and W. Liao, "Improving Scalability of Parallel CNN Training by Adjusting Mini-Batch Size at Run-Time," in Proceedings of 2019 IEEE International Conference on Big Data (BigData), 2019, pp. 830–839.

119. A. Paul, D. Jha, R. Al-Bahrani, W. Liao, A. Choudhary, and **A. Agrawal**, "Transfer Learning Using Ensemble Neural Nets for Organic Solar Cell Screening," in Proceedings of International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8.

120. A. Paul, M. Mozaffar, Z. Yang, W. Liao, A. Choudhary, J. Cao, and **A. Agrawal**, "A real-time iterative machine learning approach for temperature profile prediction in additive manufacturing processes," in Proceedings of 6th IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2019, pp. 541–550.

121. Z. Yang, R. Al-Bahrani, A. Reid, S. Papanikolaou, S. Kalidindi, W. Liao, A. Choudhary, and **A. Agrawal**, "Deep learning based domain knowledge integration for small datasets: Illustrative applications in materials informatics," in Proceedings of International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8.

122. D. Han, **A. Agrawal**, W. Liao, and A. Choudhary, "Parallel DBSCAN Algorithm Using a Data Partitioning Strategy with Spark Implementation," in Proceedings of 2018 IEEE Conference on Big Data (BigData), 2018, pp. 305–312.

123. Q. Kang, J. L. Träff, R. Al-Bahrani, **A. Agrawal**, A. Choudhary, and W. Liao, "Full-Duplex Inter-Group All-to-All Broadcast Algorithms with Optimal Bandwidth," in Proceedings of the 25th European MPI Users' Group Meeting (EuroMPI 2018), 2018, p. 10.

124. X. Li, Z. Yang, L. C. Brinson, A. Choudhary, **A. Agrawal**, and W. Chen, "A Deep Adversarial Learning Methodology for Designing Microstructural Material Systems," in Proceedings of the ASME 2018 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference (IDETC 2018), 2018, p. V02BT03A008.

125. R. Al-Bahrani, M. K. Danilovich, W. Liao, A. Choudhary, and **A. Agrawal**, "Towards Identifying Informal Caregivers of Alzheimer's and Dementia Patients in Social Media," in Proceedings of the Fifth International Conference on Healthcare Informatics (ICHI), 2017, pp. 324–324.

126. K. Lee, **A. Agrawal**, and A. Choudhary, "Forecasting Influenza Levels using Real-Time Social Media Streams," in Proceedings of the Fifth International Conference on Healthcare Informatics (ICHI), 2017, pp. 409–414.

127. K. Lee, S. A. Hasan, O. Farri, A. Choudhary, and **A. Agrawal**, "Medical Concept Normalization for Online User-Generated Texts," in Proceedings of the Fifth International Conference on Healthcare Informatics (ICHI), 2017, pp. 462–469.

128. S. Lee, D. Jha, **A. Agrawal**, A. Choudhary, and W. Liao, "Parallel Deep Convolutional Neural Network Training by Exploiting the Overlapping of Computation and Communication," in Proceedings of 24th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC'17), 2017, pp. 183–192.

129. E. Rangel, N. Frontiere, S. Habib, K. Heitmann, W. Liao, **A. Agrawal**, and A. Choudhary, "Building Halo Merger Trees from the Q Continuum Simulation," in Proceedings of 24th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC'17), 2017, pp. 398–407.

130. Y. Xie, Z. Chen, **A. Agrawal**, and A. Choudhary, "Distinguish polarity in bag-of-words visualization," in Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), 2017, pp. 3344–3350.

131. **A. Agrawal** and A. Choudhary, "A Fatigue Strength Predictor for Steels Using Ensemble Data Mining," in Proceedings of 25th ACM International Conference on Information and Knowledge Management (CIKM) (Demo), 2016, pp. 2497–2500.

132. **A. Agrawal**, J. Mathias, D. Baker, and A. Choudhary, "Identifying HotSpots in Five Year Survival Electronic Health Records of Older Adults," in Proceedings of 6th IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), 2016, pp. 1–6.

133. **A. Agrawal**, J. Mathias, D. Baker, and A. Choudhary, "Five Year Life Expectancy Calculator for Older Adults," in Proceedings of IEEE International Conference on Data Mining (ICDM) (Demo), 2016, pp. 1280–1283.

134. **A. Agrawal**, B. Meredig, C. Wolverton, and A. Choudhary, "A Formation Energy Predictor for Crystalline Materials Using Ensemble Data Mining," in Proceedings of IEEE International Conference on Data Mining (ICDM) (Demo), 2016, pp. 1276–1279.

135. R. Al-Bahrani, M. K. Danilovich, **A. Agrawal**, and A. Choudhary, "Towards Informal Caregiver Identification in Social Media," in Proceedings of IEEE International Conference on Healthcare Informatics (ICHI) (Poster), 2016, pp. 414–414.

136. Q. Kang, W. Liao, **A. Agrawal**, and A. Choudhary, "A Filtering-based Clustering Algorithm for Improving Spatio-temporal Kriging Interpolation Accuracy," in Proceedings of 25th ACM International Conference on Information and Knowledge Management (CIKM), 2016, pp. 2209–2214.

137. D. Palsetia, W. Hendrix, S. Lee, **A. Agrawal**, W. Liao, and A. Choudhary, "Parallel Community Detection Algorithm Using a Data Partitioning Strategy with Pairwise Subdomain Duplication," in High Performance Computing, 31st International Conference, ISC High Performance 2016, Frankfurt, Germany, June 19-23, 2016, Proceedings, 2016, pp. 98–115.

138. E. Rangel, N. Li, S. Habib, T. Peterka, **A. Agrawal**, W. Liao, and A. Choudhary, "Parallel DTFE Surface Density Field Reconstruction," in Proceedings of IEEE Cluster, 2016, pp. 30–39. Won the best paper award.

139. Z. Yuan, W. Hendrix, S. W. Son, C. Federrath, **A. Agrawal**, W. Liao, and A. Choudhary, "Parallel Implementation of Lossy Data Compression for Temporal Data Sets," in Proceedings of 23rd Annual International Conference on High Performance Computing, Data, and Analytics (HiPC'16), 2016, pp. 62–71.

140. Y. Cheng, **A. Agrawal**, H. Liu, and A. Choudhary, "Legislative prediction with dual uncertainty minimization from heterogeneous information," in Proceedings of the 15th SIAM International Conference on Data Mining (SDM), 2015, pp. 361–369.

141. C. Jin, Z. Chen, W. Hendrix, **A. Agrawal**, and A. Choudhary, "Incremental, Distributed Single-Linkage Hierarchical Clustering Algorithm Using MapReduce," in Proceedings of the 23rd High Performance Computing Symposium (HPC), 2015, pp. 83–92.

142. C. Jin, Q. Fu, H. Wang, W. Hendrix, Z. Chen, **A. Agrawal**, A. Banerjee, and A. Choudhary, "Running MAP Inference on Million Node Graphical Models: A High Performance Computing Perspective," in Proceedings of the 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), 2015, pp. 565–575.

143. C. Jin, R. Liu, Z. Chen, W. Hendrix, **A. Agrawal**, and A. Choudhary, "A Scalable Hierarchical Clustering Algorithm Using Spark," in Proceedings of IEEE International Conference on Big Data Computing Service and Applications (BigDataService), 2015, pp. 418–426.

144. K. Lee, **A. Agrawal**, and A. Choudhary, "Mining Social Media Streams to Improve Public Health Allergy Surveillance," in Proceedings of IEEE/ACM International Conference on Social Networks Analysis and Mining (ASONAM), 2015, pp. 815–822.

145. R. Liu, **A. Agrawal**, Z. Chen, W. Liao, and A. Choudhary, "Pruned Search: A Machine Learning Based Meta-Heuristic Approach for Constrained Continuous Optimization," in Proceedings of 8th IEEE International Conference on Contemporary Computing (IC3), 2015, pp. 13–18.

146. V. Rastogi and **A. Agrawal**, "All Your Google and Facebook Logins are Belong to Us: A Case for Single Sign-off," in Proceedings of 8th IEEE International Conference on Contemporary Computing (IC3), 2015, pp. 416–421.

147. Y. Xie, P. Daga, Y. Cheng, K. Zhang, **A. Agrawal**, and A. Choudhary, "Reducing infrequent-token perplexity via variational corpora," in Proceedings of the 53rd Annual Meeting of the Association of Computational Linguistics (ACL) and the 7th International Joint Conference on Natural Language Processing, 2015, pp. 609–615.

148. Y. Cheng, **A. Agrawal**, H. Liu, and A. Choudhary, "Social Role Identification via Dual Uncertainty Minimization Regularization," in Proceedings of International Conference on Data Mining (ICDM), 2014, pp. 767–772.

149. Z. Chen, S. W. Son, W. Hendrix, **A. Agrawal**, W. Liao, and A. Choudhary, "NUMARCK: Machine Learning Algorithm for Resiliency and Checkpointing," in Proceedings of 26th International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC'14), 2014, pp. 733–744.

150. D. Palsetia, M. Patwary, W. Hendrix, **A. Agrawal**, and A. Choudhary, "Clique Guided Community Detection," in Proceedings of 2014 IEEE International Conference on Big Data (BigData), 2014, pp. 500–509.

151. Y. Xie, Z. Chen, D. Palsetia, **A. Agrawal**, and A. Choudhary, "Indexing Bipartite Memberships in Web Graphs," in Proceedings of IEEE/ACM International Conference on Social Networks Analysis and Mining (ASONAM) 2014, 2014, pp. 166–173.

152. Y. Xie, D. Palsetia, G. Trajcevski, **A. Agrawal**, and A. Choudhary, "Silverback: Scalable Association Mining For Temporal Data in Columnar Probabilistic Databases," in Proceedings of 30th IEEE International Conference on Data Engineering (ICDE), Industrial and Applications Track, 2014, pp. 1072–1083.

153. Y. Cheng, Z. Chen, L. Liu, J. Wang, **A. Agrawal**, and A. Choudhary, "Feedback-Driven Multiclass Active Learning for Data Streams," in Proceedings of 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, USA, Oct. 2013, 2013, pp. 1311–1320.

154. Y. Cheng, Z. Chen, J. Wang, **A. Agrawal**, and A. Choudhary, "Bootstrapping Active Name Disambiguation with Crowdsourcing," in Proceedings of 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, USA, Oct. 2013 (Poster paper), 2013, pp. 1213–1216.

155. Z. Chen, Y. Xie, Y. Cheng, K. Zhang, **A. Agrawal**, W. Liao, N. Samatova, and A. Choudhary, "Forecast Oriented Classification of Spatio-Temporal Extreme Events," in Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI), 2013, pp. 2952–2954.

156. Y. Cheng, Y. Xie, Z. Chen, **A. Agrawal**, A. Choudhary, and S. Guo, "JobMiner: A Real-time System for Mining Job-related Patterns from Social Media," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD) (Demo paper), 2013, pp. 1450–1453.

157. P. D. Deshpande, B. P. Gautham, A. Cecen, S. Kalidindi, **A. Agrawal**, and A. Choudhary, "Application of Statistical and Machine Learning Techniques for Correlating Properties to Composition and Manufacturing Processes of Steels," in 2nd World Congress on Integrated Computational Materials Engineering, July 7-11, 2013, Salt Lake City, Utah, 2013, pp. 155–160.

158. W. Hendrix, D. Palsetia, M. M. A. Patwary, **A. Agrawal**, W. Liao, and A. Choudhary, "A Scalable Algorithm for Single-Linkage Hierarchical Clustering on Distributed Memory Architectures," in Proceedings of 3rd IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV 2013), Atlanta GA, USA, Oct. 2013, 2013, pp. 7–13.

159. K. Lee, **A. Agrawal**, and A. Choudhary, "Real-Time Disease Surveillance using Twitter Data: Demonstration on Flu and Cancer," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD) (Demo paper), 2013, pp. 1474–1477.

160. L. Liu, J. Tang, Y. Cheng, **A. Agrawal**, W. Liao, and A. Choudhary, "Mining Diabetes Complication and Treatment Patterns for Clinical Decision Support," in Proceedings of 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, USA, Oct. 2013, 2013, pp. 279–288.

161. M. Patwary, D. Palsetia, **A. Agrawal**, W. Liao, F. Manne, and A. Choudhary, "Scalable Parallel OPTICS Data Clustering Using Graph Algorithmic Techniques," in Proceedings of 25th International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC'13), 2013, pp. 1–12. Article No. 49.

162. Y. Xie, Z. Chen, **A. Agrawal**, A. Choudhary, and L. Liu, "Random Walk-based Graphical Sampling in Unbalanced Heterogeneous Bipartite Social Graphs," in Proceedings of 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013), San Francisco, USA, Oct. 2013 (Poster paper), 2013, pp. 1473–1476.

163. Y. Xie, Z. Chen, **A. Agrawal**, W. Liao, and A. Choudhary, "Caranx: Scalable Social Image Index Using Phylogenetic Tree of Hashtags," in Proceedings of 25th International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing, SC'13) (Poster paper), 2013, pp. 1–2.

164. Y. Xie, Z. Chen, Y. Cheng, K. Zhang, **A. Agrawal**, W. Liao, and A. Choudhary, "Detecting and Tracking Disease Outbreaks by Mining Social Media Data," in Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI), 2013, pp. 2958–2960.

165. Y. Xie, Z. Chen, K. Zhang, M. M. A. Patwary, Y. Cheng, H. Liu, **A. Agrawal**, and A. Choudhary, "Graphical Modeling of Macro Behavioral Targeting in Social Networks," in Proceedings of the 13th SIAM International Conference on Data Mining (SDM), 2013, pp. 740–748.

166. Y. Xie, Z. Chen, K. Zhang, Y. Cheng, C. Jin, **A. Agrawal**, and A. Choudhary, "Elver: Recommending Facebook Pages in Cold Start Situation Without Content Features," in Proceedings of IEEE International Conference on Big Data (BigData), 2013, pp. 475–479.

167. K. Zhang, D. Downey, Z. Chen, Y. Xie, Y. Cheng, **A. Agrawal**, W. Liao, and A. Choudhary, "A Probabilistic Graphical Model for Brand Reputation Assessment in Social Networks," in Proceedings of IEEE/ACM International Conference on Social Networks Analysis and Mining (ASONAM), Niagara Falls, Canada, Aug. 2013, 2013, pp. 223–230.

168. Y. Cheng, K. Zhang, Y. Xie, **A. Agrawal**, and A. Choudhary, "On active learning in hierarchical classification," in Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM) (Poster paper), 2012, pp. 2467–2470.

169. W. Hendrix, M. M. A. Patwary, **A. Agrawal**, W. Liao, and A. Choudhary, "Parallel Hierarchical Clustering on Shared Memory Platforms," in Proceedings of the 19th International Conference on High Performance Computing (HiPC), 2012, pp. 1–9.

170. M. Patwary, D. Palsetia, **A. Agrawal**, W. Liao, F. Manne, and A. Choudhary, "A new scalable parallel DBSCAN algorithm using the disjoint-set data structure," in ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC), 2012, pp. 1–11. Article No. 62.

171. Y. Xie, D. Honbo, A. Choudhary, K. Zhang, Y. Cheng, and **A. Agrawal**, "VOXSUP: a social engagement framework," in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD) (Demo paper), 2012, pp. 1556–1559.

172. K. Zhang, Y. Xie, Y. Cheng, D. Honbo, D. Downey, **A. Agrawal**, W. Liao, and A. Choudhary, "Sentiment identification by incorporating syntax, semantics and context information," in Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval (Poster paper), 2012, pp. 1143–1144.

173. **A. Agrawal**, S. Misra, A. Choudhary, and K. Bilimoria, "Risk prediction for post-operative adverse outcomes in colorectal cancer surgery," in Proc. of IEEE 1st International Conference on Computational Advances in Bio and Medical Sciences (ICCABS) (Poster paper), 2011, pp. 232–232.

174. **A. Agrawal**, S. Misra, R. Narayanan, L. Polepeddi, and A. Choudhary, "A lung cancer mortality risk calculator based on SEER data," in Proc. of IEEE 1st International Conference on Computational Advances in Bio and Medical Sciences (ICCABS) (Poster paper), 2011, pp. 233–233.

175. Y. Zhang, S. Misra, D. Honbo, **A. Agrawal**, W. Liao, and A. Choudhary, "Efficient pairwise statistical significance estimation for local sequence alignment using GPU," in IEEE 1st International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), 2011, pp. 226–231.

176. **A. Agrawal**, A. Choudhary, and X. Huang, "Derived Distribution Points Heuristic for Fast Pairwise Statistical Significance Estimation," in Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology (BCB), 2010, pp. 312–321.

177. **A. Agrawal**, A. Choudhary, and X. Huang, "Non-conservative Pairwise Statistical Significance of Local Sequence Alignment Using Position-Specific Substitution Matrices," in Proceedings of BIOCOMP, 2010, pp. 262–268.

178. D. Honbo, **A. Agrawal**, and A. Choudhary, "Efficient pairwise statistical significance estimation using FPGAs," in Proceedings of BIOCOMP, 2010, pp. 571–577.

179. **A. Agrawal**, V. Brendel, and X. Huang, "Pairwise Statistical Significance Versus Database Statistical Significance for Local Alignment of Protein Sequences," in Bioinformatics Research and Applications, 2008, vol. 4983, pp. 50–61.

180. **A. Agrawal**, A. Ghosh, and X. Huang, "Estimating Pairwise Statistical Significance of Protein Local Alignments Using a Clustering-Classification Approach Based on Amino Acid Composition," in Bioinformatics Research and Applications, 2008, vol. 4983, pp. 62–73.

181. **A. Agrawal** and X. Huang, "Pairwise Statistical Significance of Local Sequence Alignment Using Substitution Matrices with Sequence-Pair-Specific Distance," in Proc. of International Conference on Information Technology, ICIT, 2008, pp. 94–99. **Won the best student paper award**.

182. **A. Agrawal** and X. Huang, "Pairwise DNA Alignment with Sequence Specific Transition-Transversion Ratio Using Multiple Parameter Sets," in Proc. of International Conference on Information Technology, ICIT, 2008, pp. 89–93.

183. **A. Agrawal** and X. Huang, "Conservative, Non-conservative and Average Pairwise Statistical Significance of Local Sequence Alignment," in Proc. of IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2008, pp. 433–436.

184. **A. Agrawal** and X. Huang, "DNAlignTT: Pairwise DNA Alignment with Sequence Specific Transition-Transversion Ratio," in Proc. of IEEE International Conference on Electro/Information Technology (EIT), 2008, pp. 453–455.

185. **A. Agrawal** and S. K. Khaitan, "A new heuristic for multiple sequence alignment," in Proc. of IEEE International Conference on Electro/Information Technology (EIT) (Poster paper), 2008, pp. 215–217.

186. **A. Agrawal**, A. Mittal, R. Jain, and R. Takkar, "An adaptive fuzzy thresholding algorithm for exon prediction," in Proc. of IEEE International Conference on Electro/Information Technology (EIT) (Poster paper), 2008, pp. 211–214.

187. **A. Agrawal** and A. Mittal, "Identifying temporal gene networks using signal processing metrics on time-series gene expression data," in Proc. of 3rd International Conference on Intelligent Sensing and Information Processing (ICISIP), 2005, pp. 86–92.

188. A. Mittal, S. Gupta, **A. Agrawal**, and L. F. Cheong, "Camera Motion Characterization of Establishment Shots and Video Signatures," in Proc. of 12th International Conference on Advanced Computing and Communications (ADCOM), 2004, pp. 663–668.

189. **A. Agrawal**, A. Mittal, and S. Gupta, "Identifying Temporal Gene Networks by Mining Gene Expression Data," in Proc. of 12th International Conference on Advanced Computing and Communications (ADCOM), 2004, pp. 194–200.

Peer-Reviewed Workshop Publications

190. D. Jha, K. V. L. V. Narayanachari, R. Zhang, J. Liao, D. T. Keane, W. Liao, A. Choudhary, Y.-W. Chung, M. J. Bedzyk, and **A. Agrawal**, "An Incremental Phase Mapping Approach for X-ray Diffraction Patterns using Binary Peak Representations," in SDM Workshop on Domain-Driven Data Mining (DDDM), 2021, p. 11.

191. Z. Yang, T. Watari, D. Ichigozaki, A. Mitsutoshi, H. Takahashi, Y. Suga, W. Liao, A. Choudhary, and **A. Agrawal**, "Heterogeneous feature fusion based machine learning on shallow-wide and heterogeneous-sparse industrial datasets," in ICPR Workshop on Industrial Machine Learning (IML), 2021, pp. 566–577.

192. Z. Yang, D. Jha, A. Paul, W. Liao, A. Choudhary, and **A. Agrawal**, "A General Framework Combining Generative Adversarial Networks and Mixture Density Networks for Inverse Modeling in Microstructural Materials Design," in Proceedings of 2020 NeurIPS workshop on Machine Learning for Engineering Modeling, Simulation, and Design (ML4Eng), 2020, pp. 1–8.

193. Q. Kang, **A. Agrawal**, A. Choudhary, A. Sim, K. Wu, R. Kettimuthu, P. Beckman, Z. Liu, and W. Liao, "Spatiotemporal Real-Time Anomaly Detection for Supercomputing Systems," in Proceedings of IEEE BigData Workshop on Big Data Predictive Maintenance using Artificial Intelligence, 2019, pp. 4381–4389.

194. Z. Yang, T. Watari, D. Ichigozaki, K. Morohoshi, Y. Suga, W. Liao, A. Choudhary, and **A. Agrawal**, "Data-driven insights from predictive analytics on heterogeneous experimental data of industrial magnetic materials," in Proceedings of 2019 ICDM Workshop on Learning and Mining with Industrial Data (LMID), 2019, pp. 806–813.

195. K. Hou, R. Al-Bahrani, E. Rangel, **A. Agrawal**, R. Latham, R. Ross, A. Choudhary, and W. Liao, "Integration of Burst Buffer in High-level Parallel I/O Library for Exa-scale Computing Era," in Proceedings of SC workshop on Parallel Data Storage and Data Intensive Scalable Computing Systems (PDSW-DISCS 2018), 2018, pp. 1–12.

196. Q. Kang, **A. Agrawal**, A. Choudhary, and W. Liao, "Optimal Algorithms for Half-Duplex Inter-Group All-to-All Broadcast on Fully Connected and Ring Topologies," in Proceedings of 2018 SC Workshop on Exascale MPI (ExaMPI), 2018, pp. 1–11.

197. R. Kettimuthu, Z. Liu, I. Foster, P. Beckman, A. Sim, K. Wu, W. Liao, Q. Kang, **A. Agrawal**, and A. Choudhary, "Towards Autonomic Science Infrastructure: Architecture, Limitations, and Open Issues," in Proceedings of the 1st International Workshop on Autonomous Infrastructure for Science, 2018.

198. S. Lee, **A. Agrawal**, P. Balaprakash, A. Choudhary, and W. Liao, "Communication-Efficient Parallelization Strategy for Deep Convolutional Neural Network Training," in Proceedings of SC workshop on Machine Learning in HPC Environments (MLHPC 2018), 2018, pp. 47–56.

199. A. Paul, D. Jha, R. Al-Bahrani, W. Liao, A. Choudhary, and **A. Agrawal**, "CheMixNet: Mixed DNN Architectures for Predicting Chemical Properties using Multiple Molecular Representations," in Proceedings of 2018 NIPS workshop on Machine Learning for Molecules and Materials (MLMM), 2018, pp. 1–13.

200. R. Al-Bahrani, M. K. Danilovich, W. Liao, A. Choudhary, and **A. Agrawal**, "Analyzing Informal Caregiving Expression in Social Media," in IEEE ICDM Workshop on Sentiment Elicitation from Natural Text for Information Retrieval and Extraction (SENTIRE), 2017, pp. 342–349.

201. Q. Kang, W. Liao, **A. Agrawal**, and A. Choudhary, "A Hybrid Training Algorithm for Recurrent Neural Network Using Particle Swarm Optimization-based Preprocessing and Temporal Error Aggregation," in IEEE ICDM Workshop on Optimization Based Techniques for Emerging Data Mining Problems (OEDM), 2017, pp. 812–817.

202. D. Han, **A. Agrawal**, W. Liao, and A. Choudhary, "A novel scalable DBSCAN algorithm with Spark," in Proceedings of 5th IEEE IPDPS Workshop on Parallel and Distributed Computing for Large Scale Machine Learning and Big Data Analytics (ParLearning), 2016, pp. 1393–1402.

203. A. Krishna, **A. Agrawal**, and A. Choudhary, "Predicting the Outcome of Startups: Less Failure, More Success," in Proceedings of IEEE ICDM Workshop on Data Market for Co-evolution of Sciences and Business (MoDAT), 2016, pp. 798–805.

204. R. Liu, **A. Agrawal**, W. Liao, M. D. Graef, and A. Choudhary, "Materials Discovery: Understanding Polycrystals from Large-Scale Electron Patterns," in Proceedings of IEEE BigData Workshop on Advances in Software and Hardware for Big Data to Knowledge Discovery (ASH), 2016, pp. 2261–2269.

205. S. Lee, W. Liao, **A. Agrawal**, N. Hardavellas, and A. Choudhary, "Evaluation of K-Means Data Clustering Algorithm on Intel Xeon Phi," in Proceedings of IEEE BigData Workshop on Advances in Software and Hardware for Big Data to Knowledge Discovery (ASH), 2016, pp. 2251–2260.

206. R. Liu, D. Palsetia, A. Paul, R. Al-Bahrani, D. Jha, W. Liao, **A. Agrawal**, and A. Choudhary, "PinterNet: A Thematic Label Curation Tool for Large Image Datasets," in Proceedings of IEEE BigData Workshop on Open Science in Big Data (OSBD), 2016, pp. 2353–2362.

207. R. Liu, L. Ward, C. Wolverton, **A. Agrawal**, W. Liao, and A. Choudhary, "Deep Learning for Chemical Compound Stability Prediction," in Proceedings of ACM SIGKDD Workshop on Large-scale Deep Learning for Data Mining (DL-KDD), 2016, pp. 1–7.

208. A. Paul, **A. Agrawal**, W. Liao, and A. Choudhary, "AnonyMine: Mining anonymous social media posts using psycho-lingual and crowd-sourced dictionaries," in Proceedings of ACM SIGKDD Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM), 2016.

209. S. Gupta, D. Palsetia, M. M. A. Patwary, **A. Agrawal**, and A. Choudhary, "A New Parallel Algorithm for Two-Pass Connected Component Labeling," in Proceedings of IEEE IPDPS Workshop on Multithreaded Architectures and Applications (MTAAP), 2014, pp. 1355–1362.

210. R. Liu, **A. Agrawal**, W. Liao, and A. Choudhary, "Enhancing Financial Decision-Making Using Social Behavior Modeling," in Proceedings of 8th KDD Workshop on Social Network Mining and Analysis for Business, Consumer and Social Insights (SNAKDD), 2014, pp. 13:1–13:5. Article No. 13.

211. R. Liu, **A. Agrawal**, W. Liao, and A. Choudhary, "Search Space Preprocessing in Solving Complex Optimization Problems," in Proceedings of IEEE BigData Workshop on Complexity for Big Data (C4BD), 2014, pp. 1–5.

212. Y. Xie, Y. Cheng, **A. Agrawal**, and A. Choudhary, "Estimating online user location distribution without GPS location," in Proceedings of ICDM Workshop on Connecting Online and Offline Social Network Analysis (COOL-SNA), 2014, pp. 936–943.

213. **A. Agrawal**, R. Al-Bahrani, R. Merkow, K. Bilimoria, and A. Choudhary, "Colon Surgery Outcome Prediction Using ACS NSQIP Data," in Proceedings of the KDD Workshop on Data Mining for Healthcare (DMH), 2013, pp. 1–6.

214. **A. Agrawal**, R. Al-Bahrani, J. Raman, M. J. Russo, and A. Choudhary, "Lung Transplant Outcome Prediction using UNOS Data," in Proceedings of the IEEE Big Data Workshop on Bioinformatics and Health Informatics (BHI), 2013, pp. 1–8.

215. **A. Agrawal** and A. Choudhary, "An Analysis of Variation in Hospital Billing Using Medicare Data," in Proceedings of the KDD Workshop on Data Mining for Healthcare (DMH), 2013, pp. 1–6.

216. **A. Agrawal**, J. Raman, M. J. Russo, and A. Choudhary, "Heart Transplant Outcome Prediction using UNOS Data," in Proceedings of the KDD Workshop on Data Mining for Healthcare (DMH), 2013, pp. 1–6.

217. R. Al-Bahrani, **A. Agrawal**, and A. Choudhary, "Colon cancer survival prediction using ensemble data mining on SEER data," in Proceedings of the IEEE Big Data Workshop on Bioinformatics and Health Informatics (BHI), 2013, pp. 9–16.

218. C. Jin, Q. Fu, H. Wang, **A. Agrawal**, W. Hendrix, W. Liao, M. M. A. Patwary, A. Banerjee, and A. Choudhary, "Solving Combinatorial Optimization Problems using Relaxed Linear Programming: A High Performance Computing Perspective," in Proceedings of the KDD Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine), 2013, pp. 39–46. **Won the best paper award**.

219. C. Jin, M. M. A. Patwary, **A. Agrawal**, W. Hendrix, W. Liao, and A. Choudhary, "DiSC: A Distributed Single-Linkage Hierarchical Clustering Algorithm using MapReduce," in Proceedings of the 4th International SC Workshop on Data Intensive Computing in the Clouds (DataCloud) 2013, 2013, pp. 1–6.

220. K. Lee, **A. Agrawal**, and A. Choudhary, "Real-Time Digital Flu Surveillance using Twitter Data," in Proceedings of the SDM Workshop on Data Mining for Medicine and Healthcare (DMMH), 2013, pp. 19–27.

221. Y. Cheng, Y. Xie, K. Zhang, **A. Agrawal**, and A. Choudhary, "CluChunk: clustering large scale user-generated content incorporating chunklet information," in Proceedings of the KDD Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine), 2012, pp. 12–19.

222. Y. Cheng, Y. Xie, K. Zhang, **A. Agrawal**, and A. Choudhary, "How Online Content is Received by Users in Social Media: A Case Study on Facebook. com Posts," in Proceedings of the KDD Workshop on Social Media Analytics (SOMA), 2012, pp. 1–8.

223. D. Palsetia, M. M. A. Patwary, K. Zhang, K. Lee, C. Moran, Y. Xie, D. Honbo, **A. Agrawal**, W. Liao, and A. Choudhary, "User-Interest based Community Extraction in Social Networks," in Proceedings of the KDD Workshop on Social Network Mining and Analysis (SNAKDD), 2012, pp. 1–4.

224. Y. Xie, Y. Cheng, D. Honbo, K. Zhang, **A. Agrawal**, A. Choudhary, Y. Gao, and J. Gou, "Probabilistic macro behavioral targeting," in Proceedings of the CIKM workshop on Data-driven user behavioral modelling and mining from social media (DUBMMSM), 2012, pp. 7–10.

225. Y. Xie, Y. Cheng, D. Honbo, K. Zhang, **A. Agrawal**, and A. Choudhary, "Crowdsourcing recommendations from social sentiment," in Proceedings of KDD Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM), 2012, pp. 9:1–9:8.

226. **A. Agrawal** and A. Choudhary, "Identifying HotSpots in Lung Cancer Data Using Association Rule Mining," in 2nd IEEE ICDM Workshop on Biological Data Mining and its Applications in Healthcare (BioDM), 2011, pp. 995–1002.

227. **A. Agrawal**, S. Misra, R. Narayanan, L. Polepeddi, and A. Choudhary, "A lung cancer outcome calculator using ensemble data mining on SEER data," in Proceedings of the Tenth International Workshop on Data Mining in Bioinformatics (BIOKDD), New York, NY, USA, 2011, pp. 1–9.

228. Y. Cheng, K. Zhang, Y. Xie, **A. Agrawal**, W. Liao, and A. Choudhary, "Learning to Group Web Text Incorporating Prior Information," in 6th IEEE ICDM Workshop on Optimization Based Techniques for Emerging Data Mining Problems, (OEDM), 2011, pp. 212–219.

229. W. Hendrix, I. Tetteh, **A. Agrawal**, F. Semazzi, W. Liao, and A. Choudhary, "Community Dynamics and Analysis of Decadal Trends in Climate Data," in 3rd IEEE ICDM Workshop on Knowledge Discovery from Climate Data, (ClimKD), 2011, pp. 9–14.

230. K. Lee, D. Palsetia, R. Narayanan, M. Patwary, **A. Agrawal**, and A. Choudhary, "Twitter Trending Topic Classification," in 6th IEEE ICDM Workshop on Optimization Based Techniques for Emerging Data Mining Problems, (OEDM), 2011, pp. 251–258.

231. N. Nakka, **A. Agrawal**, and A. N. Choudhary, "Predicting Node Failure in High Performance Computing Systems from Failure and Usage Logs," in IEEE International Symposium on Parallel and Distributed Processing Symposium (IPDPS) Workshops, 2011, pp. 1557–1566.

232. L. Polepeddi, **A. Agrawal**, and A. Choudhary, "Poll: A Citation-Text-Based System for Identifying High-Impact Contributions of an Article," in IEEE ICDM Workshop on Data Mining in Networks, (DaMNet), 2011, pp. 965–968.

233. K. Zhang, Y. Cheng, Y. Xie, **A. Agrawal**, D. Palsetia, K. Lee, and A. Choudhary, "SES: Sentiment Elicitation System for Social Media Data," in IEEE ICDM Workshop on Sentiment Elicitation from Natural Text for Information Retrieval and Extraction, (SENTIRE), 2011, pp. 129–136.

234. Y. Zhang, M. Patwary, S. Misra, **A. Agrawal**, W. Liao, and A. N. Choudhary, "Enhancing parallelism of pairwise statistical significance estimation for local sequence alignment," in HiPC Workshop on Hybrid Multi-core Computing (WHMC), 2011, pp. 1–8.

235. **A. Agrawal**, S. Misra, D. Honbo, and A. Choudhary, "MPIPairwiseStatSig: Parallel Pairwise Statistical Significance Estimation of Local Sequence Alignment," in Proceedings of the HPDC Workshop on Emerging Computational Methods for the Life Sciences (ECMLS), 2010, pp. 470–476.

236. **A. Agrawal** and X. Huang, "Pairwise Statistical Significance of Local Sequence Alignment Using Multiple Parameter Sets," in Proc. of ACM 2nd International Workshop on Data and Text Mining in Bioinformatics (DTMBIO), 2008, pp. 53–60.

Scientific Presentations and Talks

1. (Keynote) AI+HPC for Accelerating Science and Engineering, *2nd International Conference on Recent Advances in Artificial Intelligence, Computer Vision & Smart Systems (ICRACS)*, April 16, 2025, Poornima Institute of Engineering & Technology (PIET), Jaipur, Rajasthan, India.

2. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Science, *5th International Conference on AI, Machine Learning and Applications (AIMLA 2025)*, February 15, 2025, Chennai, India (Virtual).

3. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Science, *15th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2025*, January 17, 2025, Amity University, Noida, Delhi NCR, India.
4. (Invited) AI for Materials: Research Landscape of AI and Design", *IIN Workshop on AI for Materials*, November 22, 2024, Evanston, IL, USA.
5. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Science, *SciX 2024*, October 21, 2024, Raleigh, NC, USA.
6. (Keynote) AI+HPC for Accelerating Science and Engineering, *International Conference on Artificial Intelligence and Emerging Technology (Global AI Summit) 2024*, September 06, 2024, Bennett University, Greater Noida, UP, India.
7. (Keynote) AI+HPC for Accelerating Science and Engineering, *14th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2024*, January 19, 2024, Amity University, Noida, Delhi NCR, India.
8. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Science and Engineering, *Computers, Materials & Continua (CMC) Webinar*, December 21, 2023 (Virtual).
9. A Deep Learning Framework for Time-Series Processing-Microstructure-Property Prediction, *22nd IEEE International Conference on Machine Learning and Applications (ICMLA) 2023*, December 15, 2023, Jacksonville, FL, USA.
10. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Science and Engineering, *7th International Conference on Image Information Processing (ICIIP) 2023*, November 23, 2023, Jaypee University of Information Technology, Solan, HP, India.
11. Physics-based Data-Augmented Deep Learning for Enhanced Autogenous Shrinkage Prediction on Experimental Dataset, *International Conference on Contemporary Computing (IC3 2023)*, August 04, 2023, Noida, Delhi NCR, India.
12. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *Freudenberg Visit*, July 11, 2023, Evanston, IL, USA.
13. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *JARVIS School and CHiMaD Training*, June 15, 2023, Lemont, IL, USA.
14. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *JARVIS School and CHiMaD Training*, June 14, 2023, Evanston, IL, USA.
15. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *7th World Congress on Integrated Computational Materials Engineering (ICME 2023)*, May 22, 2023, Orlando, FL, USA.
16. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *6th Forum of Materials Genome Engineering (ForMGE)*, February 18, 2023, Hangzhou City, Zhejiang Province, China.
17. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Science and Engineering, *13th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2023*, January 20, 2023, Amity University, Noida, Delhi NCR, India.
18. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Science and Engineering, *Accelerated Materials Design and Additive Manufacturing: Scientific and Technological Perspectives (AMDAM), 76th Annual Technical Meeting of The Indian Institute of Metals*, November 13, 2022, Hyderabad, Telangana, India.
19. Artificial Intelligence and High-Performance Data Mining: Overview and Updates, *CHiMaD Annual Meeting*, November 08, 2022, University of Chicago, Chicago, IL, USA.

20. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *3rd International Conference on Advances and Applications of Artificial Intelligence & Machine Learning (ICAAAIML 2022)*, September 16, 2022, Sharda University, Greater Noida, Delhi NCR, India.
21. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *9th International Conference Signal Processing and Integrated Networks(SPIN 2022)*, August 26, 2022, Amity University, Noida, Delhi NCR, India (Virtual).
22. Machine Learning for Materials Science (MLMS), *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2022)*, August 15, 2022, Washington DC, USA.
23. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *3rd Biennial International Conference on Future Learning Aspects of Mechanical Engineering (FLAME 2022)*, August 04, 2022, Amity University, Noida, Delhi NCR, India.
24. AI/ML/DL for Forward and Inverse Problems in Science and Engineering, *International Conference on Dynamical Systems, Control and their Applications (ICDSCA 2022)*, July 01, 2022, Indian Institute of Technology Roorkee (IITR), Roorkee, Uttarakhand, India.
25. (Keynote) AI + HPC for Accelerating Science, *IEEE International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Perspectives and Prospects (COM-IT-Con 2022)*, May 26, 2022, Manav Rachna International Institute of Research and Studies, Faridabad, Delhi NCR, India.
26. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Materials Discovery and Design, *Workshop on Data-driven Modeling and Optimization*, May 19, 2022, Fraunhofer Institute for Mechanics of Materials, IWM, Germany (Virtual).
27. (Invited) AI/ML/DL Approaches for Accelerating Materials Discovery and Design, *TMS 2022: AI/Data Informatics*, February 28, 2022, Anaheim, CA, USA.
28. (Invited) Harnessing the Power of FAIR Materials Data, *MaRDA Annual Meeting*, February 23, 2022 (Virtual).
29. (Invited) AI-Driven Learning of the Science and Engineering of Materials from Simulations, *NAFEMS Seminar on AI, Data Driven Models & Machine Learning*, February 23, 2022 (Virtual).
30. (Keynote) AI + HPC for Accelerating Science, *12th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2022*, January 28, 2022, Amity University, Noida, Delhi NCR, India (Virtual).
31. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *AI for Interdisciplinary Scientific Discovery*, January 27, 2022, Institute of Advanced Studies, Loughborough University, England, UK.
32. Artificial Intelligence and High-Performance Data Mining: Overview and Updates, *CHiMaD Annual Meeting*, January 24, 2022, Northwestern University, USA (Virtual).
33. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *US Army DEVCOM CBC Seminar on AI/ML Applications*, November 10, 2021 (Virtual).
34. (Invited) Introduction to Machine Learning and Deep Learning for Materials Science, *TMS Online Course: Artificial Intelligence in Materials Science and Engineering*, November 02, 2021 (Virtual).
35. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *BiGmax Summer School on Harnessing Big Data in Materials Science from Theory to Experiment*, September 15, 2021, Max Plank Institute, Germany (Virtual).
36. (Invited) AI for Accelerating Materials Discovery and Design, *Nanocombinatorics Workshop on Science of AI*, August 19, 2021, Northwestern University, USA (Virtual).

37. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *ULTRA DOE-EFRC Seminar*, June 28, 2021, Arizona State University, USA (Virtual).
38. (Invited) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *NSF Workshop on Accelerating Materials Discovery, Design, and Synthesis: A Grand Challenge for Artificial Intelligence (AIMS): AI Panel*, April 09, 2021, Pennsylvania State University, USA (Virtual).
39. CHiMaD AI and High Performance Data Mining: Tool Group Updates, *CHiMaD Executive Meeting*, February 12, 2021, Northwestern University, USA (Virtual).
40. (Invited) Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *TMS Webinar Series: Artificial Intelligence in Materials: Research, Design, and Manufacturing*, February 04, 2021, Virtual.
41. (Keynote) Artificial Intelligence and High-Performance Data Mining for Accelerating Scientific Discovery, *11th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2021*, January 29, 2021, Amity University, Noida, Delhi NCR, India (Virtual).
42. (Invited) AI and High-Performance Data Mining: Illustrative Applications in Materials Science, *Indian Symposium on Machine Learning (IndoML)*, December 16, 2020, Indian Institute of Technology (IIT) Gandhinagar, India (Virtual).
43. (Invited) Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *MS&T 2020: Materials Design through AI Composition and Process Optimization*, November 02, 2020 (Virtual).
44. Data-Driven Analytics on Navy Hull Steels Data, *DLA/SFSA Naval Structural Steels Review*, July 24, 2020 (Virtual).
45. Artificial Intelligence and High-Performance Data Mining, *CHiMaD Annual Meeting*, June 09, 2020, Northwestern University, USA (Virtual).
46. (Invited) Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *CHiMaD Phase Field Workshop 2020*, April 21, 2020, Northwestern University, USA (Virtual).
47. (Invited) Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *TMS 2020: ICME Gap Analysis in Materials Informatics*, February 26, 2020, San Diego, CA, USA.
48. (Keynote) High Performance Data Mining: An Essential Paradigm for AI and Big Data Analytics, *10th International Conference on Cloud Computing, Data Science & Engineering (CONFLUENCE) 2020*, January 29, 2020, Amity University, Noida, Delhi NCR, India.
49. Data-driven Analytics for Understanding Materials Properties, *Toyota Visit*, Toyota Motor Corporation, January 23, 2020, Nagoya, Aichi, Japan.
50. A Real-time Iterative Machine Learning Approach for Temperature Profile Prediction in Additive Manufacturing Processes, *6th IEEE International Conference on Data Science and Advanced Analytics (DSAA) 2019*, October 07, 2019, Washington DC, USA.
51. Martensite Start Temperature Predictor for Steels Using Ensemble Data Mining, *6th IEEE International Conference on Data Science and Advanced Analytics (DSAA) 2019*, October 07, 2019, Washington DC, USA.
52. CHiMaD Data Mining and Analytics, *CHiMaD Annual Meeting and Phase II Kickoff*, September 24, 2019, Chicago, IL, USA.
53. (Keynote) Materials Image Informatics Using Deep Learning: Challenges and Opportunities, *LANL Workshop on Advanced Probes and Materials for Enabling 3-D Imaging,* , Los Alamos National Laboratory, August 27, 2019, Santa Fe, NM, USA.

54. (Keynote) High Performance Data Mining: An Essential Paradigm for Big Data Analytics and AI, *5th International Conference on Artificial Intelligence and Security (ICAIS) 2019*, July 27, 2019, Brooklyn, NY, USA.

55. (Invited) Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *CHiMaD Materials Design & Data Informatics Workshop*, July 26, 2019, Evanston, IL, USA.

56. Deep Materials Informatics: Illustrative Applications of Deep Learning in Materials Science, *5th World Congress on Integrated Computational Materials Engineering (ICME 2019)*, July 22, 2019, Indianapolis, IN, USA.

57. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *KAIST Faculty Visit*, May 16, 2019, Evanston, IL, USA.

58. (Invited) Illustrative Microstructure Analytics @CHiMaD, *NIST/NSF Microstructures Workshop*, May 14, 2019, National Cybersecurity Center of Excellence, Rockville, MD, USA.

59. (Invited) CHiMaD Data Analytics, *SRG 2019*, March 25, 2019, Evanston, IL, USA.

60. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *TMS 2019: Computational Approaches for Big Data, Artificial Intelligence and Uncertainty Quantification in Computational Materials Science - Big Data*, March 13, 2019, San Antonio, TX, USA.

61. (Invited) High Performance Data Mining: An Essential Paradigm for Applied Mathematics and Interdisciplinary Big Data Analytics, *RAMSA 2019*, January 18, 2019, Noida, Delhi NCR, India.

62. Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *ACS Midwest Regional Meeting (MWRM 2018)*, October 23, 2018, Ames, IA, USA.

63. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *XXVII International Materials Research Congress (MRS-Mexico)*, August 20, 2018, Cancun, Mexico.

64. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *Artificial Intelligence for Materials Science Workshop (AIMS)*, National Institute of Standards and Technology, August 07, 2018, Gaithersburg, MD, USA.

65. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *13th World Congress on Computational Mechanics (WCCM)*, July 25, 2018, New York City, NY, USA.

66. (Invited) High Performance Data Mining: An Essential Paradigm for Interdisciplinary Big Data Analytics, *Tata Consultancy Services Visit*, TRDDC/TCS, July 18, 2018, Pune, Maharashtra, India.

67. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *Toyota Visit*, Toyota Motor Corporation, June 13, 2018, Toyota-shi, Japan.

68. The investigation of machine learning for material development, *Toyota Visit*, Toyota Motor Corporation, June 13, 2018, Toyota-shi, Japan.

69. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *Machine Learning in Science & Engineering*, Carnegie Mellon University, June 08, 2018, Pittsburgh, PA, USA.

70. Deep Transfer Learning Based Pavement and Structural Health Monitoring , *Machine Learning in Science & Engineering*, Carnegie Mellon University, June 08, 2018, Pittsburgh, PA, USA.

71. Big Data Analytics in Medicine and Healthcare: Analyzing Electronic Healthcare Records, Sequence Data, Social Media, and More, *Machine Learning in Science & Engineering*, Carnegie Mellon University, June 07, 2018, Pittsburgh, PA, USA.

72. CHiMaD Data Mining, *CHiMaD Site-Visit*, April 16, 2018, Evanston, IL, USA.

73. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *Machine Learning in Materials Science Workshop*, University of Utah, April 06, 2018, Salt Lake City, UT, USA.
74. Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *MRS Spring Meeting & Exhibit*, April 04, 2018, Phoenix, AZ, USA.
75. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *6th NU-NIMS Materials Genome Workshop*, March 28, 2018, Evanston, IL, USA.
76. (Invited) CHiMaD Data Mining, *SRG 2018*, March 27, 2018, Evanston, IL, USA.
77. (Invited) Data-Driven Approaches for Steel Fatigue Strength Prediction, *TMS 2018: Fatigue in Materials: Fundamentals, Multiscale Modeling and Prevention - Data-driven Investigations of Fatigue*, March 12, 2018, Phoenix, AZ, USA.
78. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *Materials Science Division Colloquium, Argonne National Lab*, February 15, 2018, Lemont, IL, USA.
79. Deep Learning Models for Structure-Property Linkages in High Contrast Composites, *MURI Final Review*, December 18, 2017, Arlington, VA, USA.
80. Learning Crystal Orientations of Polycrystalline Materials from Electron Backscatter Diffraction Experiments using Convolutional Neural Networks, *MURI Final Review*, December 18, 2017, Arlington, VA, USA.
81. (Invited) Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *International Conference on Nano and Functional Materials (NFM 2017)*, November 17, 2017, BITS Pilani, Rajasthan, India.
82. (Invited) Materials Informatics on Images: Structure Characterization, Crack Detection, Localization, and More, *Machine Learning Applied to Materials Imaging Workshop, Northwestern-Argonne Institute of Science and Engineering*, October 30, 2017, Evanston, IL, USA.
83. (Invited) High Performance Data Mining: An Essential Paradigm for Interdisciplinary Big Data Analytics, *2017 IEEE Region 4 Workshop on Big Data*, October 25, 2017, Evanston, IL, USA.
84. Materials Informatics and Big Data: Realization of "Fourth Paradigm" of Science in Materials Science, *MS&T 2017: Data and Tools for Materials Discovery and Design: Data Science Methods in Materials Discovery and Development*, October 11, 2017, Pittsburgh, PA, USA.
85. Data-Driven Approaches for Predicting Thermoelectric Properties, *MS&T 2017: In-situ Characterization of Energy Materials*, October 10, 2017, Pittsburgh, PA, USA.
86. Classification of Scientific Journal Articles to Support Automated Data Extraction and Curation, *MS&T 2017: Recent Advances in Computer-aided Materials Design: Emerging Approaches of Material Design*, October 10, 2017, Pittsburgh, PA, USA.
87. Data-Driven Approaches for Predicting Fatigue Strength of Steels, *MS&T 2017: Shaping & Forming of Advanced High Strength Steels: Performance*, October 10, 2017, Pittsburgh, PA, USA.
88. Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *Materials Research and Data Science Conference*, September 25, 2017, Rockville, MD, USA.
89. (Invited) Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *MAPEX Symposium 2017, University of Bremen*, September 15, 2017, Bremen, Germany.

90. (Keynote) Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *Materials Genome Symposium, Chinese Materials Conference (CMC 2017), Chinese Materials Research Society (C-MRS)*, July 11, 2017, Yinchuan, Ningxia, China.
91. Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *4th World Congress on Integrated Computational Materials Engineering (ICME 2017), ICME Success Stories and Applications*, May 24, 2017, Ypsilanti, MI, USA.
92. (Invited) Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *SJTU/MaGIC Faculty Visit*, April 27, 2017, Evanston, IL, USA.
93. (Invited) Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science: Steel Fatigue Strength Predictor, *5th NU-NIMS Materials Genome Workshop*, March 28, 2017, Evanston, IL, USA.
94. CHiMaD Data Mining, *CHiMaD Annual Meeting*, March 27, 2017, Evanston, IL, USA.
95. (Invited) CHiMaD Data Mining: Fatigue, *SRG 2017*, March 23, 2017, Evanston, IL, USA.
96. Data Science Approaches for Predicting Fatigue Strength of Steels, *TMS 2017*, February 27, 2017, San Diego, CA, USA.
97. Data Science Approaches for Predicting Thermoelectric Properties, *TMS 2017*, February 27, 2017, San Diego, CA, USA.
98. Parallel Implementation of Lossy Data Compression for Temporal Data Sets, *23rd Annual International Conference on High Performance Computing, Data, and Analytics (HiPC)*, December 20, 2016, Hyderabad, India.
99. Five Year Life Expectancy Calculator for Older Adults, *IEEE International Conference on Data Mining (ICDM)*, December 13, 2016, Barcelona, Spain.
100. A Formation Energy Predictor for Crystalline Materials Using Ensemble Data Mining, *IEEE International Conference on Data Mining (ICDM)*, December 13, 2016, Barcelona, Spain.
101. Predicting the Outcome of Startups: Less Failure, More Success, *IEEE ICDM Workshop on Data Market for Co-evolution of Sciences and Business (MoDAT)*, December 12, 2016, Barcelona, Spain.
102. (Invited) Materials Informatics and Big Data: Realization of the "Fourth Paradigm" of Science in Materials Science, *CHiMaD Summit on Data & Analytics for Materials Research*, November 02, 2016, Evanston, IL, USA.
103. A Fatigue Strength Predictor for Steels Using Ensemble Data Mining, *25th ACM International Conference on Information and Knowledge Management (CIKM)*, October 26, 2016, Indianapolis, IN, USA.
104. (Invited) High Performance Data Mining: An Essential Paradigm for Big Data Analytics and Knowledge Discovery, *3M Visit*, October 19, 2016, Evanston, IL, USA.
105. Identifying HotSpots in Five Year Survival Electronic Health Records of Older Adults, *6th IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*, October 13, 2016, Atlanta, GA, USA.
106. Software tools for sequence comparison, sequence mapping, and patient-specific healthcare outcome prediction, *7th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB)*, October 05, 2016, Seattle, WA, USA.
107. Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *International Conference on Fatigue Damage of Structural Materials (FATD)*, September 22, 2016, Hyannis, MA, USA.
108. CHiMaD Data Mining: An Update, *CHiMaD Executive Meeting*, September 12, 2016, Evanston, IL, USA.

109. Materials Informatics and Big Data: Realization of 4th Paradigm of Science in Materials Science, *Theory and Applications of Computational Chemistry (TACC)*, August 30, 2016, University of Washington, Seattle, WA, USA.

110. Deep Learning Based Big Data Analytics in Materials Science, *MURI Annual Meeting*, August 22, 2016, Caltech, Pasadena, CA, USA.

111. (Invited) Data-driven materials science enabling large-scale property prediction and optimization, *APS/CNM Users Meeting 2016*, May 10, 2016, Argonne National Lab, Chicago, IL, USA.

112. (Invited) High Performance Data Mining: An Essential Paradigm for Big Data Analytics and Knowledge Discovery, *Invited seminar*, May 05, 2016, Golden, CO, USA.

113. (Invited) Big data analytics in medicine and healthcare: Analyzing electronic healthcare records, sequence data, social media, and more, *Outcomes Research Workshop*, April 27, 2016, University of Chicago, Chicago, IL, USA.

114. CHiMaD Data Mining, *CHiMaD Annual Meeting*, March 23, 2016, Evanston, IL, USA.

115. (Invited) CHiMaD Data Mining: Fatigue, *SRG 2016*, March 21, 2016, Evanston, IL, USA.

116. (Invited) High Performance Data Mining: An Essential Paradigm for Big Data Analytics and Knowledge Discovery, *Invited seminar*, March 02, 2016, Auburn, AL, USA.

117. (Invited) Towards Better Efficiency and Accuracy: Data Mining for Prediction and Optimization in Materials System Design, *TMS 2016*, February 16, 2016, Nashville, TN, USA.

118. CHiMaD Data Mining: An Update, *CHiMaD Executive Meeting*, October 19, 2015, Evanston, IL, USA.

119. Parallel Distributed-Memory Based Community Detection for Large Graphs, *DARPA GRAPHS / SIMPLEX Workshop: Data, Algorithms and Problems on Graphs (DAPG) 2015*, September 28, 2015, New York, NY, USA.

120. (Invited) Big Data Analytics and Discovery in Medicine and Healthcare, *NUS Surgical Faculty Visit*, September 18, 2015, Chicago Innovation Exchange, Chicago, IL, USA.

121. Pruned Search: A Machine Learning Based Meta-Heuristic Approach for Constrained Continuous Optimization, *8th International Conference on Contemporary Computing (IC3) 2015*, August 20, 2015, Noida, Delhi NCR, India.

122. All Your Google and Facebook Logins are Belong to Us: A Case for Single Sign-off, *8th International Conference on Contemporary Computing (IC3) 2015*, August 20, 2015, Noida, Delhi NCR, India.

123. (Invited) Towards an Infrastructure for Materials Data Analytics: An Overview of Tools and Platforms for Code Development, Collaboration, and Data Analytics, *Materials Research Collaboration Environment Workshop*, August 13, 2015, Dayton, OH, USA.

124. Application of Machine Learning to Materials Discovery and Development, *MURI 3-Year Review*, June 23, 2015, Arlington, VA, USA.

125. Optimization of Microstructures in Magnetoelastic Alloys, *MURI 3-Year Review*, June 23, 2015, Arlington, VA, USA.

126. Mining of Process-Structure-Property Linkages Using Data Science Tools, *MURI 3-Year Review*, June 23, 2015, Arlington, VA, USA.

127. Towards Better Efficiency and Accuracy: Data Mining for Optimization and Prediction in Materials System Design, *AFOSR Program Review*, May 21, 2015, Arlington, VA, USA.

128. CHiMaD Data Mining, *CHiMaD Annual Meeting*, May 01, 2015, Evanston, IL, USA.

129. Data-driven Analytics for Materials Science: Realization of the Fourth Paradigm of Science, *DARPA SIMPLEX Pre-Kickoff Meeting*, April 14, 2015, Evanston, IL, USA.

130. (Invited) CHiMaD Data Mining: Fatigue, *SRG 2015*, March 23, 2015, Evanston, IL, USA.

131. (Invited) High Performance Data Mining: An Essential Paradigm for Big Data Analytics and Knowledge Discovery, *Invited seminar*, March 09, 2015, Lawrence, KS, USA.
132. CHiMaD Data Mining, *CHiMaD Executive Meeting*, December 01, 2014, Evanston, IL, USA.
133. (Invited) High Performance Big Data Analytics for Data-Driven Discovery in Natural Sciences, *GBMF Symposium*, July 29, 2014, Palo Alto, CA, USA.
134. (Invited) CHiMaD Data Mining, *MTL/SRG 2014*, March 24, 2014, Evanston, IL, USA.
135. High Performance Big Data Clustering, *SDAV All-Hands Meeting*, February 25, 2013, Atlanta, GA, USA.
136. (Invited) An Overview of Essential Concepts in Data Mining, *MURI Program Review*, January 21, 2014, Dayton, OH, USA.
137. Multi-objective Optimization and Multimodal Prediction in the Design of Materials System, *MURI Program Review*, January 21, 2014, Dayton, OH, USA.
138. (Invited) Data-Driven Analytics and Discovery in Medicine and Healthcare, *Research Mela*, November 23, 2013, Rush Hospital, Chicago, IL, USA.
139. An Analysis of Variation in Hospital Billing Using Medicare Data, *KDD Workshop on Data Mining for Healthcare (DMH)*, August 11, 2013, Chicago, IL, USA.
140. Heart Transplant Outcome Prediction using UNOS Data, *KDD Workshop on Data Mining for Healthcare (DMH)*, August 11, 2013, Chicago, IL, USA.
141. Colon Surgery Outcome Prediction Using ACS NSQIP Data, *KDD Workshop on Data Mining for Healthcare (DMH)*, August 11, 2013, Chicago, IL, USA.
142. High Performance Big Data Clustering, *SDAV All-Hands Meeting*, February 20, 2013, San Fransisco, CA, USA.
143. Parallel Hierarchical Clustering on Shared Memory Platforms, *International Conference on High Performance Computing (HiPC)*, December 19, 2012, Pune, Maharashtra, India.
144. Data-driven Analytics and Applications - Realization of the Fourth Paradigm of Science, *MURI Kickoff Meeting*, October 19, 2012, Carnegie Mellon University, Pittsburgh, PA, USA.
145. Supporting Computational Data Model Representation with High-performance I/O in Parallel netCDF, *International Conference on High Performance Computing (HiPC)*, December 19, 2011, Bangalore, Karnataka, India.
146. Enhancing Parallelism of Pairwise Statistical Significance Estimation for Local Sequence Alignment, *2nd HiPC Workshop on Hybrid Multi-Core Computing*, December 18, 2011, Bangalore, Karnataka, India.
147. Identifying HotSpots in Lung Cancer Data Using Association Rule Mining, *2nd IEEE ICDM Workshop on Biological Data Mining and its Applications in Healthcare, BioDM 2011*, December 11, 2011, Vancouver, Canada.
148. Community Dynamics and Analysis of Decadal Trends in Climate Data, *3rd IEEE ICDM Workshop on Knowledge Discovery from Climate Data, ClimKD 2011*, December 11, 2011, Vancouver, Canada.
149. Derived Distribution Points Heuristic for Fast Pairwise Statistical Significance Estimation, *ACM International Conference on Bioinformatics and Computational Biology (ACM-BCB) 2010*, August 03, 2010, Niagara, NY, USA.
150. MPIPairwiseStatSig: Parallel Pairwise Statistical Significance Estimation of Local Sequence Alignment, *HPDC ECMLS 2010*, June 22, 2010, Chicago, IL, USA.
151. Pairwise Statistical Significance of Local Sequence Alignment Using Substitution Matrices with Sequence-Pair-Specific Distance, *IEEE International Conference on Information Technology (ICIT 2008)*, December 20, 2008, Bhubaneswar, Orissa, India.

152. Pairwise DNA Alignment with Sequence Specific Transition-Transversion Ratio Using Multiple Parameter Sets, *IEEE International Conference on Information Technology (ICIT 2008)*, December 20, 2008, Bhubaneswar, Orissa, India.

153. Conservative, Non-Conservative and Average Pairwise Statistical Significance of Local Sequence Alignment, *IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2008)*, November 5, 2008, Philadelphia, PA, USA.

154. Pairwise Statistical Significance of Local Sequence Alignment Using Multiple Parameter Sets, *2nd International Workshop on Data and Text Mining in Bioinformatics (DTMBIO 2008)*, October 30, 2008, Napa Valley, CA, USA.

155. DNAlignTT: Pairwise DNA Alignment with Sequence Specific Transition-Transversion Ratio, *IEEE International Conference on EIT 2008*, May 21, 2008, Ames, IA, USA.

156. Pairwise Statistical Significance Versus Database Statistical Significance for Local Alignment of Protein Sequences, *International Symposium on Bioinformatics Research and Applications (ISBRA 2008)*, May 7, 2008, Atlanta, GA, USA.

157. Estimating Pairwise Statistical Significance of Protein Local Alignments Using a Clustering-Classification Approach Based on Amino Acid Composition, *International Symposium on Bioinformatics Research and Applications (ISBRA 2008)*, May 7, 2008, Atlanta, GA, USA.

158. Identifying Temporal Gene Networks Using Signal Processing Metrics on Time-Series Gene Expression Data, *IEEE Third International Conference on Intelligent Sensing and Information Processing*, December 14, 2005, Bangalore, Karnataka, India.

159. Identifying Temporal Gene Networks by Mining Gene Expression Data, *IEEE 12th International Conference on Advanced Computing and Communications (ADCOM) 2004*, December 17, 2004, Ahmedabad, Gujarat, India.

Software

1. Parallel Data Clustering Algorithms: Software for parallel DBSCAN clustering in MPI and OpenMP, along with sample dataset is available at http://cucis.ece.northwestern.edu/projects/Clustering

2. Pruned Search Optimization: A general optimization package that performs search space reduction given an objective function and mathematical constraints, available at http://cucis.ece.northwestern.edu/projects/MURI/workopt.html

3. Parallel MEP: A parallel community detection algorithm for graphs, based on idea of maximizing equilibrium and purity of communities, available at http://cucis.ece.northwestern.edu/projects/PMEP/

4. Steel Fatigue Strength Predictor: The online steel fatigue strength predictor is available at http://info.eecs.northwestern.edu/SteelFatigueStrengthPredictor

5. Formation Energy Predictor: The online formation energy predictor is available at http://info.eecs.northwestern.edu/FEpredictor

6. ThermoEl Toolkit: The online thermoelectric toolkit for predicting thermoelectric properties is available at http://info.eecs.northwestern.edu/ThermoEl

7. Seebeck Coefficient Predictor: The online Seebeck coefficient predictor is available as part of the ThermoEl toolkit at http://info.eecs.northwestern.edu/SeebeckCoefficientPredictor

8. Bulk Modulus Predictor: The online bulk modulus predictor is available as part of the ThermoEl toolkit at http://info.eecs.northwestern.edu/BulkModulusPredictor

9. Martensite Start Temperature (MsT) Predictor: The online MsT predictor is available at http://info.eecs.northwestern.edu/MsTpredictor

10. Organic Photovoltaic Predictor: The online OPV predictor is available at http://info.eecs.northwestern.edu/OPVPredictor

11. Macroscale (Effective) Stiffness Predictor: The macroscale (effective) stiffness predictor of high-contrast two-phase three-dimensional composite materials is available at http://cucis.ece.northwestern.edu/projects/MURI/workpred.html
12. Stress-Strain Curve Predictor: The online stress-strain curve predictor is available at http://info.eecs.northwestern.edu/StressStrainCurvePredictor
13. Materials Property Predictor: The online materials property predictor is available at http://ai.eecs.northwestern.edu/MPpredictor
14. Five Year Life Expectancy Calculator: The online five year life expectancy calculator for older adults is available at http://info.eecs.northwestern.edu/FiveYearLifeExpectancyCalculator
15. Lung Cancer Outcome Calculator: The online lung cancer outcome calculator is available at http://info.eecs.northwestern.edu/LungCancerOutcomeCalculator
16. Colon Cancer Outcome Calculator: The online colon cancer outcome calculator is available at http://info.eecs.northwestern.edu/ColonCancerOutcomeCalculator
17. Analyzing the Variation in Hospital Billing using Medicare Data: Visually depicted results in the form of US state heat maps can be found at http://users.eecs.northwestern.edu/~ankitag/hospitalbilling/
18. Sentiment Analysis for Social Media Data: An API for sentiment analysis service and benchmark data is available at http://cucis.ece.northwestern.edu/projects/Social/sentiment_api.html
19. Real-Time Digital Flu Surveillance: The real-time disease surveillance tool for flu is available at http://pulse.eecs.northwestern.edu/~kml649/flu/
20. Real-Time Digital Cancer Surveillance: The real-time disease surveillance tool for cancer is available at http://pulse.eecs.northwestern.edu/~kml649/cancer/
21. Real-Time Digital Allergy Surveillance: The real-time disease surveillance tool for allergy is available at http://pulse.eecs.northwestern.edu/~kml649/allergy/
22. Pairwise Statistical Significance Estimation: Sequential codes for pairwise statistical significance estimation and its variants are available at http://www.cs.iastate.edu/~ankitag/projects.html. Parallel codes for optimized HPC implementations (on MPI, OpenMP, Hybrid, GPU) of pairwise statistical significance estimation are available at http://cucis.ece.northwestern.edu/projects/PSSE/
23. AGILE: Long Read Sequence Mapping: Software and data for AGILE are available at http://cucis.ece.northwestern.edu/projects/NGS/agile.html
24. DNAlignTT: DNA Alignment with Sequence-Specific Transition-Transversion Ratio: Source code for DNAlignTT is available at http://www.cs.iastate.edu/~ankitag/DNAlignTT.html
25. FATBEP: Fuzzy Adaptive Thresholding Based Exon Predictor: An implementation of the approach as a user friendly GUI in MATLAB is available at http://www.cs.iastate.edu/~ankitag/FATBEP.html

Research Supervision

Postdoctoral Fellows and/or Research Associates

- Jagat Sesh Challa (2020-2021)
- Reda Al-Bahrani (2018-2019)
- Dianwei Han (2015-2017)
- Al'ona Furmanchuk (2015-2016)
- Lu Liu (2013-2014)
- Zhengzhang Chen (2012-2014)
- Mostofa Patwary (2011-2013)
- William Hendrix (2011-2012)

Graduate Students (as Advisor/Co-advisor/Preceptor/Committee Co-Chair)

- Sayak Chakrabarty, PhD (2024-present)
- Youjia Li, PhD (2023-present)
- Muhammed Nur Talha Kilic, PhD (2022-present)
- Alec Peltekian, PhD (2022-present)
- Alexandra Day, Dissertation Title: "Development of Artificial Intelligence Techniques for Automated Nanoparticle Characterization" PhD (2020-2025)
- Yuwei Mao, PhD, Dissertation Title: "Artificial Intelligence Methodologies for Prediction and Optimization Problems in Materials Informatics" (2020-2025)
- Kewei Wang, PhD, Dissertation Title: "Scalable Training Strategy for Parallel Neural Network and Interpretable Materials Modeling" (2019-2025)
- Vishu Gupta, PhD, Dissertation Title: "Deep Learning Methodologies for Limited Resources Scenarios in Materials Informatics" (2019-2023)
- Zijiang Yang, PhD, Dissertation Title: "Deep Learning for Computer Vision and Applications in Materials Informatics" (2016-2021)
- Dipendra Jha, PhD, Dissertation Title: "Deep Learning Methodologies for Scientific Knowledge Discovery" (2016-2019)
- Arindam Paul, PhD, Dissertation Title: "Machine Learning and Data-driven Optimization for Applications in Scientific Discovery" (2015-2019)
- Reda Al-Bahrani, PhD, Dissertation Title: "Knowledge Discovery for Health Informatics from Structured and Textual Data" (2012-2018)
- Kathy Lee, PhD, Disseration Title: "Mining Social Media for Healthcare Intelligence" (2010-2017)
- Ruoqian (Rosanne) Liu, PhD, Dissertation Title: "Multi-Contextual Representation and Learning with Applications in Materials Knowledge Discovery" (2012-2016)
- Yusheng (Yves) Xie, PhD, Dissertation Title: "Recommender Systems for Social Networks" (2011-2015)
- Diana Palsetia, PhD, Dissertation Title: "High Quality, Scalable Community Detection for Large Graphs" (2010-2015)

Graduate Students (Other)

- Zanhua Huang, PhD (2023-2024)
- Claire Lee, PhD, Dissertation Title: "Data Management Challenges for Scaling Graph Neural Networks in Scientific Workflows" (2019-2024)
- Kai Yuan Hou, PhD, Dissertation Title: "Scalable Parallel I/O in the Exa-scale Era" (2016-2022)
- Qiao Kang, PhD, Dissertation Title: "Scalable Communication and I/O Algorithms for High Performance Computing Systems" (2015-2020)
- Sunwoo Lee, PhD, Dissertation Title: "Scalable Parallelization Strategy for Large-Scale Deep Learning" (2015-2020)
- Esteban (Steve) M. Rangel, PhD, Dissertation Title: "High Performance Algorithms for Data Analysis in Computational Cosmology" (2012-2018)
- Chen Jin, PhD (2013-2018)
- Amar Krishna, MS (2013-2016)
- Sharifah Ummu Kulthumm, MS (2015)
- Zheng (Dobbie) Yuan, PhD (2015-2016)
- Yu Cheng, PhD, Dissertation Title: "Mining and Understanding Professional Social Network: Challenges and Solutions" (2010-2015)
- Yuhong Zhang, PhD (2009-2011)
- Daniel Honbo, PhD (2009-2014)
- Kunpeng Zhang, PhD, Dissertation Title: "Big Social Media Data Mining for Marketing Intelligence" (2010-2013)
- Sanchit Misra, PhD, Dissertation Title: "High Throughput Sequence Mapping for Next Generation DNA Sequencing" (2009-2011)

Undergraduate Students

- Justin Liao (2018-2018)
- Pranjal Daga (2015-2015)
- Siddharth Gupta (2013-2013)
- Lalith Polepeddi (2009-2010)

Graduate Student Exam Committees

- Kevin Yao (Chemical and Biological Engineering Major), PhD Research Proposal (2024)
- Sayak Chakrabarty, PhD Qualifying Examination (2024)
- Alexandra Day, PhD Research Proposal (2024), PhD Thesis Defense Exam (2025)
- Yuwei Mao, PhD Research Proposal (2023), PhD Thesis Defense Exam (2024)
- Kewei Wang, PhD Research Proposal (2023), PhD Thesis Defense Exam (2024)
- Claire Lee, PhD Research Proposal (2023), PhD Thesis Defense Exam (2024)
- Vishu Gupta, PhD Research Proposal (2022), PhD Thesis Defense Exam (2023)
- Kai Yuan Hou, PhD Research Proposal (2020), PhD Thesis Defense Exam (2022)
- Zijiang Yang, PhD Research Proposal (2019), PhD Thesis Defense Exam (2021)
- Yu-Chin Chan (Mechanical Engineering Major), PhD Qualifying Examination (2019), PhD Thesis Defense Exam (2022)
- Wenyu Han (Mechanical Engineering Major), MS Thesis Defense Exam (2019)
- Qiao Kang, PhD Research Proposal (2018), PhD Thesis Defense Exam (2020)
- Sunwoo Lee, PhD Research Proposal (2018), PhD Thesis Defense Exam (2020)
- Dipendra Jha, PhD Research Proposal (2017), PhD Thesis Defense Exam (2019)
- Arindam Paul, PhD Research Proposal (2017), PhD Thesis Defense Exam (2019)
- Xiaolin Li (Mechanical Engineering Major), PhD Qualifying Examination (2016), PhD Thesis Defense Exam (2018)
- Reda Al-Bahrani, PhD Qualifying Examination (2015), PhD Research Proposal (2017), PhD Thesis Defense Exam (2018)
- Ruoqian (Rosanne) Liu, PhD Research Proposal (2015), PhD Thesis Defense Exam (2016)
- Yusheng (Yves) Xie, PhD Research Proposal (2014), PhD Thesis Defense Exam (2015)
- Yu Cheng, PhD Research Proposal (2014), PhD Thesis Defense Exam (2015)
- Esteban (Steve) M. Rangel, PhD Qualifying Examination (2014), PhD Research Proposal (2015, 2016), PhD Thesis Defense Exam (2018)
- Diana Palsetia, PhD Research Proposal (2014), PhD Thesis Defense Exam (2015)
- Chen Jin, PhD Research Proposal (2013)
- Kathy Lee, PhD Qualifying Examination (2013), PhD Research Proposal (2014), PhD Thesis Defense Exam (2017)
- Kunpeng Zhang, PhD Qualifying Examination (2012), PhD Research Proposal (2012), PhD Thesis Defense Exam (2013)
- Daniel Honbo, PhD Research Proposal (2011), PhD Thesis Defense Exam (2014)
- Sanchit Misra, PhD Research Proposal (2011), PhD Thesis Defense Exam (2011)

Professional Service

Editorial board and organizing committees

- Editor-in-Chief, Computers, Materials & Continua, 2019-present
- Specialty Chief Editor, Frontiers in High Performance Computing: HPC Applications, 2022-present
- Guest Editor, MRS Communications, Special Issue on Artificial Intelligence, 2019
- Guest Editor, Integrating Materials and Manufacturing Innovation (IMMI), Special Issue on "Using Digital Data in Materials Science and Engineering", 2013
- Editorial Oversight Committee Member, Integrating Materials and Manufacturing Innovation (IMMI), 2020-present
- Editorial Board, Integrating Materials and Manufacturing Innovation (IMMI), 2022-present
- Editorial Board, Scientific Reports, 2023-present
- Editorial Board, BMC/Springer Big Data Analytics, 2016-2020

- Organizing Committee Member, First Workshop on Machine Learning for Materials Science (MLMS) in conjunction with 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, August 15, 2022
- Organizing Committee Member, Data & Analytics for Materials Research Summit, hosted by the National Institute of Standards and Technology and its Center of Excellence for Hierarchical Materials Design (CHiMaD), October 31 - November 2, 2016.
- Advisory Committee Member, Materials Data Analytics: A Path-Finding Workshop, hosted by ASM International's CMD Network and OSU with support from NIST, October 8-9, 2015.
- Session Chair, TMS 2019, 2016

Invited program committee member

- IEEE International Conference on Data Mining (ICDM) 2025, 2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014
- IEEE International Conference on Big Data (BigData) 2025, 2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014, 2013
- ACM International Conference on Information and Knowledge Management (CIKM) 2022, 2021, 2020, 2019, 2013, 2012
- IEEE International Conference on Machine Learning and Applications (ICMLA) 2022, 2021
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2016
- IEEE International Conference on Machine Learning and Data Science (ICMLDS) 2017
- ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC) 2017
- IEEE International Parallel & Distributed Processing Symposium (IPDPS) 2016, 2015
- IEEE IPDPS Workshop on Parallel and Distributed Processing for Computational Social Systems (ParSocial) 2022, 2021, 2020, 2018, 2017, 2016
- IEEE ICDM Workshop on Data Mining in Networks (DaMNet) 2019, 2017, 2016, 2015, 2014, 2013, 2012
- ACM/IEEE SC Workshop on Big Data Analytics: Challenges and Opportunities (BDAC) 2015
- IEEE International Conference on Systems, Man, Cybernetics (SMC) 2015
- IEEE International Conference on Big Data Science and Engineering (BDSE) 2014, 2013

Invited grant proposal reviewer/panelist

- US National Science Foundation (NSF)
- US Department of Energy (DOE)
- US Air Force Office of Scientific Research (AFOSR)
- US Army Research Office (ARO)
- Northwestern Data Science Initiative (DSI)
- Singapore Agency for Science, Technology and Research (A*STAR)
- European Research Council (ERC)
- Dutch Research Council (NWO)
- Austrian Science Fund (FWF)
- Mitacs Canada

Invited paper reviewer

- ACM Computing Surveys
- ACM Transactions on Knowledge Discovery from Data (TKDD)
- ACS Combinatorial Science
- ARC American Institute of Aeronautics and Astronautics Journal (AIAA)
- BMC Big Data Analytics
- BMC Bioinformatics
- BMC Medical Informatics and Decision Making
- BMC Research Notes

- Cambridge MRS Advances
- Cambridge MRS Communications
- Elsevier Acta Materialia
- Elsevier Additive Manufacturing (AM)
- Elsevier Calphad (Computer Coupling of Phase Diagrams and Thermochemistry)
- Elsevier Computers in Biology and Medicine (CBM)
- Elsevier Computer Methods and Programs in Biomedicine (CMPB)
- Elsevier Computer Communications (COMCOM)
- Elsevier Computational Materials Science (COMMAT)
- Elsevier Current Opinion in Solid State & Materials Science (COSSMS)
- Elsevier Composites Science and Technology (CSTE)
- Elsevier Data and Knowledge Engineering (DKE)
- Elsevier Digital Signal Processing (DSP)
- Elsevier Information Sciences
- Elsevier Journal of Computational Science (JOCS)
- Elsevier Journal of Parallel and Distributed Computing (JPDC)
- Elsevier Materials Science and Engineering: A (MSEA)
- Elsevier Neural Networks (NEUNET)
- Hindawi BioMed Research International
- Hindawi Scientifica
- IEEE Transactions on Computational Biology and Bioinformatics (TCBB)
- IEEE Transactions on Emerging Topics in Computing (TETC)
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Parallel and Distributed Systems (TPDS)
- IEEE Transactions on Very Large Scale Integration Systems (TVLSI)
- IET Healthcare Technology Letters
- Inderscience International Journal of Computer Aided Engineering and Technology (IJCAET)
- Inderscience International Journal of High Performance Computing and Networking (IJHPCN)
- MDPI Algorithms
- MDPI Crystals
- MDPI Entropy
- Nature Communications
- npj Computational Materials
- OxfordUP Bioinformatics
- PLOS One
- Springer Artificial Intelligence Review
- Springer Data Mining and Knowledge Discovery (DMKD)
- Springer Integrating Materials and Manufacturing Innovation (IMMI)
- Springer International Journal of Precision Engineering and Manufacturing (IJPEM)
- Springer Journal of Materials Science (JMSC)
- Springer Language Resources and Evaluation (LREV)
- T&F International Journal of Parallel, Emergent and Distributed Systems (IJPEDS)
- T&F Journal of Civil Engineering and Management (JCEM)
- T&F Transport
- Wiley Concurrency and Computation: Practice and Experience (CCPE)
- World Scientific Journal of Circuits, Systems, and Computers (JCSC)
- SC 2013
- ICDM 2012
- ICDM DaMNet 2011
- IPDPS 2011
- PAKDD 2010

References

Available upon request