

Employment

University of Virginia <i>Assistant Professor of Data Science</i>	Aug 2023 – present
MIT, CSAIL <i>Visiting Assistant Professor</i>	Aug 2023 – Jun 2024
MIT, CSAIL <i>Postdoctoral Associate</i>	Jan 2022 – Aug 2023
Microsoft <i>Intern</i>	May 2021 – Aug 2021
UMass Medical School <i>Intern</i>	Aug 2018 – Sep 2019

Education

Worcester Polytechnic Institute <i>Ph.D. in Data Science</i> Committee: Elke Rundensteiner, Xiangnan Kong, Randy Paffenroth, Jenna Wiens	2021
Worcester Polytechnic Institute <i>M.S. in Data Science</i>	2018
SUNY Geneseo <i>B.A. in Applied Mathematics</i>	2016

Grants & Awards

CapitalOne Data Design Fellowship, PI (\$55K) <i>Learning to Reason about Financial Time Series</i>	2025
Fellowship in AI Research, PI (\$200K) UVA Darden School of Business <i>Detecting Implicit Bias in Natural Language with Large Language Models</i>	2024
Research Innovation Grant – UCLA Initiative to Study Hate, Co-PI (\$25K) <i>Community Grounded Hate Speech Detection</i>	2024
Microsoft Accelerate Foundation Models Program, PI (\$30K compute) <i>Evaluation Framework for Continually Detoxifying LLMs</i>	2024
UVA Environmental Institute Climate Fellowship, Co-PI (\$150K) <i>Using Advanced AI for Simultaneous Bias-Correction and Downscaling of Climate Model Projections for Localized Policy Relevance</i>	2023
Microsoft Accelerate Foundation Models Program, PI (\$20K MSFT Azure Credits) <i>Enabling Large Language Models to Reason about Time Series</i>	2023

Selected Honors & Achievements

Oral Presentation , ICML'25 Position Paper Track (top ~2%)	2025
Outstanding Paper Award , AAAI'25 KnowFM Workshop	2025
Best Paper , ICML 2023 workshop on Interpretable Machine Learning in Healthcare	2023
Best Poster , International Conference on Health Informatics	2020
Outstanding Graduate Research Award , WPI	2019
Best Poster , Graduate Research Innovation and Exchange, WPI	2019
Best Student Paper Runner Up , International Conference on Health Informatics	2018
People's Choice Poster Award , Graduate Research Showcase, WPI	2017
GAANN PhD Fellowship , U.S. Dept. of Ed.	2016 – 2021

Publications

* denotes equal contribution

-
- 2025 79. *Math Neurosurgery: Isolating LLMs' Math Reasoning Abilities Using Only Forward Passes*.
Bryan Christ, Zachary Gottesman, Jonathan Kropko, **Thomas Hartvigsen**.
ACL - *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2025.
- 2025 78. *Lifelong Model Editing with Graph-Based External Memory*.
Yash Kumar Atri, Ahmed Alaa, **Thomas Hartvigsen**.
ACL - *Findings of the Annual Meeting of the Association for Computational Linguistics*, 2025.
- 2025 77. *Low-Bit Quantization Favors Undertrained LLMs*.
Xu Ouyang, Tao Ge, **Thomas Hartvigsen**, Zhisong Zhang, Haitao Mi, Dong Yu.
ACL - *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2025.
- 2025 76. *Efficient Knowledge Editing via Minimal Precomputation*.
Akshat Gupta, Maochuan Lu, **Thomas Hartvigsen**, Gopala Anumanchipalli.
ACL - *Proceedings of the Annual Meeting of the Association for Computational Linguistics, Short Paper*, 2025.
- 2025 75. *Understanding the Limits of Lifelong Knowledge Editing in LLMs*.
Lukas Thede, Karsten Roth, Matthias Bethge, Zeynep Akata, **Thomas Hartvigsen**.
ICML - *International Conference on Machine Learning*, 2025.
- 2025 74. *BalancEdit: Dynamically Balancing the Generality-Locality Trade-off in Multi-modal Model Editing*.
Dongliang Guo, Mengxuan Hu, Zihan Guan, **Thomas Hartvigsen**, Sheng Li.
ICML - *International Conference on Machine Learning*, 2025.
- 2025 73. *Position: Medical Large Language Model Benchmarks Should Prioritize Construct Validity*.
Ahmed Alaa, **Thomas Hartvigsen**, Niloufar Golchini, Shiladitya Dutta, Frances Dean, Inioluwa Deborah Raji, Travis Zack.
ICML - *International Conference on Machine Learning*, 2025. **Oral**.
- 2025 72. *Composable Interventions for Language Models*.
Arinbjörn Kolbeinsson, * Kyle O'Brien, * Tianjin Huang, * Shanghua Gao, Jonathan Richard Schwarz, Shiwei Liu, Anurag Vaidya, Faisal Mahmood, Marinka Zitnik, Tianlong Chen, **Thomas Hartvigsen**.
ICLR - *International Conference on Learning Representations*, 2025.
- 2025 71. *Learning under Temporal Label Noise*.
Sujay Nagaraj, Walter Gerych, Sana Tonekaboni, Anna Goldenberg, Berk Ustun, **Thomas Hartvigsen**.
ICLR - *International Conference on Learning Representations*, 2025.
- 2025 70. *PolyGuard: A Multilingual Safety Moderation Tool for 17 Languages*.
Priyanshu Kumar, Devansh Jain, Akhila Yerukola, Liwei Jiang, Himanshu Beniwal, **Thomas Hartvigsen**, Maarten Sap.
COLM - *Conference on Language Modeling*, 2025.

- 2025 69. *Disability Portrayals in Artificial Intelligence Text-to-Image Generation: Influence of Context and the Medicalization of Disability Rehabilitation Psychology*.
Benjamin Ertman, Bridget Xia, Mona Sloane, **Thomas Hartvigsen**, Paul B. Perrin.
Rehabilitation Psychology
- 2025 68. *Sparse MoE as a New Treatment: Addressing Forgetting, Fitting, and Learning Issues in Multi-Modal Multi-Task Learning*.
Jie Peng, Kaixiong Zhou, Ruida Zhou, **Thomas Hartvigsen**, Yanyong Zhang, Zhangyang Wang, Tianlong Chen.
CPAL - Conference on Parsimony and Learning, 2025.
- 2025 67. *Decoding the Rule Book: Extracting Hidden Moderation Criteria from Reddit Communities*.
Youngwoo Kim, Himanshu Beniwal, Steven Johnson, **Thomas Hartvigsen**.
preprint.
- 2025 66. *ModelCitizens: Representing Community Voices in Online Safety*.
Ashima Suvarna, Christina Chance, Karolina Naranjo, Hamid Palangi, Sophie Hao, **Thomas Hartvigsen**, Saadia Gabriel.
preprint.
- 2025 65. *Inferring Events from Time Series using Language Models*.
Mingtian Tan, Mike Merrill, Zachary Gottesman, Tim Althoff, David Evans, **Thomas Hartvigsen**.
preprint.
- 2025 64. *Continually Self-Improving Language Models for Bariatric Surgery Question-Answering*.
Yash Kumar Atri, Thomas Shin, **Thomas Hartvigsen**.
preprint.
- 2025 63. *Breaking mBad! Supervised Fine-tuning for Cross-Lingual Detoxification*.
Himanshu Beniwal, Youngwoo Kim, Maarten Sap, Soham Dan, **Thomas Hartvigsen**.
preprint.
- 2025 62. *Disparities in LLM Reasoning Accuracy and Explanations: A Case Study on African American English*.
Miles Zhou, * Guangya Wan, * Saadia Gabriel, Alex Gates, Sheng Li, Maarten Sap, **Thomas Hartvigsen**.
preprint.
- 2025 61. *Lifelong Sequential Knowledge Editing Requires Better Regularization*.
Akshat Gupta, Phudish Prateepamornkul, Maochuan Lu, Ahmed Alaa, **Thomas Hartvigsen**, Gopala Anu-
manchipalli.
preprint.
- 2025 60. *Constrained Discrete Diffusion*.
Michael Cardei, Jacob K Christopher, **Thomas Hartvigsen**, Bhavya Kailkhura, Ferdinando Fioretto.
preprint.
- 2025 59. *SAINT: Attention-Based Modeling of Sub-Action Dependencies in Multi-Action Policies*.
Matthew Landers, Taylor Killian, **Thomas Hartvigsen**, Afsaneh Doryab.
preprint.
- 2025 58. *How Can Time Series Analysis Benefit From Multiple Modalities? A Survey and Outlook*.
Haoxin Liu, Harshavardhan Kamarthi, Zhiyuan Zhao, Shangqing Xu, Shiyu Wang, Qingsong Wen, **Thomas Hartvigsen**, Fei Wang, B. Aditya Prakash.
preprint.
- 2025 57. *MedBrowseComp: Benchmarking Medical Deep Research and Computer Use*.
Shan Chen, Pedro Moreira, Yuxin Xiao, Sam Schmidgall, Jeremy Warner, Hugo Aerts, **Thomas Hartvigsen**,

Jack Gallifant, Danielle S. Bitterman.
preprint.

- 2024 56. [Are Language Models Actually Useful for Time Series Forecasting?](#)
Mingtian Tan, Michael Merrill, Vinayak Gupta, Tim Althoff, **Thomas Hartvigsen**.
NeurIPS - Advances in Neural Information Processing Systems, 2024. **Spotlight**.
- 2024 55. [UniTS: A Unified Multi-Task Time Series Model](#).
Shanghai Gao, Owen Queen, Teddy Koker, **Thomas Hartvigsen**, Theodoros Tsiligkaridis, Marinka Zitnik.
NeurIPS - Advances in Neural Information Processing Systems, 2024.
- 2024 54. [BendVLM: Test-Time Debiasing of Vision-Language Embeddings](#).
Walter Gerych, Haoran Zhang, Kimia Hamidieh, **Thomas Hartvigsen**, Marzyeh Ghassemi.
NeurIPS - Advances in Neural Information Processing Systems, 2024.
- 2024 53. [Language Models Still Struggle to Zero-Shot Reason about Time Series](#).
Michael Merrill, Mingtian Tan, Vinayak Gupta, **Thomas Hartvigsen**, Tim Althoff.
EMNLP - Empirical Methods for Natural Language Processing, Findings Track, 2024.
- 2024 52. [MATHWELL: Generating Educational Math Word Problems with Teacher Annotations](#).
Bryan Christ, Jonathan Kropko, **Thomas Hartvigsen**.
EMNLP - Empirical Methods for Natural Language Processing, Findings Track, 2024.
- 2024 51. [Language Models are Surprisingly Fragile to Drug Names in Biomedical Benchmarks](#)
Jack Gallifant, Shan Chen, Pedro Jos Ferreira Moreira, Nikolaj Munch, Mingye Gao, Jackson Pond, Leo Anthony Celi, Hugo Aerts, **Thomas Hartvigsen**, Danielle Bitterman.
EMNLP - Empirical Methods for Natural Language Processing, Short Paper, Findings Track, 2024.
- 2024 50. [TAXI: Evaluating Categorical Knowledge Editing for Language Models](#).
Derek Powell, Walter Gerych, **Thomas Hartvigsen**.
ACL - Annual Meeting of the Association for Computational Linguistics, Short Paper, Findings Track, 2024.
- 2024 49. [Demographic Bias in Misdiagnosis by Computational Pathology Models](#).
Anurag Vaidya, Richard Chen, Drew Williamson, Andrew Song, Guillaume Jaume, **Thomas Hartvigsen**, Yuzhe Yang, Emma Dyer, Ming Yang Lu, Jana Lipkova, Muhammad Shaban, Tiffany Y. Chen, Faisal Mahmood.
Nature Medicine, 30, pages 1174-1190, 2024.
- 2024 48. [Improving Black-box Robustness with In-Context Rewriting](#).
Kyle O'Brien, Nathan Ng, Isha Puri, Jorge Mendez, Hamid Palangi, Marzyeh Ghassemi, Yoon Kim, **Thomas Hartvigsen**.
TMLR - Transactions on Machine Learning Research, 2024.
- 2024 47. [Explaining Deep Multi-Class Time Series Classifiers](#).
Ramesh Doddaiiah, Prathyush Parvatharaju, Elke Rundensteiner, **Thomas Hartvigsen**.
KAIS - Knowledge and Information Systems, 2024.
- 2024 46. [Identifying Implicit Biases in Vision-Language Models](#).
Kimia Hamidieh, Haoran Zhang, Walter Gerych, **Thomas Hartvigsen**, Marzyeh Ghassemi.
AIES - AAAI Conference on AI, Ethics, and Society, 2024.
- 2024 45. [PolygloToxicityPrompts: Multilingual Evaluation of Neural Toxic Degeneration in Large Language Models](#).
Devansh Jain, Priyanshu Kumar, Sam Gehman, Xuhui Zhou, **Thomas Hartvigsen**, Maarten Sap.
COLM - Conference on Language Modeling, 2024.
- 2024 44. [FedMedICL: Towards Holistic Evaluation of Distribution Shifts in Federated Medical Imaging](#).
Kumail Alhamoud, Yasir M. Ghunaim, Motasem Alfarra, **Thomas Hartvigsen**, Philip Torr, Bernard Ghanem,

Adel Bibi, Marzyeh Ghassemi.

MICCAI - Medical Image Computing and Computer Assisted Intervention, 2024.

- 2024 43. *SkipSNN: Efficiently Classifying Noisy Spike Trains with Event Attention*.
Hang Yin, Yao Su, Liping Liu, **Thomas Hartvigsen**, Xin Dai, Xiangnan Kong.
IEEE Big Data - IEEE International Conference on Big Data, 2024.
- 2024 42. *ProCyon: A Foundation Model of Human Protein Phenotypes*.
Owen Queen, Yepeng Huang, Robert Calef, Valentina Giunchiglia, Tianlong Chen, George Dasoulas, LeAnn Tai, Yasha Ektefaie, Ayush Noori, Joseph Brown, Tom Cobley, Karin Hrovatin, **Thomas Hartvigsen**, Bradley Pentelute, Vikram Khurana, Manolis Kellis, Marinka Zitnik.
preprint.
- 2024 41. *BraVE: Offline Reinforcement Learning for Discrete Combinatorial Action Spaces*.
Matthew Landers, Taylor Killian, Hugo Barnes, **Thomas Hartvigsen**, Afsaneh Doryab.
preprint.
- 2024 40. *Backdoor in Seconds: Unlocking Vulnerabilities in Large Pre-trained Models via Model Editing*.
Dongliang Guo,* Mengxuan Hu,* Zihan Guan, **Thomas Hartvigsen**, Sheng Li.
preprint.
- 2024 39. *Wait, but Tylenol is Acetaminophen... Investigating and Improving Language Models' Ability to Resist Requests for Misinformation*
Shan Chen, Mingye Gao, Kuleen Sasse, **Thomas Hartvigsen**, Brian Anthony, Lizhou Fan, Hugo Aerts, Jack Gallifant, Danielle Bitterman.
preprint.
- 2024 38. *Dr-LLaVA: Visual Instruction Tuning with Symbolic Clinical Grounding*.
Shenghuan Sun, Alex Schubert, Zhiqing Sun, Greg Goldgof, **Thomas Hartvigsen**, Atul Butte, Ahmed Alaa.
preprint.
- 2024 37. *AnnotateGPT: Using Large Language Models for Social Determinants of Health Annotation*.
Bernardo Scapini Consoli, Song Wang, Xizhi Wu, Xinyu Zhao, Yanshan Wang, Justin Rousseau, **Thomas Hartvigsen**, Li Shen, Huanmei Xu, Qi Long, Tianlong Chen, Yifan Peng, Ying Ding.
preprint.
- 2023 36. *Aging with GRACE: Lifelong Model Editing with Discrete Key-Value Adaptors*.
Thomas Hartvigsen, Swami Sankaranarayanan, Hamid Palangi, Yoon Kim, Marzyeh Ghassemi.
NeurIPS - Advances in Neural Information Processing Systems, 2023.
- 2023 35. *Encoding Time-Series Explanations through Self-Supervised Model Behavior Consistency*.
Owen Queen, **Thomas Hartvigsen**, Huan He, Teddy Koker, Theodoros Tsiligkaridis, Marinka Zitnik.
NeurIPS - Advances in Neural Information Processing Systems, 2023. **Spotlight**.
- 2023 34. *Dissecting the Heterogeneity of "In-the-Wild" Stress from Multimodal Sensor Data*.
Sujoy Nagaraj, Sarah Goodday, **Thomas Hartvigsen**, Adrien Boch, Luca Foschini, Marzyeh Ghassemi, Stephen Friend, and Anna Goldenberg.
npj Digital Medicine.
- 2023 33. *Taking Off with AI: Lessons from Aviation for Healthcare*.
Elizabeth Bondi-Kelly, **Thomas Hartvigsen**, Lindsay Sanneman, Swami Sankaranarayanan, Lauren Oakden-Rayder, Leo Anthony Celi, Julie Shah, Marzyeh Ghassemi.
EAMMO - ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, 2023.
- 2023 32. *Knowledge Amalgamation for Multi-Label Classification via Label Dependency Transfer*.
Jidapa Thadajarassiri, **Thomas Hartvigsen**, Walter Gerych, Xiangnan Kong, Elke Rundensteiner.
AAAI - AAAI Conference on Artificial Intelligence, 2023.

- 2023 31. *Algorithmic Fairness in Chest X-ray Diagnosis*.
Haoran Zhang, **Thomas Hartvigsen**, Marzyeh Ghassemi.
SERC - MIT Case Studies in Social and Ethical Responsibilities of Computing, 2023.
- 2023 30. *Multi-State Brain Network Discovery*.
Hang Yin, Yao Su, Xinyue Liu, **Thomas Hartvigsen**, Yanhua Li, and Xiangnan Kong.
IEEE BigData - IEEE International Conference on Big Data, 2023.
- 2023 29. *Stabilizing Adversarial Training for Generative Networks*.
Walter Gerych, Kevin Hickey, **Thomas Hartvigsen**, Luke Buquicchio, Abdulaziz Alajaji, Kavin Chandrasekaran, Hamid Mansoor, Emmanuel Agu, and Elke Rundensteiner.
IEEE BigData - IEEE International Conference on Big Data, 2023.
- 2023 28. *Finding Short Signals in Long Irregular Time Series with Continuous-Time Attention Policy Networks*.
Thomas Hartvigsen, Jidapa Thadajarassiri, Xiangnan Kong, Elke Rundensteiner.
preprint.
- 2023 27. *Interpretable Unified Language Checking*.
Tianhua Zhang, Hongyin Luo, Yung-Sung Chuang, Wei Fang, Luc Gaitskell, **Thomas Hartvigsen**, Xixin Wu, Danny Fox, Helen Meng, James Glass.
preprint.
- 2022 26. *ToxiGen: A Large-Scale Machine-Generated Dataset for Adversarial and Implicit Hate Speech Detection*.
Thomas Hartvigsen, Saadia Gabriel, Hamid Palangi, Maarten Sap, Dipankar Ray, Ece Kamar.
ACL - Annual Meeting of the Association for Computational Linguistics, 2022.
- 2022 25. *Stop&Hop: Early Classification of Irregular Time Series*.
Thomas Hartvigsen, Walter Gerych, Jidapa Thadajarassiri, Xiangnan Kong, Elke Rundensteiner.
CIKM - ACM International Conference on Information and Knowledge Management, 2022.
- 2022 24. *Class-Specific Explainability for Deep Time Series Classifiers*.
Ramesh Doddaiiah, Prathyush Parvatharaju, Elke Rundensteiner, **Thomas Hartvigsen**.
ICDM - International Conference on Data Mining, 2022.
- 2022 23. *Robust Recurrent Classifier Chains For Multi-Label Learning With Missing Labels*.
Walter Gerych, **Thomas Hartvigsen**, Emmanuel Agu, Elke Rundensteiner.
CIKM - ACM International Conference on Information and Knowledge Management, 2022.
- 2022 22. *Recovering the Propensity Score from Biased Positive Unlabeled Data*.
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Emmanuel Agu, Elke Rundensteiner.
AAAI - AAAI Conference on Artificial Intelligence, 2022. **Oral**.
- 2022 21. *The Road to Explainability is Paved with Bias: Measuring the Fairness of Explanations*.
Aparna Balagopalan, Haoran Zhang, Kimia Hamidieh, **Thomas Hartvigsen**, Frank Rudzicz, Marzyeh Ghassemi.
FAccT - ACM Conference on Fairness, Accountability, and Transparency, 2022.
- 2022 20. *Positive Unlabeled Learning with a Sequential Selection Bias*.
Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Kavin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, Elke Rundensteiner, Emmanuel Agu.
SDM - SIAM International Conference on Data Mining, 2022.
- 2022 19. *TWEET-FID: An Annotated Dataset for Multiple Foodborne Illness Detection Tasks*.
Ruofan Hu, Dongyu Zhang, Dandan Tao, **Thomas Hartvigsen**, Hao Feng, Elke Rundensteiner.
LREC - Language Resources and Evaluation Conference, 2022.
- 2021 18. *Recurrent Bayesian Classifier Chains for Exact Multi-label Classification*.

Walter Gerych, **Thomas Hartvigsen**, Luke Buquicchio, Emmanuel Agu, Elke Rundensteiner.
NeurIPS - Advances in Neural Information Processing Systems, 2021.

- 2021 17. [Energy-Efficient Models for High-Dimensional Spike Train Classification using Sparse Spiking Neural Networks](#).
Hang Yin, John Boaz Lee, Xiangnan Kong, **Thomas Hartvigsen**, Sihong Xie.
KDD - ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2021.
- 2021 16. [Semi-Supervised Knowledge Amalgamation for Sequence Classification](#).
Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
AAAI - AAAI Conference on Artificial Intelligence, 2021.
- 2021 15. [Learning Saliency Maps to Explain Deep Time Series Classifiers](#).
Prathyush Parvatharaju, Ramesh Doddaiiah, **Thomas Hartvigsen***, Elke Rundensteiner*.
CIKM - ACM International Conference on Information and Knowledge Management, 2021.
- 2021 14. [Variational Open-Set Recognition](#).
Luke Buquicchio, Walter Gerych, Kavin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, **Thomas Hartvigsen**,
Elke Rundensteiner, Emmanuel Agu.
IEEE BigData - IEEE International Conference on Big Data, 2021.
- 2021 13. [Human-like Explanation for Text Classification with Limited Attention Supervision](#).
Dongyu Zhang, Cansu Sen, Jidapa Thadajarassiri, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
IEEE BigData - IEEE International Conference on Big Data, 2021.
- 2020 12. [Recurrent Halting Chain for Early Multi-label Classification](#).
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
KDD - ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2020.
- 2020 11. [Human Attention Maps for Text Classification: Do Humans and Neural Networks Focus on the Same Words?](#)
Cansu Sen, **Thomas Hartvigsen**, Biao Yin, Xiangnan Kong, Elke Rundensteiner.
ACL - Annual Meeting of the Association for Computational Linguistics, 2020.
- 2020 10. [Learning to Selectively Update State Neurons in Recurrent Networks](#).
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
CIKM - ACM International Conference on Information and Knowledge Management, 2020.
- 2020 9. [Learning Similarity Preserving Word Meta-Embedding](#).
Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
BigData - IEEE International Conference on Big Data, 2020.
- 2020 8. [Clinical Performance Evaluation of a Machine Learning System for Predicting Hospital-Acquired Clostridium Difficile Infection](#).
Erin Teeple, **Thomas Hartvigsen**, Cansu Sen, Kajal Claypool, Elke Rundensteiner.
International Conference on Health Informatics, 2020. 🏆 **Best Poster**.
- 2019 7. [Adaptive-Halting Policy Network for Early Classification](#).
Thomas Hartvigsen, Cansu Sen, Xiangnan Kong, Elke Rundensteiner.
KDD - ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2019.
- 2019 6. [Patient-Level Classification of Clinical Note Sequences Guided by Attributed Hierarchical Attention](#).
Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
BigData - IEEE International Conference on Big Data, 2019.
- 2019 5. [Learning Temporal Relevance in Longitudinal Medical Notes](#).

- Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
BigData - IEEE International Conference on Big Data, 2019.
- 2019 4. *Comparing General and Locally-Learned Word Embeddings for Clinical Text Mining*.
Jidapa Thadajarassiri, Cansu Sen, **Thomas Hartvigsen**, Xiangnan Kong, Elke Rundensteiner.
IEEE International Conference on Biomedical and Health Informatics, 2019.
- 2018 3. *Detecting MRSA Infections by Fusing Structured and Unstructured Electronic Health Record Data*.
Thomas Hartvigsen, Cansu Sen, Elke Rundensteiner.
International Joint Conference on Biomedical Engineering Systems and Technologies, 2018.
- 2018 2. *Early Prediction of MRSA Infections using Electronic Health Records*.
Thomas Hartvigsen, Cansu Sen, Sarah Brownell, Erin Teeple, Xiangnan Kong, Elke Rundensteiner.
International Conference on Health Informatics, 2018. 🏆 **Best Student Paper runner up**.
- 2017 1. *CREST - Risk Prediction for Clostridium Difficile Infection Using Multimodal Data Mining*.
Cansu Sen, **Thomas Hartvigsen**, Kajal Claypool, Elke Rundensteiner.
ECML - European Conference on Machine Learning, 2017.

Peer-reviewed Workshop Publications

- 2025 6. *Norm Growth and Stability Challenges in Localized Sequential Knowledge Editing*.
Akshat Gupta, Phudish Prateepamornkul, Christine Fang, Atahan Ozdemir, Maochuan Lu, **Thomas Hartvigsen**,
Gopala Anumanchipalli.
AAAI 2025 Workshop on Knowledgeable Foundation Models. **Outstanding Paper Award**.
- 2023 5. *Continuous Time Evidential Distributions for Irregular Time Series*.
Taylor Killian, Haoran Zhang, **Thomas Hartvigsen**, Ava Amini.
ICML 2023 Workshop on Interpretable ML in Healthcare.
- 2023 4. *A Pipeline for Interpretable Clinical Subtyping with Deep Metric Learning*.
Haoran Zhang, * Qixuan Jin, * **Thomas Hartvigsen**, * Miriam Udler, Marzyeh Ghassemi.
ICML 2023 Workshop on Interpretable Machine Learning in Healthcare. **Best Paper**.
- 2022 3. *Real-world Relevance of Generative Counterfactual Explanations*.
Swami Sankaranarayanan, **Thomas Hartvigsen**, Lauren Oakden-Rayner, Marzyeh Ghassemi, Philip Isola.
NeurIPS 2022 Workshop on Trustworthy and Socially Responsible Machine Learning.
- 2022 2. *On Detecting COVID-Risky Behavior from Smartphones*.
Thomas Hartvigsen, * Walter Gerych, * Marzyeh Ghassemi.
KDD 2022 epiDAMIK Workshop. **Spotlight**.
- 2022 1. *Learning Fair Multimodal Checklists for Interpretable Disease Classification*.
Qixuan Jin, Haoran Zhang, **Thomas Hartvigsen**, Marzyeh Ghassemi.
NeurIPS 2022 Workshop on Learning from Time Series for Health.

Invited Talks

UVA Health System Center for Advanced Medical Analytics	2025
UVA Darden Fellowship in AI Research Symposium	2025
IBM Research	2024
Arizona State University	2024
UMass Amherst, Machine Learning seminar	2024
Dartmouth College, CS Colloquium	2024
UC Berkeley/UCSF	2024
University of Alabama, Birmingham	2024

UVA Genome Sciences Seminar	2024
UVA Cancer Center Seminar	2024
UVA Medicine, <i>Dept. of Anesthesiology, Grand Rounds seminar</i>	2023
Cornell University, <i>ML in Medicine seminar series</i>	2023
University of Oxford, <i>CSML seminar series</i>	2023
CodaMetrix	2023
Stanford, <i>MedAI seminar series</i>	2023
IIIT Delhi, <i>LC2S seminar series</i>	2023
Harvard Medical School	2023
Rochester Institute of Technology	2023
AJCAI Workshop on Toxic Language Detection, <i>invited speaker</i>	2022
Worcester Polytechnic Institute, <i>Computer Science Colloquium</i>	2022
Northeastern	2022
Google Research, <i>Responsible AI seminar series</i>	2022
MIT, <i>Horizons Spotlight series</i>	2022

Service & Mentorship

Professional Organizing:

- NeurIPS'25 Workshop on Learning from Time Series for Health	2025
- EAAMO Faculty Network	2024
- ICLR'24 Workshop on Learning from Time Series for Health	2024
- General Chair , Machine Learning for Health (ML4H) Symposium	2023
- Conference on Health, Informatics, and Learning (CHIL), Doctoral Symposium	2023
- ICML'23 Workshop on Deployment Challenges in Generative AI, Chair	2023
- NeurIPS'22 Workshop on Learning from Time Series for Health, Chair	2022
- MIT EECS GAAP mentor, <i>Helping underrepresented students apply to graduate school</i>	2022

Conference Area Chair: NeurIPS 2025, ICML 2025, MLHC 2025, ML4H 2024

Peer Reviewing: AAAI, ACL, EMNLP, NeurIPS, ICLR, CVPR, ICCV, KDD, CHIL, WSDM, Nature Machine Intelligence, NeurIPS Workshop Proposals (2023–2025)

Patents

Hamid Palangi, Saadia Gabriel, Thomas Hartvigsen, Dipankar Ray, Maarten Sap, and Ece Kamar. *Adversarial language imitation with controlled exemplars*. Patent ID: #US12033617B2