Yi He

UPDATE: **OCT 2026**
Email: yihe@wm.edu
Address: 540 Landrum Dr, ISC 3385
Williamsburg, VA 23185
Tel: (337)-255-3817

APPOINTMENT

Assistant Professor (Tenure-Track) July, 2024 – present
Dept. of Data Science, College of William and Mary, Williamsburg, Virginia, USA

Director of Research Engagement & Outreach October, 2026 – present
Dept. of Data Science, College of William and Mary, Williamsburg, Virginia, USA

Assistant Professor (Tenure-Track) July, 2021 – July 2024
Dept. of Computer Science, Old Dominion University, Norfolk, Virginia, USA

EDUCATION

Ph.D. in Computer Science Jan. 2016–Dec. 2020
University of Louisiana at Lafayette, Lafayette, Louisiana, USA

B.E. in Transportation Engineering Sept. 2009–June 2013
Harbin Institute of Technology, Harbin, Heilongjiang Province, China

RESEARCH INTERESTS

Data Mining, Data Streams, Big Data Analytics, Real-Time Decision-Making, Graph Learning, Stochastic Models

SPONSORED PROJECTS AND GRANTS

NSF CAREER: Harnessing Graph Dynamics for Learning in Changing Environments
Grant #: IIS-2544353
Role: Principal Investigator (**Sole PI**)
Amount: $\$499,928$ Duration: 07/15/2026 – 06/30/2031
Sponsor: National Science Foundation

NOAA AI-Enhanced Processing of HabCam Imagery to Improve Atlantic Sea Scallop Survey Products for Stock Assessment
Grant #: NA26NMFX063G0010-T1-01
Role: Principal Investigator (**W&M PI**)
collaborating with VIMS Lead-PI: Ming Sun and PI: Roger Mann
Total Amount: $\$1,713,701$ W&M Share: $\$158,279$ Duration: 09/01/2026 – 08/31/2028
Sponsor: National Oceanic and Atmospheric Administration

NSF Collaborative Research: PACSP TOOLS: EPICS: Explainable AI Driven Individual Photo-Identification and Tracking for Cost-effective Conservation Study
Grant #: IOS-2446522
Role: Principal Investigator (**W&M PI**)
collaborating with FAU Lead-PI: Xingquan Zhu, Mote Lab PI: Catherine Walsh
Total Amount: $\$1,299,710$ W&M Share: $\$269,740$ Duration: 10/01/2024 – 09/30/2029
Sponsor: National Science Foundation & Allen Family Philanthropies

NSF Collaborative Research: III: SMALL: Taming Large-Scale Streaming Graphs in an Open World

Grant #: IIS-2505719, IIS-2236579
Role: Principal Investigator (**Lead PI**)
collaborating with UAZ Co-PI: Hongyi Wu, FAU PI: Xingquan Zhu, FAU Co-PI: Laurent Cherubin
Total Amount: $600,000 W&M Share: $281,308 Duration: 07/15/2023 – 06/30/2026
Sponsor: National Science Foundation

NSF CRII: III: Pursuing Interpretability in Utilitarian Online Learning Models
Grant #: IIS-2441449
Role: Principal Investigator (**Sole PI**)
Amount: $175,000 Duration: 07/01/2023 – 06/30/2025
Sponsor: National Science Foundation

NSF Travel: NSF Student Travel Support for the 2024 IEEE International Conference on Data Mining (IEEE ICDM 2024)
Grant #: IIS-2440725
Role: Principal Investigator (**Sole PI**)
Amount: $25,000 Duration: 08/15/2024 – 07/31/2025
Sponsor: National Science Foundation

SCEMFIS AI-powered Underwater Optical Survey for Benthic Biomass and Sediments Assessment
Role: Principal Investigator (**PI**)
Amount: $30,000 Duration: 04/01/2025 – 12/31/2025
Sponsor: Science Center for Marine Fisheries

CCI Secure AI for Marine Ecosystems (SAME): Modeling Benthic Biomass with Image-Graph Fusion under Data Poisoning Threats
Role: Principal Investigator (**PI**)
Amount: $99,948 Duration: 06/02/2025 – 06/30/2026
Sponsor: Commonwealth Cyber Initiative

CCI Mitigating Ethical AI Threats: Dynamic Benchmarks for Securing Multimodal Social Intelligence in Large Language Models
Role: Principal Investigator (**PI**)
Amount: $15,000 Duration: 03/01/2025 – 02/28/2026
Sponsor: Commonwealth Cyber Initiative

CCI Spotlighting and Mitigating Cyber-Attacks in AIoT-enabled Maritime Transportation Systems
Role: Principal Investigator (**PI**)
Amount: $100,000 Duration: 02/01/2022 – 09/30/2024
Sponsor: Commonwealth Cyber Initiative

PUBLICATIONS

(✉ indicates corresponding authorship; * indicates my current or former student advisees and mentees.)

Peer-Reviewed Conference Papers

IEEE BigData'26 Zhiqiang Wang, Sheri L. Barton, Jennifer H. Johnson, **Yi He**, and Xingquan Zhu, "Augmentation with Reinforcement Transformation for Low-textured Scientific Sketch Image Retrieval", accepted by *IEEE International Conference on Big Data*, Phoenix, Arizona, USA, December 14–17, 2026.

IEEE BigData'26 Sophia Sultan*, Heng Lian, Roger Mann, and **Yi He** ✉, "U-Net-Guided Detection and Measurement of Growth Bands in Atlantic Surfclam Shell Images", accepted by *IEEE International Conference on Big Data*, Phoenix, Arizona, USA, December 14–17, 2026.

NeurIPS'26 Haoyang Jiang*, Yuheng Li, Di Yang, Yanhai Xiong, Haipeng Chen, and **Yi He** ✉, "VTV-FM: Flow Matching through Variational Terminal-Velocity Closure", accepted by *the 40th Annual Conference on Neural Information Processing Systems*, Atlanta, Georgia, USA, December 9–11, 2026.

NeurIPS'26 Yufei Jin, Jiajun Shen, **Yi He**, and Xingquan Zhu, "Graph Learning from Label Proportions using Topology-Aware Pseudo Labeling and Aggregation", accepted by *the 40th Annual Conference on Neural Information Processing Systems*, Sydney, Australia, December 8–10, 2026.

NeurIPS'26 Jiheng Liang*, Chen Zhao, Di Wu, Chenyang Bu, Yunpeng Hong*, Xingquan Zhu, and **Yi He** ✉, "ColdDDI: Evaluating Knowledge Utilization in Cold-Start Drug-Drug Interaction Prediction", accepted by *the 40th Annual Conference on Neural Information Processing Systems*, Atlanta, Georgia, USA, December 9–11, 2026.

ICDM'26 Haoyang Jiang*, Zhengui Wang, Shenghan Gao, Y. Joseph Zhang, Xingquan Zhu, and **Yi He** ✉, "Physics-Refined Spatiotemporal Forecasting on Open-Boundary Hydrologic Graphs", accepted by *the 26th IEEE International Conference on Data Mining*, Shenyang, Liaoning, China, November 12–15, 2026. ***Best Paper Award***

ICDM'26 Yang Yan, Chen Zhao, **Yi He**, Zhenyi Wang, Hao Gao, and Zhong Chen, "An ensemble-based network architecture search framework for 3D dose prediction in radiotherapy", accepted by *the 26th IEEE International Conference on Data Mining*, Shenyang, Liaoning, China, November 12–15, 2026.

SDM'26 Kaibu Feng*, Heitor M. Gomes, Xingquan Zhu, and **Yi He** ✉, "Online Decision Trees for Varying Feature Spaces via Gradient-Based Surrogates", accepted by *the 2026 SIAM International Conference on Data Mining*, Salt Lake City, Utah, USA, November 19–20, 2026.

KDD'26 Yulin Wang, **Yi He**, Dianlong You, and Di Wu, "Online Learning for Wild Streaming Data with Delayed Feedback", published in *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, pp. 5196–5207, Jeju, Republic of Korea, August 9–13, 2026. [DOI]

KDD'26 Yunpeng Hong*, Chenyang Bu, Di Wu, **Yi He**, and Xindong Wu, "Implicit Fine-tuning via Context Engineering: A Curriculum Learning Framework for Multimodal Entity Alignment", published in *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, pp. 1698–1709, Jeju, Republic of Korea, August 9–13, 2026. [DOI]

KDD'26 Yunpeng Hong*, Chenyang Bu, Jie Zhang, **Yi He**, Di Wu, and Xindong Wu, "PSQE: A Theoretical-Practical Approach to Pseudo Seed Quality Enhancement for Unsupervised Multimodal Entity Alignment", published in *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*, pp. 464–475, Jeju, Republic of Korea, August 9–13, 2026. [DOI]

WWW'26 Chengjun Yu, Di Wu, **Yi He**, Jia Chen, and Xin Luo, "Federated Latent Factor Learning for Privacy-Preserving Spatio-Temporal Signal Recovery", published in *Proceedings of the ACM Web Conference 2026*, pp. 2905–2916, Dubai, United Arab Emirates, June 29–July 3, 2026. [DOI]

SIGMOD'26 Di Wu, Zetong Tang, **Yi He**, and Xin Luo, "SchemaRAG: A Schema-aware Retrieval-Augmented Generation Framework for Text-to-SQL", published in *Proceedings of the ACM on Management of Data*, Vol. 4, No. 1 (SIGMOD 2026), Article 82, pp. 82:1–82:26, 2026. [DOI]

AAAI'26 Dong Li, Zhengzhang Chen, Xujiang Zhao, Linlin Yu, Zhong Chen, **Yi He**, Haifeng Chen, and Chen Zhao, "MARLIN: Multi-Agent Reinforcement Learning for Incremental DAG Discovery", published in *Proceedings of the 40th AAAI Conference on Artificial Intelligence*, Vol. 40, No. 27, pp. 22869–22877, Singapore, January 20–27, 2026. [DOI]

AAAI'26 Manzong Huang*, Chenyang Bu, **Yi He**, Xingrui Zhuo, and Xindong Wu, "Relink: Constructing Query-Driven Evidence Graph On-the-Fly for GraphRAG", published in *Proceedings of the 40th AAAI Conference on Artificial Intelligence*, Vol. 40, No. 37, pp. 31202–31210, Singapore, January 20–27, 2026. [DOI]

ICDM'25 Xiulin Zheng, Pei-Pei Li, **Yi He**, and Xindong Wu, "Link Completion for Dynamic Knowledge Graphs with Entity Drift", published in *Proceedings of the 25th IEEE International Conference on Data Mining*, pp. 1722–1731, Washington, DC, USA, November 12–15, 2025. [DOI]

ICDM'25 Jiajun Shen, Yufei Jin, **Yi He**, and Xingquan Zhu, "LHGEL: Large Heterogeneous Graph Ensemble Learning using Batch View Aggregation", published in *Proceedings of the 25th IEEE International Conference on Data Mining*, pp. 713–722, Washington, DC, USA, November 12–15, 2025. [DOI]

ICML'25 Haoyang Jiang*, Jindong Wang, Xingquan Zhu, and **Yi He** ✉, "Topology-aware Neural Flux Prediction Guided by Physics", published in *Proceedings of the 42nd International Conference on Machine Learning*, PMLR Vol. 267, pp. 27670–27684, Vancouver, Canada, July 13–19, 2025. [Paper]

IJCAI'25 Jiajun Shen, Yufei Jin, Kaibu Feng, **Yi He**, and Xingquan Zhu, "HGEN: Heterogeneous Graph Ensemble Networks", published in *Proceedings of the 34th International Joint Conference on Artificial Intelligence*, pp. 6156–6163, Montreal, Canada, August 16–22, 2025. [DOI]

IJCAI'25 Manzong Huang*, Chenyang Bu, **Yi He**, and Xindong Wu, "How to Mitigate Information Loss in Knowledge Graphs for GraphRAG: Leveraging Triple Context Restoration and Query-Driven Feedback", published in *Proceedings of the 34th International Joint Conference on Artificial Intelligence*, pp. 8104–8112, Montreal, Canada, August 16–22, 2025. [DOI]

IJCAI'25 Zhong Chen, **Yi He** ✉, Di Wu, Chen Zhao, and Meikang Qiu, "A Novel Sparse Active Online Learning Framework for Fast and Accurate Streaming Anomaly Detection Over Data Streams", published in *Proceedings of the 34th International Joint Conference on Artificial Intelligence*, pp. 2740–2748, Montreal, Canada, August 16–22, 2025. [DOI]

ACL'25 Chenyang Bu, Guojie Chang, Zihao Chen, CunYuan Dang, Zhize Wu, **Yi He**, and Xindong Wu, "Query-Driven Multimodal GraphRAG: Dynamic Local Knowledge Graph Construction for Online Reasoning", published in *Findings of the Association for Computational Linguistics: ACL 2025*, pp. 21360–21380, Vienna, Austria, July 27–August 1, 2025. [DOI]

AAAI'25 Heng Lian*, Chen Zhao, Zhong Chen, Xingquan Zhu, My T. Thai, and **Yi He** ✉, "Metric-Agnostic Continual Learning for Sustainable Group Fairness", published in *Proceedings of the 39th AAAI Conference on Artificial Intelligence*, Vol. 39, No. 18, pp. 18648–18657, Philadelphia, PA, USA, February 25–March 4, 2025. [DOI]

SDM'25 Zhong Chen, **Yi He**, Di Wu, Wenbin Zhang and Zhiqiang Deng, "$\ell_{1,\infty}$-Mixed Norm Promoted Row Sparsity for Fast Online CUR Decomposition Learning in Varying Feature Spaces", published in *Proceedings of the 2025 SIAM International Conference on Data Mining*, pp. 124–133, Alexandria, VA, USA, May 1–3, 2025. [DOI]

CSCWD'25 Shihui Li, Di Wu, **Yi He** and Xin Luo, "Federated Projection Hamming Hash Latent Factor Learning Model", published in *Proceedings of the 28th International Conference on Computer Supported Cooperative Work in Design*, pp. 467–472, Compiègne, France, May 5–7, 2025. [DOI]

BigData'24 Zhong Chen, **Yi He**, Di Wu, Liudong Zuo, Keren Li, Wenbin Zhang, and Zhiqiang Deng, "$\ell_{1,2}$-Norm and CUR Decomposition based Sparse Online Active Learning for Data Streams with Streaming Features", published in *Proceedings of the 2024 IEEE International Conference on Big Data*, pp. 384–393, Washington, DC, USA, December 15–18, 2024. [DOI]

NeurIPS'24 Yufei Jin, Heng Lian*, **Yi He**, and Xingquan Zhu, "HGDL: Heterogeneous Graph Label Distribution Learning", published in *Advances in Neural Information Processing Systems*, Vol. 37, pp. 40792–40830, Vancouver, Canada, December 10–15, 2024. [DOI]

ICDM'24 Heng Lian*, Yu Huang, Xingquan Zhu, and **Yi He** ✉, "Utilitarian Online Learning from Open-World Soft Sensing", published in *Proceedings of the 2024 IEEE International Conference on Data Mining*, pp. 241–250, Abu Dhabi, United Arab Emirates, December 9–12, 2024. [DOI]

SMC'24 Leizhen Zhang*, Lusi Li, Di Wu, Sheng Chen, and **Yi He** ✉, "Fairness-Aware Streaming Feature Selection with Causal Graphs", published in *Proceedings of the 2024 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 858–865, Kuching, Sarawak, Malaysia, October 6–10, 2024. [DOI]

ECOOP'24 Mohammad Wahiduzzaman Khan, Sheng Chen, and **Yi He**, "Learning Gradual Typing Performance", published in *Proceedings of the 38th European Conference on Object-Oriented Programming*, LIPIcs Vol. 313, Article 21, pp. 21:1–21:27, Vienna, Austria, September 16–20, 2024. [DOI]

AAAI'24 Yufei Jin, Richard Gao, **Yi He**, and Xingquan Zhu, "GLDL: Graph Label Distribution Learning", published in *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, Vol. 38, No. 11, pp. 12965–12974, Vancouver, Canada, February 20–27, 2024. [DOI]

AAAI'24 Di Wu, Wu Sun, **Yi He** ✉, Zhong Chen, and Xin Luo, "MKG-FENN: A Multimodal Knowledge Graph Fused End-to-end Neural Network for Accurate Drug-Drug Interaction Prediction", published in *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, Vol. 38, No. 9, pp. 10216–10224, Vancouver, Canada, February 20–27, 2024. [DOI]

SDM'24 Zhong Chen, **Yi He**, Di Wu, Huixin Zhan, Victor S. Sheng, and Kun Zhang, "Robust Sparse Online Learning for Data Streams with Streaming Features", published in *Proceedings of the 2024 SIAM International Conference on Data Mining*, pp. 181–189, Houston, TX, USA, April 18–20, 2024. [DOI]

ICDM'23 Xiaoyu Zhu, Chenyang Bu, Bingbing Dong, Shengwei Ji, **Yi He**, and Xindong Wu, "IKGN: Intention-aware Knowledge Graph Network for POI Recommendation", published in *Proceedings of the 23rd IEEE International Conference on Data Mining*, pp. 908–917, Shanghai, China, December 1–4, 2023. [DOI]

ECML-PKDD'23 Cheng Liang, Di Wu, **Yi He**, Teng Huang, Zhong Chen, and Xin Luo, "MMA: Multi-Metric-Autoencoder for Analyzing High-Dimensional and Incomplete Data", published in *Machine Learning and Knowledge Discovery in Databases: Research Track (ECML PKDD 2023)*, pp. 3–19, Torino, Italy, September 18–22, 2023. [DOI]

IJCAI'23 **Yi He**, Christian Schreckenberger, Heiner Stuckenschmidt, and Xindong Wu, "Towards Utilitarian Online Learning – A Review of Online Algorithms in Open Feature Space", published in *Proceedings of the 32nd International Joint Conference on Artificial Intelligence*, pp. 6647–6655, Macao, China, August 19–25, 2023. [DOI]

AAAI'23 Di Wu, Shengda Zhuo, Yu Wang, Zhong Chen, and **Yi He** ✉, "Online Semi-Supervised Learning with Mix-Typed Streaming Features", published in *Proceedings of the 37th AAAI Conference on Artificial Intelligence*, Vol. 37, No. 4, pp. 4720–4728, Washington, DC, USA, February 7–14, 2023. [DOI]

AAAI'23 Christian Schreckenberger*, **Yi He**, Stefan Lüdtke, Christian Bartelt, and Heiner Stuckenschmidt, "Online Random Feature Forests for Learning in Varying Feature Spaces", published in *Proceedings of the 37th AAAI Conference on Artificial Intelligence*, Vol. 37, No. 4, pp. 4587–4595, Washington, DC, USA, February 7–14, 2023. [DOI]

ACM MM'22 Heng Lian*, John Scovi Atwood, Bo-Jian Hou, Jian Wu, and **Yi He** ✉, "Online Deep Learning from Doubly-Streaming Data", published in *Proceedings of the 30th ACM International Conference on Multimedia*, pp. 3185–3194, Lisbon, Portugal, October 10–14, 2022. [DOI]

ICNSC'22 Feilong Chen, Di Wu, Jie Yang, and **Yi He**, "Online Sparse Streaming Feature Selection with Uncertainty", published in *Proceedings of the 2022 IEEE International Conference on Networking, Sensing and Control*, pp. 1–6, Shanghai, China, December 15–18, 2022. [DOI]

UIC'22 Cheng Liang, **Yi He**, Teng Huang, and Di Wu, "AutoRec++: Incorporating Debias Methods Into Autoencoder-Based Recommender System", published in *Proceedings of the 19th IEEE International Conference on Ubiquitous Intelligence and Computing*, pp. 1876–1881, Haikou, China, December 15–18, 2022. [DOI]

UIC'22 Yuanpeng Hu, Xianmin Wang, Cheng Liang, Jing Li, Di Wu, and **Yi He** ✉, "OR-AutoRec: An Outlier-Resilient Autoencoder-Based Recommendation Model", published in *Proceedings of the 19th IEEE International Conference on Ubiquitous Intelligence and Computing*, pp. 1918–1923, Haikou, China, December 15–18, 2022. [DOI]

ICMR'22 Chao Jiang*, **Yi He** ✉, Richard Chapman, and Hongyi Wu ✉, "Camouflaged Poisoning Attack on Graph Neural Networks", published in *Proceedings of the 2022 ACM International Conference on Multimedia Retrieval*, pp. 451–461, Newark, NJ, USA, June 27–30, 2022. [DOI]

EDGE'22 Zhikai Yu, Di Wu, and **Yi He**, "A Robust Latent Factor Analysis Model for Incomplete Data Recovery in Wireless Sensor Networks", published in *Proceedings of the 2022 IEEE International Conference on Edge Computing and Communications*, pp. 178–183, Barcelona, Spain, July 10–16, 2022. [DOI]

DASFAA'22 Bo Sun, Di Wu, Mingsheng Shang, and **Yi He**, "Toward Auto-Learning Hyperparameters for Deep Learning-Based Recommender Systems", published in *Proceedings of the 27th International Conference on Database Systems for Advanced Applications*, pp. 323–331, Hyderabad, India, April 11–14, 2022. [DOI]

ICBK'21 Peng Zhang, **Yi He**, and Di Wu, "An Ensemble Latent Factor Model for Highly Accurate Web Service QoS Prediction", published in *Proceedings of the 12th IEEE International Conference on Big Knowledge*, pp. 361–368, Virtual conference (Auckland, New Zealand), December 7–8, 2021. [DOI]

ICDM'21 **Yi He**, Jiaxian Dong, Bo-Jian Hou, Yu Wang, and Fei Wang, "Online Learning in Variable Feature Spaces with Mixed Data", published in *Proceedings of the 2021 IEEE International Conference on Data Mining*, pp. 181–190, Virtual conference (Auckland, New Zealand), December 7–10, 2021. [DOI]

IJCAI'21 **Yi He**, Fudong Lin, Xu Yuan, and Nian-Feng Tzeng, "Interpretable Minority Synthesis for Imbalanced Classification", published in *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, pp. 2542–2548, Virtual conference, August 19–26, 2021. [DOI]

WWW'21 **Yi He**, Sheng Chen, Baijun Wu, Xu Yuan, and Xindong Wu, "Unsupervised Lifelong Learning with Curricula", published in *Proceedings of the Web Conference 2021*, pp. 3534–3545, Virtual conference (Ljubljana, Slovenia), April 19–23, 2021. [DOI]

AAAI'21 **Yi He**, Xu Yuan, Sheng Chen, and Xindong Wu, "Online Learning in Variable Feature Spaces under Incomplete Supervision", published in *Proceedings of the 35th AAAI Conference on Artificial Intelligence*, Vol. 35, No. 5, pp. 4106–4114, Virtual conference, February 2–9, 2021. [DOI]

ICDM'20 **Y. He**, Xu Yuan, Nian-Feng Tzeng, and Xindong Wu, "Active Learning with Multi-Granular Graph Auto-Encoder", published in *Proceedings of the 2020 IEEE International Conference on Data Mining*, pp. 1058–1063, Virtual conference (Sorrento, Italy), November 17–20, 2020. [DOI]

ECAI'20 **Yi He**, Baijun Wu, Di Wu, and Xindong Wu, "On Partial Multi-Task Learning", published in *Proceedings of the 24th European Conference on Artificial Intelligence*, pp. 1174–1181, Virtual conference, August 29–September 8, 2020. [DOI]

ECAI'20 Ege Beyazit, **Yi He**, N.-F. Tzeng, and Xindong Wu, "Online Learning to Accelerate Neural Network Inference with Traveling Classifiers", published in *Proceedings of the 24th European Conference on Artificial Intelligence*, pp. 1016–1023, Virtual conference, August 29–September 8, 2020. [DOI]

IJCNN'20 Jing Chen, **Yi He**, and Vijay Raghavan, "Learning with Partial Multi-Outlooks", published in *Proceedings of the 2020 International Joint Conference on Neural Networks*, pp. 1–8, Virtual conference (Glasgow, United Kingdom), July 19–24, 2020. [DOI]

SDM'20 **Yi He**, Sheng Chen, Thu Nguyen, Bruce A. Wade, and Xindong Wu, "Deep Matrix Tri-Factorization: Mining Vertex-wise Interactions in Multi-Space Attributed Graphs", published in *Proceedings of the 2020 SIAM International Conference on Data Mining*, pp. 334–342, 2020. [DOI]

BigData'19 Di Wu, **Yi He**, Xin Luo, Mingsheng Shang, and Xindong Wu, "Online Feature Selection with Capricious Streaming Features: A General Framework", published in *Proceedings of the 2019 IEEE International Conference on Big Data*, pp. 683–688, Los Angeles, CA, USA, December 9–12, 2019. [DOI]

OOPSLA'19 Baijun Wu, John Peter Campora III, **Yi He**, Alexander Schlecht, and Sheng Chen, "Generating Precise Error Specifications for C: A Zero Shot Learning Approach", published in *Proceedings of the ACM on Programming Languages*, Vol. 3, No. OOPSLA, Article 160, pp. 160:1–160:30, 2019. [DOI]

IJCAI'19 **Yi He**, Baijun Wu, Di Wu, Ege Beyazit, Sheng Chen, and Xindong Wu, "Online Learning from Capricious Data Streams: A Generative Approach", published in *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pp. 2491–2497, Macao, China, August 10–16, 2019. [DOI]

PAKDD'19 Di Wu, Xin Luo, Mingsheng Shang, **Yi He**, Guoyin Wang, and Xindong Wu, "A Data-Aware Latent Factor Model for Web Service QoS Prediction", published in *Proceedings of the 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 384–399, Macao, China, April 14–17, 2019. [DOI]

ICTAI'18 **Yi He**, Di Wu, Ege Beyazit, Xiaoduan Sun, and Xindong Wu, "Supervised Data Synthesizing and Evolving – A Framework for Real-World Traffic Crash Severity Classification", published in *Proceedings of the 30th IEEE International Conference on Tools with Artificial Intelligence*, pp. 163–170, Volos, Greece, November 5–7, 2018. [DOI]

Peer-Reviewed Journal Articles

TMLR Haoyang Jiang*, Bojian Qu, Xingquan Zhu, Jifu Tan, and **Yi He** ✉, "Boundary-Consistent Graph Neural Networks for Topological Flux Prediction", published in *Transactions on Machine Learning Research*, 2026. [Paper]

TON Di Wu, Chengjun Yu, **Yi He**, and Xin Luo, "A Novel Spatiotemporally-Regularized Latent Factorization of Tensors Model for Learning Accurate Dynamics from Temporal Sensing Networks", published in *IEEE Transactions on Networking*, Early Access, 2026. [DOI]

TDSC Di Wu, Shuai Zhong, **Yi He**, Xin Luo, and Xinbo Gao, "Federated Latent Factorization of Tensors for Privacy-Preserving Representation Learning to Large-Scale Dynamic Weighted Directed Network", published in *IEEE Transactions on Dependable and Secure Computing*, Vol. 23, No. 4, pp. 8431–8446, 2026. [DOI]

TPAMI Di Wu, Shihui Li, **Yi He**, Xin Luo, and Xinbo Gao, "Non-Gradient Hash Factor Learning for High-Dimensional and Incomplete Data Representation Learning", published in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 48, No. 5, pp. 5811–5826, 2026. [DOI]

KAIS Dong Li, Guihong Wan, Xintao Wu, Xinyu Wu, **Yi He**, Zhong Chen, Ajit J. Nirmal, Christine G. Lian, Peter K. Sorger, Yevgeniy R. Semenov, and Chen Zhao, "A Survey on Computational Pathology Foundation Models: Datasets, Adaptation Strategies, and Evaluation Tasks", published in *Knowledge and Information Systems*, Vol. 68, Article 209, 2026. [DOI]

KAIS Song Deng, Mengfei Sun, Lei Duan, and **Yi He**, "SADD-RFCO: Semi-Supervised Anomalous Data Detection Based on Random Forest with Co-Training", published in *Knowledge and Information Systems*, Vol. 68, Article 91, 2026. [DOI]

TSMC Di Wu, Cheng Liang, **Yi He**, Yan Qiao, and Xin Luo, "Multimetric Autoencoder for Representing High-Dimensional and Incomplete Data", published in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 56, No. 3, pp. 1533–1546, 2026. [DOI]

PHN Qi Zhang, Bidusha Neupane, Priyanka Patel, Futun N. Alkhalifah, **Yi He**, and Leslie Hodges, "Large Language Model-Assisted Research Question Development in Public Health: A Case Study in the Special Supplemental Nutrition Program for Women, Infants, and Children", published in *Public Health Nutrition*, Vol. 29, No. 1, Article e32, 2026. [DOI]

ACCESS Zhiqiang Wang, Sheri L. Barton, Jennifer H. Johnson, **Yi He**, and Xingquan Zhu, "Interpretable Individual Manatee Identification Using Sketch Images and Deep Neural Networks", published in *IEEE Access*, Vol. 13, pp. 218154–218168, 2025. [DOI]

TMLR Dong Li, Guihong Wan, Xintao Wu, Xinyu Wu, Xiaohui Chen, **Yi He**, Zhong Chen, Peter K. Sorger, and Chen Zhao, "Multi-Modal Foundation Models for Computational Pathology: A Survey", published in *Transactions on Machine Learning Research*, 2025. [Paper]

TKDE Heng Lian*, **Yi He** ✉, Di Wu, Zhong Chen, Xingquan Zhu, and Xindong Wu, "Online Outlier Detection in Open Feature Spaces", published in *IEEE Transactions on Knowledge and Data Engineering*, Vol. 37, No. 10, pp. 6091–6106, 2025. [DOI]

TKDD Shengda Zhuo, Di Wu, **Yi He**, Shuqiang Huang, and Xindong Wu, "Online Learning from Mix-typed, Drifted, and Incomplete Streaming Features", published in *ACM Transactions on Knowledge Discovery from Data*, Vol. 19, No. 8, Article 150, pp. 150:1–150:28, 2025. [DOI]

Measurement Song Deng, Li Tao, Dong Yue, and **Yi He**, "A shaping two-stage anomaly data recovery method based on multi-norm joint optimization under energy internet", published in *Measurement*, Vol. 242, Article 115949, 2025. [DOI]

AI Song Deng, Lin Chen, and **Yi He**, "Insulator defect detection from aerial images in adverse weather conditions", published in *Applied Intelligence*, Vol. 55, No. 6, Article 365, 2025. [DOI]

Energy Song Deng, Xia Dong, Li Tao, Junjie Wang, **Yi He**, and Dong Yue, "Multi-type load forecasting model based on random forest and density clustering with the influence of noise and load patterns", published in *Energy*, Vol. 307, Article 132635, 2024. [DOI]

TNNLS Di Wu, Zechao Li, Zhikai Yu, **Yi He**, and Xin Luo, "Robust Low-Rank Latent Feature Analysis for Spatiotemporal Signal Recovery", published in *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 36, No. 2, pp. 2829–2842, 2025. [DOI]

TKDE Heng Lian*, Di Wu, Bo-Jian Hou, Jian Wu, and **Yi He** ✉, "Online Learning from Evolving Feature Spaces with Deep Variational Models", published in *IEEE Transactions on Knowledge and Data Engineering*, Vol. 36, No. 8, pp. 4144–4162, 2024. [DOI]

TSC Di Wu, Peng Zhang, **Yi He**, and Xin Luo, "MMLF: Multi-Metric Latent Feature Analysis for High-Dimensional and Incomplete Data", published in *IEEE Transactions on Services Computing*, Vol. 17, No. 2, pp. 575–588, 2024. [DOI]

TETC Di Wu, **Yi He**, and Xin Luo, "A Graph-Incorporated Latent Factor Analysis Model for High-Dimensional and Sparse Data", published in *IEEE Transactions on Emerging Topics in Computing*, Vol. 11, No. 4, pp. 907–917, 2023. [DOI]

TSC Song Deng, Yujia Zhai, Di Wu, Dong Yue, Xiong Fu, and **Yi He**, "A Lightweight Dynamic Storage Algorithm With Adaptive Encoding for Energy Internet", published in *IEEE Transactions on Services Computing*, Vol. 16, No. 5, pp. 3115–3128, 2023. [DOI]

TCE Teng Huang, Cheng Liang, Di Wu, and **Yi He**, "A Debiasing Autoencoder for Recommender System", published in *IEEE Transactions on Consumer Electronics*, Vol. 70, No. 1, pp. 3603–3613, 2024. [DOI]

CEE Song Deng, Qicong Hu, Di Wu, and **Yi He**, "BCTC-KSM: A blockchain-assisted threshold cryptography for key security management in power IoT data sharing", published in *Computers and Electrical Engineering*, Vol. 108, Article 108666, 2023. [DOI]

TNNLS Di Wu, Xin Luo, **Yi He**, and MengChu Zhou, "A Prediction-Sampling-Based Multilayer-Structured Latent Factor Model for Accurate Representation to High-Dimensional and Sparse Data", published in *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 35, No. 3, pp. 3845–3858, 2024. [DOI]

TSC Di Wu, Peng Zhang, **Yi He**, and Xin Luo, "A Double-Space and Double-Norm Ensembled Latent Factor Model for Highly Accurate Web Service QoS Prediction", published in *IEEE Transactions on Services Computing*, Vol. 16, No. 2, pp. 802–814, 2023. [DOI]

TII Song Deng, Jiantang Zhang, Di Wu, **Yi He**, Xiangpeng Xie, and Xindong Wu, "A Quantitative Risk Assessment Model for Distribution Cyber-Physical System Under Cyberattack", published in *IEEE Transactions on Industrial Informatics*, Vol. 19, No. 3, pp. 2899–2908, 2023. [DOI]

CEE Song Deng, Fulin Chen, Di Wu, **Yi He**, Hui Ge, and Yuan Ge, "Quantitative combination load forecasting model based on forecasting error optimization", published in *Computers and Electrical Engineering*, Vol. 101, Article 108125, 2022. [DOI]

TKDE Zijie Pan, Li Hu, Weixuan Tang, Jin Li, **Yi He**, and Zheli Liu, "Privacy-Preserving Multi-Granular Federated Neural Architecture Search – A General Framework", published in *IEEE Transactions on Knowledge and Data Engineering*, Vol. 35, No. 3, pp. 2975–2986, 2023. [DOI]

TSMC Di Wu, **Yi He**, Xin Luo, and MengChu Zhou, "A Latent Factor Analysis-Based Approach to Online Sparse Streaming Feature Selection", published in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 52, No. 11, pp. 6744–6758, 2022. [DOI]

TKDE Di Wu, Xin Luo, Mingsheng Shang, **Yi He**, Guoyin Wang, and Xindong Wu, "A Data-Characteristic-Aware Latent Factor Model for Web Services QoS Prediction", published in *IEEE Transactions on Knowledge and Data Engineering*, Vol. 34, No. 6, pp. 2525–2538, 2022. [DOI]

TNNLS **Yi He**, Baijun Wu, Di Wu, Ege Beyazit, Sheng Chen, and Xindong Wu, "Toward Mining Capricious Data Streams: A Generative Approach", published in *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 32, No. 3, pp. 1228–1240, 2021. [DOI]

TSC Di Wu, Qiang He, Xin Luo, Mingsheng Shang, **Yi He**, and Guoyin Wang, "A Posterior-neighborhood-regularized Latent Factor Model for Highly Accurate Web Service QoS Prediction", published in *IEEE Transactions on Services Computing*, Vol. 15, No. 2, pp. 793–805, 2022. [DOI]

TSMC Di Wu, Xin Luo, Mingsheng Shang, **Yi He**, Guoyin Wang, and MengChu Zhou, "A Deep Latent Factor Model for High-Dimensional and Sparse Matrices in Recommender Systems", published in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 51, No. 7, pp. 4285–4296, 2021. [DOI]

App. Sci. Dianlong You, Xindong Wu, Limin Shen, **Yi He**, Xu Yuan, Zhen Chen, Song Deng, and Chuan Ma, "Online Streaming Feature Selection via Conditional Independence", published in *Applied Sciences*, Vol. 8, No. 12, Article 2548, 2018. [DOI]

TEACHING

William & Mary

DATA601 Computational Foundations of Data Science (Fall 2026)
Developer and Instructor for a foundational Ph.D. course covering computational linear algebra, optimization, numerical computing, and scalable scientific computation. Emphasizes computational thinking and the translation of mathematical formulations into efficient, stable, and trustworthy implementations.

DATA440/691 Graph Learning (Spring 2026, Fall 2025, Spring 2025)
Instructor for an undergraduate and graduate course on graph representations, ranking algorithms, graph neural networks, and spatiotemporal learning. Adapted the course from CS795/895, which I developed at ODU.

DATA301 Applied Machine Learning (Fall 2024)
Instructor for a course on applied machine learning, including supervised and unsupervised learning, model selection, and evaluation. Emphasizes the assumptions, strengths, and limitations of different methods.

Old Dominion University

CS620 Introduction to Data Science and Analytics (Spring 2024, Fall 2023, Spring 2023)
Instructor for an introductory course covering analytical, statistical, computational, and data visualization methods for extracting knowledge from complex datasets.

CS600 Algorithms and Data Structures (Fall 2023)
Instructor for a course on algorithm design and analysis, asymptotic complexity, sorting and graph algorithms, and supporting data structures.

CS390 Introduction to Theoretical Computer Science (Fall 2022, Fall 2021)
Instructor for a course on formal languages, automata, and computability, including finite automata, context-free grammars, pushdown automata, Turing machines, and undecidability.

CS795/895 Deep Learning on Graphs (Spring 2022)
Developer and Instructor for a graduate course on graph representation learning and deep learning for graph data, including graph neural networks and spatiotemporal models.

HONORS AND RECOGNITION

- NSF Faculty Early Career Development (CAREER) Award, 2026.
- Senior Member, IEEE, 2026.
- NSF CRII Award, 2023.
- IEEE TCII Volunteer Award, 2022, for outstanding service as Registration Chair of IEEE ICDM 2022.
- Outstanding Graduates, University of Louisiana at Lafayette, 2017.
- Third-Class People's Scholarship, Harbin Institute of Technology, 2012.

PROFESSIONAL SERVICE

Conference Organizing Committee

- Workshop Co-Chair: Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2027
- Proceedings Co-Chair: IEEE International Conference on Big Data 2026
- Doctoral Forum Co-Chair: IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2026
- Proceedings Co-Chair: IEEE International Conference on Data Mining (ICDM) 2025
- Workshop Co-Chair: IEEE International Conference on Data Mining (ICDM) 2024
- NSF Travel Award Chair: IEEE International Conference on Data Mining (ICDM) 2024
- Workshop Co-Chair: IEEE International Conference on Data Mining (ICDM) 2023
- Publicity Chair: IEEE International Conference on Knowledge Graph (ICKG) 2023
- Registration Chair: IEEE International Conference on Data Mining (ICDM) 2022

Technical Program Committee

(* indicates AC or Senior PC service)

- ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD) 2023, 2024, 2025, 2026
- The Web Conference 2025
- SIAM International Conference on Data Mining (SDM) 2024, 2025
- Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2023, 2024
- IEEE International Conference on Data Mining (ICDM) 2022, 2023, 2024, 2025, 2026
- ACM International Conference on Multimedia (ACM MM) 2023, 2024, 2025, 2026
- ACM International Conference on Information and Knowledge Management (CIKM) 2022
- AAAI Conference on Artificial Intelligence (AAAI) 2021, 2022, 2023, 2024, 2025, 2026*
- International Joint Conference on Artificial Intelligence (IJCAI) 2022, 2023, 2024, 2025, 2026*
- International Conference on Learning Representations (ICLR) 2025, 2026
- Annual Conference on Neural Information Processing Systems (NeurIPS) 2026

Journal Associate Editor

- Journal of Data, Information and Management (Springer Nature): 2026 – now
- Big Data Research (Elsevier): 2024 – now

Journal Guest Editor of Special Topics

- AI and Data Security for the Industrial Internet, hosted on Computers, Materials & Continua (CMC)
- Towards Human-Centric Computing and Equitable Artificial Intelligence, hosted on Frontiers in Neurorobotics.
- Neurorobotics and Neurological Diseases, hosted on Frontiers in Neurorobotics.

Journal Referee

- IEEE Transactions on Knowledge and Data Engineering
- ACM Transactions on Knowledge Discovery from Data
- IEEE Transactions on Neural Networks and Learning Systems
- IEEE Transactions on Mobile Computing
- IEEE Transactions on Services Computing
- IEEE Transactions on Fuzzy Systems
- IEEE Transactions on Intelligent Transportation systems
- IEEE Internet of Things Journal
- Frontiers in Neurorobotics
- IEEE/CAA Journal of Automatica Sinica
- Elsevier Neurocomputing
- Health Information Science and Systems
- Neural Computation (MIT Press)