

Research Interests

Areas Artificial Intelligence (AI), Machine Learning (ML), Computer Systems, Computer Architecture
Topics AI/ML Systems, Sustainability, Human-Centered Computing

Academic Employment

8/2028-present **Purdue University**, West Lafayette, IN, USA
Assistant Professor in Elmore Family School of Electrical and Computer Engineering
Affiliated Faculty, Institute for a Sustainable Future (ISF)
Affiliated Faculty, Institute for Physical Artificial Intelligence (IPAI)
1/2021-8/2023 **Massachusetts Institute of Technology**, Cambridge, MA, USA
Postdoctoral Associate & NSF Computing Innovation Fellow. Mentor: Michael Carbin.

Education

2015–2020 **The University of Chicago**, Chicago, IL, USA
PhD & MS in Computer Science. Advisor: Henry Hoffmann
PhD Dissertation: *Learning Structure for Computer Systems Management*
2013–2015 **Nanyang Technological University**, Singapore
Doctoral Student in Computer Science. Passed Qualification Exam.
2008–2012 **Beijing Jiaotong University**, Beijing, China
B.E. in Electronic Science and Technology. Graduated with *Highest Honor*.

Industry Employment

10/2021-12/2022 **Meta**, Cambridge, MA, USA
Visiting Researcher. Led a team to improve maintenance efficiency of servers in hyperscale datacenters.
6/2019-9/2019 **Google**, Sunnyvale, CA, USA
Research Intern. Introduced causal analysis to fleet understanding and automated confounder discovery.

Awards and Honors

2023 Innovation Award, Quantum Computing for Drug Discovery Challenge at ICCAD
2020-2023 NSF/CRA/CCC Computing Innovation Fellowship [[Link](#)].
2021 Meta Research Award [[Link](#)]
2020 EECS Rising Stars at UC Berkeley [[Link](#)]
2017 NeurIPS Spotlight Paper (top 4% of submissions)

Publications

Underline Students advised by me; ★ Equal contribution; † Corresponding faculty author

MLSys'26 **Cost-aware Duration Prediction for Software Upgrades in Datacenters**
Yi Ding[†], Aijia Gao, Thibaud Ryden, Michal Sedlak, Essam Ewaisha, Igor Marnat, Henry Hoffmann.
The 9th Annual Conference on Machine Learning and Systems (MLSys), 2026.
ICLR'26 **ViTSP: A Vision Language Models Guided Framework for Large-Scale Traveling Salesman Problems**
Zhuoli Yin, **Yi Ding**, Reem Khir, Hua Cai.
The 14th International Conference on Learning Representations (ICLR), 2026.
SIGMETRICS'26 **Cache Your Prompt When It's Green — Carbon-Aware Caching for Large Language Model Serving**
Yuyang Tian, Desen Sun, **Yi Ding**, Sihang Liu.
ACM on Measurement and Analysis of Computer Systems (POMACS), 2026.

- EM'25 **Beyond Climate Change: A Holistic Framework for Evaluating the Environmental Impact of Computing Systems**
Yi Ding[†], [Tianyao Shi](#), [Yanran Wu](#), Inez Hua.
 The Magazine for Environmental Manager (EM), 2025.
- Dagstuhl'25 **Climate Change: What is Computing's Responsibility?** [DOI]
 Bran Knowles, Vicki L. Hanson, Christoph Becker, Mike Berners-Lee, Andrew A. Chien, Benoit Combemale, Vlad Coroama, Koen De Bosschere, **Yi Ding**, Adrian Friday, Boris Gamazaychikov, Lynda Hardman, Simon Hinterholzer, Mattias Hojer, Lynn Kaack, Lenneke Kuijer, Anne-Laure Ligozat, Jan Tobias Muehlberg, Yunmook Nah, Thomas Olsson, Anne-Cecile Orgerie, Daniel Pargman, Birgit Penzenstadler, Tom Romanoff, Emma Strubell, Colin Venters, Junhua Zhao.
 Dagstuhl Perspectives Workshop 25122. In Dagstuhl Reports, Volume 15, Issue 3, pp. 113-124, 2025.
- IEEE CAL'25 **Disaggregated Speculative Decoding for Carbon-Efficient LLM Serving** [DOI]
[Tianyao Shi^{*}](#), [Yanran Wu^{*}](#), Sihang Liu, **Yi Ding[†]**.
 IEEE Computer Architecture Letters (CAL), 2025.
- ACL'25 **Unveiling Environmental Impacts of Large Language Model Serving: A Functional Unit View** [DOI]
[Yanran Wu](#), Inez Hua, **Yi Ding[†]**.
 The 63rd Annual Meeting of the Association for Computational Linguistics (ACL), 2025.
- HotCarbon'25 **Not All Water Consumption Is Equal: A Water Stress Weighted Metric for Sustainable Computing** [DOI]
[Yanran Wu](#), Inez Hua, **Yi Ding[†]**.
 The 4th Workshop on Sustainable Computer Systems (HotCarbon) and ACM SIGENERGY Energy Informatics Review (EIR), 2025.
- HotCarbon'25 **When Servers Meet Species: A Fab-to-Grave Lens on Computing's Biodiversity Impact** [DOI]
[Tianyao Shi](#), [Ritvik Kumar](#), Inez Hua, **Yi Ding[†]**.
 The 4th Workshop on Sustainable Computer Systems (HotCarbon) and ACM SIGENERGY Energy Informatics Review (EIR), 2025.
- e-Energy'25 **EnsembleCI: Ensemble Learning for Carbon Intensity Forecasting** [DOI]
[Leyi Yan](#), [Linda Wang](#), Sihang Liu, **Yi Ding[†]**.
 The 16th ACM International Conference on Future Energy Systems (e-Energy), 2025.
- CHASE'25 **Predicting and Understanding College Student Mental Health with Interpretable Machine Learning** [DOI]
[Meghna Roy Chowdhury](#), [Wei Xuan](#), Sheyras Sen, Yixue Zhao, **Yi Ding[†]**.
 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2025.
- MOBILESoft'25 **Unlocking Mental Health: Exploring College Students' Well-being through Smartphone Behaviors** [DOI]
[Wei Xuan](#), [Meghna Roy Chowdhury](#), **Yi Ding**, Yixue Zhao.
 IEEE/ACM 12th International Conference on Mobile Software Engineering and Systems (MOBILESoft), 2025.
- EMBC'25 **SSL-SE-EEG: A Framework for Robust Learning from Unlabeled EEG Data with Self-Supervised Learning and Squeeze-Excitation Networks**
[Meghna Roy Chowdhury](#), **Yi Ding**, Shreyas Sen.
 The 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2025.
- IGSC'24 **Sustainable LLM Serving: Environmental Implications, Challenges, and Opportunities** [DOI]
Yi Ding[†], [Tianyao Shi](#).
 The 15th International Green and Sustainable Computing Conference (IGSC), 2024.
- HotCarbon'24 **Uncertainty-Aware Decarbonization for Datacenters** [DOI]
[Amy Li](#), Sihang Liu, **Yi Ding[†]**.
 The 3rd Workshop on Sustainable Computer Systems (HotCarbon) and ACM SIGENERGY Energy Informatics Review (EIR), 2024.

- HotCarbon'24 **Towards Sustainable Large Language Model Serving** [DOI]
 Sophia Nguyen*, Beihao Zhou*, Yi Ding, Sihang Liu.
 The 3rd Workshop on Sustainable Computer Systems (HotCarbon) and ACM SIGENERGY Energy Informatics Review (EIR), 2024.
- ASPLOS'23 **CAFQA: A Classical Simulation Bootstrap for Variational Quantum Algorithms** [DOI]
 Gokul Subramanian Ravi, Pranav Gokhale, Yi Ding, William Kirby, Kaitlin Smith, Jonathan M. Baker, Peter J. Love, Henry Hoffmann, Kenneth R. Brown, Frederic T. Chong.
 The 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2023.
 2023 Innovation Award, Quantum Computing for Drug Discovery Challenge at ICCAD.
- OOPSLA'23 **Turaco: Complexity-Guided Data Sampling for Training Neural Surrogates of Programs** [DOI]
 Alex Renda, Yi Ding, Michael Carbin.
 ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA), 2023.
- Biometrika'23 **Minimax Designs for Causal Effects in Temporal Experiments with Treatment Habituation** [DOI]
 Guillaume W Basse, Yi Ding, Panos Toulis.
 Biometrika, 2023.
- MLSys'22 **NURD: Negative-Unlabeled Learning for Online Datacenter Straggler Prediction**
 Yi Ding, Avinash Rao, Hyebin Song, Rebecca Willett, Henry Hoffmann.
 Machine Learning and Systems (MLSys), 2022.
- ESEC/FSE'21 **Generalizable and Interpretable Learning for Configuration Extrapolation** [DOI]
 Yi Ding, Ahsan Pervaiz, Michael Carbin, Henry Hoffmann.
 The 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE), 2021.
- Onward!'21 **Programming with Neural Surrogates of Programs** [DOI]
 Alex Renda, Yi Ding, Michael Carbin.
 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (Onward!), 2021.
- npj Urban'21 **Neighborhood Street Activity and Greenspace Usage Uniquely Contribute to Predicting Crime** [DOI]
 Kathryn E. Schertz, James Saxon, Carlos Cardenas-Iniguez, Luis Bettencourt, Yi Ding, Henry Hoffmann, Marc G. Berman.
 npj Urban Sustainability, 2021.
- AISTATS'20 **Dynamical Systems Theory for Causal Inference with Application to Synthetic Control Methods** [DOI]
 Yi Ding, Panos Toulis.
 International Conference on Artificial Intelligence and Statistics (AISTATS), 2020.
- NeurIPS'20 **A Polynomial-Time Algorithm for Learning Nonparametric Causal Graphs** [DOI]
 Ming Gao, Yi Ding, Bryon Aragam.
 The 34th Annual Conference on Advances in Neural Information Processing Systems (NeurIPS), 2020.
- ISCA'19 **Generative and Multi-Phase Learning for Computer Systems Optimization** [DOI]
 Yi Ding, Nikita Mishra, Henry Hoffmann.
 The 46th International Symposium on Computer Architecture (ISCA), '19.
- NeurIPS'17 **Multiresolution Kernel Approximation for Gaussian Process Regression** [DOI]
 Yi Ding, Risi Kondor, Jonathan Eskreis-Winkler.
 The 31st Annual Conference on Advances in Neural Information Processing Systems (NeurIPS), 2017.
 Spotlight Presentation, Top 4% of Submissions.
- ICDM'17 **Large Scale Kernel Methods for Online AUC Maximization** [DOI]
 Yi Ding, Chenghao Liu, Peilin Zhao, Steven C.H. Hoi.
 IEEE International Conference on Data Mining (ICDM), 2017.
 Long Oral, Top 8% of Submissions.

AAAI'15 **An Adaptive Gradient Method for Online AUC Maximization** [DOI]

Yi Ding, Peilin Zhao, Steven C.H. Hoi, Yew-Soon Ong.

The 29th AAAI Conference on Artificial Intelligence (AAAI), 2015.

Oral Presentation, Top 10% of Submissions.

AAAI'14 **Learning Relative Similarity by Stochastic Dual Coordinate Ascent** [DOI]

Pengcheng Wu, **Yi Ding**, Peilin Zhao, Chunyan Miao, Steven C.H. Hoi.

The 28th AAAI Conference on Artificial Intelligence (AAAI), 2014.

Grants

- | | |
|-----------|---|
| | Funder: National Science Foundation |
| | Title: Conference: DESC: Type III: A Holistic AI Computing Framework: |
| 2024-2026 | Incorporating the Water and Biodiversity Dimensions of Sustainability |
| | People: Inez Hua (PI), Yi Ding |
| | Awarded: \$99,992 (my share: 50%) |
| | Funder: Purdue University |
| | Title: Seed Funding for High-Impact Review Papers |
| 2024-2025 | |
| | People: Inez Hua (PI), Yi Ding |
| | Awarded: \$10,000 (my share: 50%) |
| | Funder: Meta |
| | Title: Negative-Unlabeled Learning for Online Datacenter Straggler Prediction |
| 2021-2022 | |
| | People: Michael Carbin (PI), Yi Ding |
| | Awarded: \$46,000 |
| | Funder: National Science Foundation |
| | Title: CRA/CCC Computing Innovation Fellowship |
| 2020-2023 | |
| | People: Yi Ding |
| | Awarded: \$295,704 |

Teaching

Instructor, Purdue University

- | | |
|-------------|---|
| Spring 2026 | Python for Data Science (ECE 20875) |
| Fall 2025 | Machine Learning in Cloud Computing (ECE 69500) |
| Spring 2025 | Python for Data Science (ECE 20875) |
| Fall 2024 | Machine Learning in Cloud Computing (ECE 69500) |
| Spring 2024 | Python for Data Science (ECE 20875) |
| Fall 2023 | Python for Data Science (ECE 20875) |

Teaching Assistant, University of Chicago

- | | |
|-------------|---|
| Spring 2017 | Machine Learning and Large Scale Data Analysis (CMSC 25025) |
| Winter 2017 | Machine Learning (CMSC 25400) |
| Spring 2016 | Machine Learning (MPCS 53111) |
| Winter 2016 | Machine Learning for Public Policy (CAPP 30255) |

Research Advising

PhD Students

- | | |
|------------|---|
| Fall 2025- | Asher Sprigler (ECE), Purdue University |
| Fall 2025- | Lauren Caccamise (ECE), Purdue University |
| Fall 2024- | Tianyao Shi (ECE), Purdue University |

PhD Thesis Committee

- | | |
|------------|---|
| Fall 2021- | Meghna Roy Chowdhury (ECE), Purdue University |
| Fall 2021- | Zhuoli Yin (IE), Purdue University |

Master Students

- | | |
|-----------|-----------------|
| 2021-2022 | Hyunji Kim, MIT |
|-----------|-----------------|

Undergraduate Students

Summer 2025 Barbara Su, Rice University (Remote Intern, SURF)
 Spring 2025 Jaewon Cho, Purdue University (DUIRI, awarded \$1,000 fellowship)
 Spring 2025 Isha Shamim, Purdue University (DUIRI, awarded \$1,000 fellowship)
 Fall 2024 Gavin Fortwendel, Purdue University (DUIRI, awarded \$1,000 fellowship; **Won 1st Place in Research Talk, College of Engineering, Fall 2024 Undergraduate Research Expo**)
 Fall 2024 Sarah Deniz, Purdue University (DUIRI, awarded \$1,000 fellowship)
 Fall 2024 Leyi Yan, University of Waterloo (Co-author, e-Energy 2025 paper)
 Fall 2024 Linda Wang, University of Waterloo (Co-author, e-Energy 2025 paper)
 Spring 2024 Amy Li, University of Waterloo (Co-author, HotCarbon 2024 paper)
 Spring 2024 Beihao Zhou, University of Waterloo (Co-author, HotCarbon 2024 paper)
 Spring 2024 Sophia Nguyen, University of Waterloo (Co-author, HotCarbon 2024 paper)
 2019-2020 Avinash Rao, University of Chicago (Co-author, MLSys 2022 paper)

Professional Service

Organizer

2026 Co-Chair, HARMONY Workshop on AI and Mental Health, co-located with IEEE/ACM CHASE 2026
 2026 Co-Chair, The 1st Workshop on Trustworthy and Adaptive LLMs for Mental and Physical Wellbeing in Recommendations, co-located with ACM UMAP 2026 (Co-Chair)
 2024 Co-Chair, NSF Workshop on Sustainable Computing: AI, Water, and Biodiversity

Invited Participant

2026 CRA Career Mentoring Workshop
 2026 DOE/AFOSR Energy Consequences of Information Workshop
 2025 Indiana Water Summit
 2025 CCC Computing Futures Symposium
 2025 CIFellows 2025 Symposium
 2025 Dagstuhl Perspectives Workshop: Climate Change: What is Computing's Responsibility?
 2024 NSF Workshop on Sustainable Computing for Sustainability
 2022 NITRD 30th Anniversary Symposium
 2022 CIFellows 2022 Workshop

Program Committee

2026 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)
 2026 ACM International Conference on Future and Sustainable Energy Systems (e-Energy)
 2026 ACM Intl. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS)
 2025 IEEE Computer Architecture Letters (CAL)
 2025 ACM Workshop on Sustainable Computer Systems (HotCarbon)
 2025 SIGOPS Asia-Pacific Workshop on Systems (APSys)
 2025 IEEE/ACM International Symposium on Computer Architecture (ISCA)
 2025 IEEE International Symposium on High-Performance Computer Architecture (HPCA)
 2024 USENIX Annual Technical Conference (ATC)
 2024 Conference on Systems and Machine Learning (MLSys)
 2023 ACM Student Research Competition at PACT
 2022 ACM SPLASH Onward!
 2022 Conference on Systems and Machine Learning (MLSys)
 2022 ACM Asia-Pacific Workshop on Systems
 2022 Journal of Systems Research

Technical Reviewing

2026 International Conference on Learning Representations (ICLR)
 2022 Neural Information Processing Systems (NeurIPS)
 2022 International Conference on Learning Representations (ICLR)
 2022 International Conference on Machine Learning (ICML)

- 2021 Neural Information Processing Systems (NeurIPS)
- 2021 AAAI Conference on Artificial Intelligence (AAAI)
- 2020 AAAI Conference on Artificial Intelligence (AAAI)
- 2019 Neural Information Processing Systems (NeurIPS)
- 2019 International Conference on Machine Learning (ICML)

Presentations

Invited Talks

- 2026 Systematic Characterization of LLM Quantization: A Performance, Energy, and Quality Perspective
 - DOE/AFOSR Energy Consequences of Information Workshop, Santa Fe, USA Feb. 2026
- 2025 Not All Water Consumption Is Equal: A Water Stress Weighted Metric for Sustainable Computing
 - Green Software Foundation (Virtual) Sep. 2025
- 2025 Not All Water Consumption Is Equal: An AI, Datacenter, and Semiconductor Perspective
 - Indiana Water Summit, Indianapolis, USA Aug. 2025
- 2024 Towards Sustainable Next Generation AI and Cloud Systems
 - Meta, Sunnyvale, USA Sep. 2024
- 2024 Uncertainty-Aware Carbon Optimization in Cloud Computing
 - SoDec Workshop at e-Energy, Singapore Jun. 2024
 - ASTAR Center for Frontier AI Research (CFAR), Singapore Jun. 2024
- 2023 A Holistic View on Machine Learning for Systems
 - University of Waterloo, Department of Computer Science Jun. 2023
 - Microsoft Research Apr. 2023
 - Texas A&M University, Department of Computer Science and Engineering Apr. 2023
 - University of Southern California, Department of Electrical and Computer Engineering Apr. 2023
 - University of Illinois, Department of Computer Science Mar. 2023
 - Cornell Tech, Department of Electrical and Computer Engineering Mar. 2023
 - Washington University in St. Louis, Department of Computer Science & Engineering Mar. 2023
 - Purdue University, Department of Electrical and Computer Engineering Mar. 2023
 - Purdue University, Department of Computer Science Mar. 2023
 - Virginia Tech, Department of Computer Science Mar. 2023
 - Indiana University Bloomington, Department of Computer Science Feb. 2023
 - University of Colorado Boulder, Department of Computer Science Feb. 2023
 - University of Massachusetts Amherst, College of Information and Computer Sciences Feb. 2023

Conference Presentations

- 2024 Sustainable LLM Serving: Environmental Implications, Challenges, and Opportunities
 - Conference presentation at IGSC, Austin, USA Oct. 2024
- 2024 Uncertainty-Aware Decarbonization for Datacenters
 - Conference presentation at HotCarbon, Santa Cruz, USA Jul. 2024
- 2022 NURD: Negative-Unlabeled Learning for Online Datacenter Straggler Prediction
 - Conference presentation at MLSys, Santa Clara, USA Aug. 2022
- 2022 Predictable Maintenance Job Planning in Datacenters
 - Meta Infrastructure Data Science Faculty Workshop at KDD, DC, USA Aug. 2022
- 2021 Generalizable and Interpretable Learning for Configuration Extrapolation
 - Conference presentation at ESEC/FSE, Virtual Nov. 2021
- 2020 Dynamical Systems Theory for Causal Inference with Application to Synthetic Controls
 - Causal Data Science Meeting, Virtual Nov. 2020
 - Conference presentation at AISTATS, Virtual Aug. 2020
- 2019 Generative and Multi-phase Learning for Computer Systems Optimization
 - Conference presentation at ISCA, Phoenix, USA Jun. 2019
- 2017 Multiresolution Kernel Approximation for Gaussian Process Regression

- Conference presentation at NeurIPS, Long Beach, USA Dec. 2017
- 2017 Large Scale Kernel Methods for Online AUC Maximization
- Conference presentation at ICDM, New Orleans, USA Nov. 2017
- 2015 An Adaptive Gradient Method for Online AUC Maximization
- Conference presentation at AAAI, Austin, USA Jan. 2015

Media Coverage

- 1/21/2026 [The Hidden Cost of AI: Making Data Centers Sustainable.](#) Purdue ECE Podcast
- 11/24/2025 [How to make data centers less thirsty.](#) Grist
- 10/15/2025 [Purdue ECE research reveals how computing impacts global biodiversity.](#) Purdue ECE News
- 10/02/2025 [Economic boom or environmental disaster? Rural Texas grapples with pros, cons of data centers.](#) The Texas Tribune
- 09/25/2025 [Data centers are thirsty for Texas' water, but state planners don't know how much they will need.](#) The Texas Tribune
- 09/16/2025 [Spain's data centre law: supporting local groups in the public consultation.](#) The Green Web Foundation
- 08/15/2025 [Big Tech's big thirst — AI's demand for Texas water.](#) Texas Public Radio
- 07/15/2025 [Companies focus on ways of achieving energy efficiency as consumption keeps increasing.](#) New Zealand Herald
- 07/14/2025 [Tech Giants Scramble To Meet AI's Looming Energy Crisis.](#) Barron's

Last updated February 27, 2026.