

Jeff M. Phillips

Associate Professor | School of Computing | University of Utah

Director, Utah Center for Data Science

Director, Data Science Program

50 S Central Campus Dr., Salt Lake City, UT 84112 | (801) 585-7775

<http://www.cs.utah.edu/~jeffp> | jeffp@cs.utah.edu

Education

Duke University

Ph.D. in Computer Science, January, 2009.

Thesis Title: *Small and Stable Descriptors of Distributions for Geometric Statistical Problems.*

Advisor: Pankaj K. Agarwal.

Rice University

Bachelor of Science in Computer Science, May 2003.

Bachelor of Arts in Mathematics, May 2003.

Academic and Research Positions

| | |
|--|----------------------|
| University of Utah, School of Computing (Associate Professor) | (2017-present) |
| Utah Center for Data Science (Director) | (2019 - present) |
| Simons Institute for Theoretical Computer Science, UC Berkeley (visiting fellow) | (Fall 2018) |
| University of Utah, School of Computing (Assistant Professor) | (2011-2017) |
| University of Utah, School of Computing (Postdoctoral CI Fellow) | (2009-2011) |
| Duke University, Department of Computer Science (Postdoctoral Associate) | (2009) |
| Duke University, Department of Computer Science (Research Assistant) | (2003-2009) |
| Yahoo! Research (Research Intern) | (Summer 2007) |
| AT&T Research (Visiting Researcher) | (Summer/Winter 2005) |
| Rice University, Department of Computer Science (Research Assistant) | (2000-2003) |
| The Charles Stark Draper Laboratory, Inc. (Research Scientist) | (2002-2003) |

Fellowships and Awards

| | |
|--|-------------|
| Alberto Mendelzon PODS 2022 Test of Time Award for most influential paper from PODS 2012. (2022) Mergeable Summaries. Agarwal, Cormode, Huang, Phillips, Wei, and Yi. <i>PODS 2012</i> | |
| Utah College of Engineering, Outstanding Service Award. | (2021) |
| NSF CAREER Award. | (2014) |
| Best Paper Award at MultiClust Workshop: Discovering, Summarizing and Using Multiple Clusterings. | (2011) |
| CCC-CRA-NSF Computing Innovation Fellowship. | (2009) |
| 2 year postdoctoral fellowship — 60 awarded among all graduating computer scientists in US | |
| Best Student Paper at International Conference on Automata, Languages and Programming (ICALP). | (2008) |
| Distinguished Department Service Award (Duke Computer Science). | (2008) |
| For 5 years of department service — never before awarded | |
| Outstanding Department Service Award (Duke Computer Science). | (2006) |
| NSF Graduate Research Fellowship. | (2004-2007) |
| 3 year full graduate fellowship | |
| James B. Duke Fellowship. | (2003-2007) |
| 4 years partial Duke graduate fellowship | |
| C. S. Draper Laboratory Fellowship. | (2003-2007) |
| 4 year full graduate fellowship — declined | |
| James S. Waters Creativity Award. | (2002) |

Annual Rice Undergraduate Research Award

NASA/Texas Space Grant Consortium Undergraduate Scholarship.

(2002)

Annual Texas Undergraduate Research Award

Brown Undergraduate Research Internship Award.

(May 2001, September 2001)

Rice Undergraduate Research Award, won twice

RESEARCH¹

Books

- [B1] Mathematical Foundations for Data Analysis.
 Jeff M. Phillips.
Springer-Nature. 2021. (mathfordata.github.io)

Book Chapters and Surveys

- [S1] Coresets and Sketches.
 Jeff M. Phillips.
Handbook of Discrete and Computational Geometry, 3rd edition, CRC Press, Chapter 48. 2016.
arXiv:1601.00617, February 2016.
- [S2] A Gentle Introduction to the Kernel Distance.
 Jeff M. Phillips and Suresh Venkatasubramanian.
arXiv:1103.1625, March 2011.

Journal Publications

- [J1] An Experimental Study On Classifying Spatial Trajectories.
 Hasan Pourm Mahmood-aghababa and Jeff M. Phillips.
Knowledge and Information Systems (KAIS). December 2022.
- [J2] Local Kernel Ridge Regression for Scalable, Interpolating, Continuous Regression.
 Mingxuan Han, Chenglong Ye, and Jeff M. Phillips.
Transactions on Machine Learning Research (TMLR). October 2022.
- [J3] Capturing Intent behind Selection In Scatterplot Visualizations.
 Kiran Gadhawe, Jochen Gortler, Zach Cutler, Carolina Nobre, Oliver Deussen, Miriah Meyer, Jeff M. Phillips, and Alexander Lex.
Information Visualization. August 2021.
- [J4] Practical and Configurable Network Traffic Classification Using Probabilistic Machine Learning.
 Jiahui Chen, Joe Breen, Jeff M. Phillips, Jacobus Van der Merwe.
Cluster Computing, accepted August 2021.
 (DOI 10.1007/s10586-021-03393-2)
- [J5] Efficient Oblivious Query Processing for Range and kNN Queries.
 Zhao Chang, Dong Xie, Feifei Li, Jeff M. Phillips, and Rajeev Balasubramanian.
Transactions on Knowledge and Data Engineering (TKDE). accepted February 2021.
- [J6] Semantic Embedding for Regions of Interest.
 Debjyoti Paul, Jeff M. Phillips, and Feifei Li.
Very Large Data Bases Journal (VLDBJ). February 2021.
 (<https://doi.org/10.1007/s00778-020-00647-0>)

¹Utah students underlined

- [J7] Inferencing Hourly Traffic Volume using Data-Driven Machine Learning and Graph Theory.
Zhiyan Yi, Xiaoyue Cathy Liu, Nikola Markovic, and Jeff M. Phillips.
Computers, Environment and Urban Systems, Vol 85. January 2021.
- [J8] Scalable Spatial Scan Statistics for Trajectories.
Michael Matheny, Dong Xie, and Jeff M. Phillips.
ACM Transactions on Knowledge Discovery from Data (TKDD) 14(6) no. 73. September 2020.
arXiv:1906.01693. June 2019.
- [J9] Closed Form Word Embedding Alignment.
Sunipa Dev, Saffia Hassan, and Jeff M. Phillips.
Knowledge and Information Systems (KAIS) 63, 867–887, 2020.
arXiv:1806.01330. June 2018.
- [J10] Near-Optimal Coresets of Kernel Density Estimates.
Jeff M. Phillips and Wai Ming Tai.
Discrete & Computational Geometry (DCG) 63, 867–887, 2020.
arxiv:1802.01751. April 2019.
- [J11] An Integrated Classification Scheme for Mapping Estimates and Errors of Estimation from the American Community Survey.
Ran Wei, Daoqin Tong, and Jeff M. Phillips.
Computers, Environment and Urban Systems (CEUS) 63, 95–103, 2017.
- [J12] Frequent Directions: Simple and Deterministic Matrix Sketching.
Mina Ghashami, Edo Liberty, Jeff M. Phillips and David P. Woodruff.
SIAM Journal of Computing (SICOMP) 45:5, 2016.
arXiv:1501.01711. January 2015.
- [J13] Nearest Neighbor Searching Under Uncertainty II.
Pankaj K. Agarwal, Boris Aronov, Sariel Har-Peled, Jeff M. Phillips, Ke Yi, Wuzhou Zhang.
ACM Transactions on Algorithms (TALG) 13:1 2016.
arXiv:1606.00112. June 2016.
- [J14] Improved Practical Matrix Sketching with Guarantees.
Ameey Desai, Mina Ghashami, and Jeff M. Phillips.
Transactions on Knowledge and Data Engineering (TKDE) 28:07, 1678–1690, 2016.
- [J15] Lower Bounds for Number-in-Hand Multiparty Communication Complexity, Made Easy.
Jeff M. Phillips, Elad Verbin, and Qin Zhang.
SIAM Journal of Computing (SICOMP) 45:1, 174–196, 2016.
- [J16] Mergeable Summaries.
Pankaj K. Agarwal, Graham Cormode, Zengfeng Huang, Jeff M. Phillips, Zhewei Wei, and Ke Yi.
invited to *ACM Transactions on Database Systems (TODS)* 38:26, 2013.
- [J17] (Approximate) Uncertain Skylines.
Peyman Afshani, Pankaj K. Agarwal, Lars Arge, Kasper Green Larsen, and Jeff M. Phillips.
Theory of Computing Systems (TOCS) 52, 342–366, 2013. (Special Issue: ICDT 2011)

Conference Publications

- [C1] Interpretable Debiasing of Vectorized Language Representations with Iterative Orthogonalization.
Prince Osei Aboagye, Yan Zheng, Jack Shunn, Chin-Chia Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen, Liang Wang, Wei Zhang, and Jeff Phillips.
International Conference on Learning Representations (ICLR). April 2023.
- [C2] Batch Multi-Fidelity Active Learning with Budget Constraints.
Shibo Li, Jeff M. Phillips, Xin Yu, Robert M. Kirby, Shandian Zhe.
Neural Information Processing Systems (NeurIPS). December 2022.

- [C3] Quantized Wasserstein Procrustes Alignment of Word Embedding Spaces.
Prince Osei Aboagye, Yan Zheng, Chin-Chia Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen,
 Liang Wang, Wei Zhang, Jeff M. Phillips.
Association for Machine Translation in the Americas (AMTA). September 2022.
- [C4] Using Existential Theory of the Reals to Bound VC Dimension.
Austin Watkins and Jeff M. Phillips.
Canadian Conference on Computational Geometry (CCCG). August 2022.
- [C5] Normalization of Language Embeddings for Cross-Lingual Alignment.
Prince Osei Aboagye, Yan Zheng, Chin-Chia Michael Yeh, Junpeng Wang, Wei Zhang, Liang Wang, Hao Yang,
 and Jeff M. Phillips.
International Conference on Learning Representations (ICLR). April 2022.
- [C6] Self-Adaptable Point Processes with Nonparametric Time Decays.
Zhimeng Pan, Zheng Wang, Jeff M. Phillips, and Shandian Zhe.
Neural Information Processing Systems (NeurIPS). December 2021.
- [C7] Approximate Maximum Halfspace Discrepancy.
Michael Matheny and Jeff M. Phillips.
International Symposium on Algorithms and Computation (ISAAC). December 2021.
arXiv:2106.13851. June 2021.
- [C8] Harms of Gender Exclusivity and Challenges in Non-Binary Representation in Language Technologies.
 Sunipa Dev, Masoud Manajatipoor, Anaelia Ovalle, Arjun Subramonian, Jeff M. Phillips, and Kai-Wei Chang.
Conference on Empirical Methods in Natural Language Processing (EMNLP). November 2021.
arXiv:2108.12084. August 2021.
- [C9] OSCaR: Orthogonal Subspace Correction and Rectification of Biases in Word Embeddings.
Sunipa Dev, Tao Li, Jeff M Phillips, and Vivek Srikumar.
Conference on Empirical Methods in Natural Language Processing (EMNLP). November 2021.
arXiv:2007.00049. July 2020.
- [C10] Constrained Non-Affine Alignment of Embeddings.
Yuwei Wang, Yan Zheng, Yanqing Peng, Michael Yeh, Zhongfang Zhuang, Das Mahashweta, Bendre Mangesh,
 Feifei Li, Wei Zhang, and Jeff M. Phillips.
International Conference on Data Mining (ICDM). December 2021.
arXiv:1910.05862. September 2021.
- [C11] Finding an Approximate Mode of a Kernel Density Estimate.
 Jasper C.H. Lee, Jerry Li, Christopher Musco, Jeff M. Phillips, and Wai Ming Tai.
European Symposium on Algorithms (ESA). September 2021.
arXiv:1912.07673. December 2019.
- [C12] Orientation-Preserving Vectorized Distance Between Curves.
 Jeff M Phillips and Hasan Pourmahmood-Aghababa.
Mathematical and Scientific Machine Learning (MSML). August 2021.
arXiv:2007.15924. July 2020.
- [C13] A Deterministic Streaming Sketch for Ridge Regression.
Benwei Shi and Jeff M. Phillips.
International Conference on Artificial Intelligence and Statistics (AISTats). April 2021.
arXiv:2002.02013. February 2020.
- [C14] Spatial Independent Range Sampling.
 Dong Xie, Jeff M. Phillips, Michael Matheny, and Feifei Li.
ACM Symposium on Management of Data (SIGMOD). June 2021.
- [C15] At-the-time and Back-in-time Persistent Sketches.
 Benwei Shi, Zhuoyue Zhao, Yanqing Peng, Feifei Li, and Jeff M. Phillips.
ACM Symposium on Management of Data (SIGMOD). June 2021.

- [C16] The GaussianSketch for Almost Relative Error Kernel Distance.
Jeff M. Phillips and Wai Ming Tai.
International Conference on Randomization and Computation (RANDOM). August 2020.
arXiv:1811.04136. December 2019.
- [C17] Sketched MinDist.
Jeff M. Phillips and Pingfan Tang.
International Symposium on Computational Geometry (SoCG). June 2020.
arXiv:1907.02171, July 2019.
- [C18] On Measuring and Mitigating Biased Inferences of Word Embeddings.
Sunipa Dev, Tao Li, Jeff M. Phillips and Vivek Srikumar.
AAAI Conference on Artificial Intelligence (AAAI). February 2020.
arXiv:1908.09369, August 2019.
- [C19] Simple Distances for Trajectories via Landmarks.
Jeff M. Phillips and Pingfan Tang.
ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL). November 2019.
arXiv:1804.11284, June 2019.
- [C20] The Kernel Spatial Scan Statistic.
Mingxuan Han, Michael Matheny, and Jeff M. Phillips.
ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL). November 2019.
arXiv:1906.09381, June 2019.
- [C21] Closed Form Word Embedding Alignment.
Sunipa Dev and Jeff M. Phillips.
IEEE International Conference on Data Mining (ICDM). November 2019.
arXiv:1806.01330, September 2018.
- [C22] On the VC Dimension of Metric Balls under Frechet and Hausdorff Distances.
Anne Driemel, Jeff M. Phillips, Ioannis Psarros.
International Symposium on Computational Geometry (SoCG). June 2019.
arXiv:1903.03211, March 2019.
- [C23] Independent Range Sampling, Revisited Again.
Peyman Afshani and Jeff M. Phillips.
International Symposium on Computational Geometry (SoCG). June 2019.
- [C24] Attenuating Bias in Word Vectors.
Sunipa Dev and Jeff M. Phillips.
International Conference on Artificial Intelligence and Statistics (AISTats). April 2019.
arXiv:1901.07656, January 2019.
- [C25] Computing Approximate Statistical Discrepancy.
Michael Matheny and Jeff M. Phillips.
International Symposium on Algorithm and Computation (ISAAC). December 2018.
arXiv:1804.11287, April 2018.
- [C26] Improved Bounds on Information Dissemination by Manhattan Random Waypoint Model.
Aria Rezaei, Jie Gao, Jeff M. Phillips, and Csaba D. Toth.
ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL). November 2018.
arXiv:1809.07392, September 2018.
- [C27] Practical Low-Dimensional Halfspace Range Space Sampling.
Michael Matheny and Jeff M. Phillips.
European Symposium on Algorithms (ESA). September 2018.
arXiv:1804.11307, April 2018.

- [C28] Approximating the Distribution of the Median and other Robust Estimators on Uncertain Data.
Kevin Buchin, Jeff M. Phillips and Pingfan Tang.
International Symposium on Computational Geometry (SOCG). June 2018.
arxiv:1601.00630, January 2016.
- [C29] Near-Optimal Coresets of Kernel Density Estimates.
Jeff M. Phillips and Wai Ming Tai.
International Symposium on Computational Geometry (SOCG). June 2018.
arxiv:1802.01751, February 2018.
- [C30] Fully Convolutional Structured LSTM Networks for Joint 4D Medical Image Segmentation.
Yang Gao, Jeff M. Phillips, Yan Zheng, Renqiang Min, P. Thomas Fletcher, and Guido Gerig.
IEEE International Symposium on Biomedical Imaging (ISBI). April 2018.
- [C31] Improved Coresets for Kernel Density Estimates.
Jeff M. Phillips and Wai Ming Tai.
ACM-SIAM Symposium on Discrete Algorithms (SODA). January 2018.
arxiv:1710.04325, October 2017.
- [C32] Relative Error Embeddings for the Gaussian Kernel Distance.
Di Chen and Jeff M. Phillips.
Algorithmic Learning Theory (ALT), October 2017.
arxiv:1602.05350, February 2016.
- [C33] Distributed Trajectory Similarity Search.
Dong Xie, Feifei Li, and Jeff M. Phillips.
International Conference on Very Large Databases (VLDB), August 2017.
- [C34] Coresets for Kernel Regression.
Yan Zheng and Jeff M. Phillips. *ACM Conference on Knowledge Discovery and Data Mining (KDD)*, August 2017.
arxiv:1702.03644, February 2017.
- [C35] Scalable Spatial Scan Statistics through Sampling.
Michael Matheny, Raghvendra Singh, Kaiqiang Wang, Liang Zhang and Jeff M. Phillips.
ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL), [18%] November 2016.
- [C36] The Robustness of Estimator Composition.
Pingfan Tang and Jeff M. Phillips.
Conference on Neural Information Processing (NIPS), [24%] December 2016.
- [C37] epsilon-Kernel Coresets for Stochastic Points.
Lingxiao Huang, Jian Li, Jeff M. Phillips, Haitao Wang.
European Symposium on Algorithms (ESA), [27%] August 2016.
arXiv:1411.0194, April 2015.
- [C38] Efficient Frequent Directions Algorithm for Sparse Matrices.
Mina Ghashami and Edo Liberty, and Jeff M. Phillips.
ACM Conference on Knowledge Discovery and Data Mining (KDD), [18%] August 2016.
arXiv:1602.00412, February 2016.
- [C39] Streaming Kernel Principal Component Analysis.
Mina Ghashami and Daniel Perry, and Jeff M. Phillips.
International Conference on Artificial Intelligence and Statistics (AISTATS), [31%] May 2016.
arXiv:1512.05059, December 2015.
- [C40] Subsampling in Smooth Range Spaces.
Jeff M. Phillips and Yan Zheng.
Algorithmic Learning Theory (ALT), [50%] October 2015.

- [C41] L_∞ Error and Bandwidth Selection for Kernel Density Estimates of Large Data.
 Jeff M. Phillips and Yan Zheng.
ACM Conference on Knowledge Discovery and Data Mining (KDD), [19%] August 2015.
- [C42] Geometric Inference on Kernel Density Estimates.
 Jeff M. Phillips, Bei Wang, and Yan Zheng.
International Symposium on Computational Geometry (SoCG), [38%] June 2015.
arXiv:1307.7760, January 2014.
- [C43] Improved Practical Matrix Sketching with Guarantees.
 Mina Ghashami, Amey Desai, and Jeff M. Phillips.
22nd Annual European Symposium on Algorithms (ESA), [25%] September 2014.
 extended version as *arXiv:1501.06561*, January 2015.
- [C44] Continuous Matrix Approximation on Distributed Data.
 Mina Ghashami, Jeff M. Phillips, and Feifei Li.
40th International Conference on Very Large Data Bases (VLDB), [22%] September 2014.
- [C45] Relative Errors for Deterministic Low-Rank Matrix Approximations.
 Mina Ghashami and Jeff M. Phillips.
25th Annual ACM-SIAM Symposium on Discrete Algorithms (SoDA), [28%] January 2014.
arXiv:1307.7454.
- [C46] Quality and Efficiency for Kernel Density Estimates in Large Data.
 Yan Zheng, Jeffrey Jests, Jeff M. Phillips, and Feifei Li.
ACM Conference on Management of Data (SIGMOD), [20%] June 2013.
- [C47] Nearest Neighbor Searching Under Uncertainty II.
 Pankaj K. Agarwal, Boris Aronov, Sarel Har-Peled, Jeff M. Phillips, Ke Yi, Wuzhou Zhang.
32nd ACM Symposium on Principles of Database Systems (PODS), [25%] June 2013.
- [C48] Range Counting Coresets for Uncertain Data.
 Amirali Abdullah, Samira Daruki, Jeff M. Phillips.
29th Annual ACM Symposium on Computational Geometry (SoCG), [35%] June 2013.
arXiv:1304.4243.
- [C49] Radio Tomographic Imaging and Tracking of Stationary and Moving People via Kernel Distance.
 Yang Zhao, Neal Patwari, Jeff M. Phillips, and Suresh Venkatasubramanian.
12th ACM-IEEE Conference on Information Processing in Sensor Networks (IPSN), [21%] April 2013.
- [C50] ϵ -Samples for Kernels.
 Jeff M. Phillips.
24th Annual ACM-SIAM Symposium on Discrete Algorithms (SoDA), [30%] January 2013.
arXiv:1112.4105.
- [C51] Efficient Protocols for Distributed Classification and Optimization.
 Hal Daume III, Jeff M. Phillips, Avishek Saha, and Suresh Venkatasubramanian.
23rd International Conference on Algorithmic Learning Theory (ALT), [49%] October 2012.
arXiv:1204.3523.
- [C52] Ranking Large Temporal Data.
 Jeffrey Jests, Jeff M. Phillips, Feifei Li, and Mingwang Tang.
38st International Conference on Very Large Databases (VLDB), [20%] August 2012.
arXiv:1208.0222.
PVLDB 5:1412-1423, 2012.
- [C53] Mergeable Summaries.
 Pankaj K. Agarwal, Graham Cormode, Zengfeng Huang, Jeff M. Phillips, Zhewei Wei, and Ke Yi.
31st ACM Symposium on Principles of Database Systems (PODS), [26%] May 2012.
 invited to *ACM Transactions on Database Systems (TODS)*.

- [C54] Protocols for Learning Classifiers on Distributed Data.
Hal Daume III, Jeff M. Phillips, Avishek Saha, and Suresh Venkatasubramanian.
15th International Conference on Artificial Intelligence and Statistics (AISTATS), [30%] April 2012.
- [C55] Efficient Threshold Monitoring for Distributed Probabilistic Data.
Mingwang Tang, Feifei Li, Jeff M. Phillips, and Jeffrey Jests.
28th IEEE International Conference on Data Engineering (ICDE), [24%] April 2012.
- [C56] Uncertainty Visualization in HARDI based on Ensembles of ODFs.
Fangxiang Jiao, Jeff M. Phillips, Yaniv Gur, and Chris R. Johnson.
5th IEEE Pacific Visualization Symposium (PacificVis), [34%] February 2012.
- [C57] Lower Bounds for Number-in-Hand Multiparty Communication Complexity, Made Easy.
Jeff M. Phillips, Elad Verbin, and Qin Zhang.
23th Annual ACM-SIAM Symposium on Discrete Algorithms (SoDA), [31%] January 2012.
arXiv:1107.2559.
- [C58] Geometric Computation on Indecisive Points.
Allan G. Jørgensen, Maarten Löffler, and Jeff M. Phillips.
Algorithms and Data Structures Symposium (WADS), [42%] August 2011.
- [C59] Computing Hulls, Centerpoints, and VC-Dimension in Positive Definite Space.
P. Thomas Fletcher, John Moeller, Jeff M. Phillips, and Suresh Venkatasubramanian.
Algorithms and Data Structures Symposium (WADS), [42%] August 2011.
arXiv:0912.1580.
- [C60] Comparing Distributions and Shapes Using the Kernel Distance.
Sarang Joshi, Raj Varma Kommaraju, Jeff M. Phillips, and Suresh Venkatasubramanian.
ACM Symposium on Computational Geometry (SoCG), [39%] June 2011.
arXiv:1001.0591.
- [C61] Spatially-Aware Comparison and Consensus for Clusterings.
Jeff M. Phillips, Parasaran Raman, and Suresh Venkatasubramanian.
SIAM International Conference on Data Mining (SDM), [25%] April 2011.
arXiv:1102.0026.
- [C62] (Approximate) Uncertain Skylines.
Peyman Afshani, Pankaj K. Agarwal, Lars Arge, Kasper Dalgaard Larsen, and Jeff M. Phillips.
14th International Conference on Database Theory (ICDT), [41%] March 2011.
invited to *Theory of Computing Systems (TOCS)* (Special Issue: ICDT 2011)
- [C63] Stability of ϵ -Kernels.
Pankaj K. Agarwal, Jeff M. Phillips, Hai Yu.
18th Annual European Symposium on Algorithms (ESA), [27%] September 2010.
- [C64] A Unified Algorithmic Framework for Multi-Dimensional Scaling.
Arvind Agarwal, Jeff M. Phillips, and Suresh Venkatasubramanian.
16th Annual ACM International Conference on Knowledge Discovery and Data Mining (KDD), [17%] July 2010.
- [C65] Lipschitz Unimodal and Isotonic Regression on Paths and Trees.
Pankaj K. Agarwal, Jeff M. Phillips, and Bardia Sadri.
9th Latin American Theoretical Informatics Symposium (LATIN), [36%] April 2010.
- [C66] Shape Fitting on Point Sets with Probability Distributions.
Maarten Löffler and Jeff M. Phillips.
17th Annual European Symposium on Algorithms (ESA), [25%] September 2009.
- [C67] An Efficient Algorithm for Euclidean 2-Center with Outliers.
Pankaj K. Agarwal and Jeff M. Phillips.
16th Annual European Symposium on Algorithms (ESA), [35%] September 2008.

- [C68] Algorithms for ε -Approximations of Terrains. **Best Student Paper**.
Jeff M. Phillips.
International Colloquium on Automata, Languages and Programming (ICALP), [24%] July 2008.
- [C69] Spatial Scan Statistics for Graph Clustering.
Bei Wang, Jeff M. Phillips, Robert Schreiber, Dennis Wilkinson, Nina Mishra, and Robert E. Tarjan.
SIAM International Conference on Data Mining (SDM), [14%] April 2008.
- [C70] Value-Based Notification Conditions in Large Publish/Subscribe Systems.
Badrish Chandramoulli, Jeff M. Phillips, and Jun Yang.
International Conference on Very Large Data Bases (VLDB), [17%] September 2007.
- [C71] Outlier Robust ICP for Minimizing Fractional RMSD.
Jeff M. Phillips, Ran Liu, and Carlo Tomasi.
International Conference on 3-D Digital Imaging and Modeling (3DIM), [38%] August 2007.
Poster/abstract in *Eurographics Symposium on Geometric Processing (SGP)*, June 2006.
- [C72] Segmenting Motifs in Protein-Protein Interface Surfaces.
Jeff M. Phillips Johannes Rudolph, and Pankaj K. Agarwal.
Workshop on Algorithms in Bioinformatics (WABI), [36%] September 2006.
- [C73] Spatial Scan Statistics: Approximations and Performance Study.
Deepak Agarwal, Andrew McGregor, Jeff M. Phillips, Suresh Venkatasubramanian, and Zhengyuan Zhu.
ACM SIGKDD International Conference on Knowledge Discovery and Data (KDD), [11%] August 2006.
- [C74] On Bipartite Matching under the RMS Distance.
Pankaj K. Agarwal and Jeff M. Phillips.
Canadian Conference on Computational Geometry (CCCG), August 2006.
- [C75] Hunting of the Bump: On Maximizing Statistical Discrepancy.
Deepak Agarwal, Jeff M. Phillips, and Suresh Venkatasubramanian.
SIAM-ACM Symposium on Discrete Algorithms (SoDA), [31%] January 2006.
- [C76] Guided Expansive Spaces Trees: A Search Technique for Motion- and Cost-Constrained State Spaces.
Jeff M. Phillips, Nazareth Bedrossian, and Lydia E. Kavvaki.
IEEE International Conference on Robotics and Automation (ICRA), [59%] April 2004.
- [C77] Simulated Knot Tying.
Jeff M. Phillips, Andrew Ladd, and Lydia E. Kavvaki.
IEEE International Conference on Robotics and Automation (ICRA), [59%] May 2002.

Tutorials

- [T1] A Visual Tour of Bias Mitigation for Word Representations.
Archit Rathore, Sunipa Dev, Jeff M. Phillips, Vivek Srikumar, Bei Wang.
 - *AAAI Conference on Artificial Intelligence*, February 2021.
 - *ACM Conference on Knowledge Discovery and Data Mining*, September 2021.

Workshop Publications

- [W1] Toward Classifying Unknown Application Traffic.
Ryan Baker, Ren Quinn, Jeff M. Phillips, and Jacobus (Kobus) Van der Merwe. *DYNAMIC and Novel Advances in Machine Learning and Intelligent Cyber Security (DYNAMICS) Workshop*. December 2018.
- [W2] Visualization of Big Spatial Data using Coresets for Kernel Density Estimates.
Yan Zheng, Yi Ou, Alexander Lex, and Jeff M. Phillips. *Visual Data Science (VDS)*. October 2017.
- [W3] Visualizing Sensor Network Coverage with Location Uncertainty.
Tim Sodergren, Jessica Hair, Jeff M. Phillips, and Bei Wang. *Visual Data Science (VDS)*. October 2017.

- [W4] Detecting and Localizing Spectrum Offenders Using Crowdsourcing.
Sneha Kasera, Neal Patwari, and Jeff M. Phillips.
IEEE Dynamic Spectrum Access Networks : Future Directions in Spectrum Management Research, October 2015.
- [W5] Subsampling in Smooth Range Spaces.
Jeff M. Phillips and Yan Zheng.
Computational Geometry : Young Researchers Forum, June 2015.
- [W6] Continuous Matrix Approximation on Distributed Data.
Mina Ghashami, Jeff M. Phillips, and Feifei Li.
6th Workshop on Massive Data Algorithmics (MASSIVE), September 2014.
- [W7] Sensor Network Localization for Moving Sensors.
Arvind Agarwal, Hal Daume III, Jeff M. Phillips, Avishek Saha, and Suresh Venkatasubramanian.
IEEE ICDM International Workshop on Data Mining in Networks (DaMNet), December 2012.
- [W8] Kernel Distance for Geometric Inference.
Jeff M. Phillips and Bei Wang.
22th Fall Workshop on Computational Geometry (FWCG), October 2012.
arXiv:1307.7760, July 2013.
- [W9] Generating A Diverse Set Of High-Quality Clusterings. **Best Paper Award.**
Jeff M. Phillips, Parasaran Raman, Suresh Venkatasubramanian.
MultiClust Workshop: Discovering, Summarizing and Using Multiple Clusterings (MULTICLUST), September 2011.
- [W10] Mergeable Coresets.
Pankaj K. Agarwal, Graham Cormode, Zengfeng Huang, Jeff M. Phillips, Zhewei Wei, and Ke Yi.
3rd Workshop on Massive Data Algorithmics (MASSIVE), June 2011.
- [W11] Johnson-Lindenstrauss Dimensionality Reduction on the Simplex.
Rasmus J. Kyng, Jeff M. Phillips, and Suresh Venkatasubramanian.
20th Fall Workshop on Computational Geometry (FWCG), October 2010.
- [W12] Metrics for Uncertainty Analysis and Visualization of Diffusion Tensor Images.
Fangxiang Jiao, Jeff M. Phillips, Jeroen Stinstra, Jens Krueger, Raj Varma Kummaraju, Edward Hsu, Julie Korenberg, Chris R. Johnson.
5th International Workshop on Medical Imaging and Augmented Reality (MIAR), September 2010.
- [W13] Incremental Multi-Dimensional Scaling.
Arvind Agarwal, Jeff M. Phillips, Hal Daume III, Suresh Venkatasubramanian.
The Learning Workshop at Snowbird, April 2010.
- [W14] The Hunting of the Bump: On Maximizing Statistical Discrepancy.
Deepak Agarwal, Jeff M. Phillips, Suresh Venkatasubramanian.
Fall Workshop on Computational Geometry (FWCG), November 2005.
- [W15] Spacecraft Rendezvous and Docking with Real-Time, Randomized Optimization.
Jeff M. Phillips, Lydia E. Kavraki, and Nazareth Bedrossian.
AIAA Guidance, Navigation, and Control, August 2003.
- [W16] Probabilistic Optimization Applied to Spacecraft Rendezvous and Docking.
Jeff M. Phillips, Lydia E. Kavraki, and Nazareth Bedrossian.
AAS/AIAA Space Flight Mechanics Meeting, February 2003.

Online Expository Articles and Preprints

- [M1] Rethinking Abstractions in Big Data: Why, Where, How, and What.
 Mary Hall, Robert M. Kirby, Feifei Li, Miriah Meyer, Valerio Pascucci, Jeff M. Phillips, Rob Ricci, Jacobus Van der Merwe, Suresh Venkatasubramanian.
University of Utah, School of Computing, Tech Report: UUCS-13-002, April 2013.
arXiv:1306.3295, June, 2013.
- [M2] Chernoff-Hoeffding Inequality and Applications.
 Jeff M. Phillips.
arXiv:1209.6396, September 2012.
-

Contributed/Invited Talks

Conference Presentations

| | |
|--|----------------------------|
| International Symposium on Algorithms and Computation | (2021) |
| European Symposium on Algorithms. | (2008,2009,2010,2014,2021) |
| Symposium on Computational Geometry. | (2011,2015,2018,2019,2020) |
| International Conference on Randomization and Computation. | (2020) |
| Visual Data Science. | (2017) |
| ACM-SIAM Symposium on Discrete Algorithms. | (2006,2013) |
| International Conference on Algorithmic Learning Theory. | (2012) |
| Yaroslavl International Conference on Discrete Geometry dedicated to A.D. Alexanderov. | (2012) |
| ACM Symposium on Principals of Database Systems. | (2012) |
| Algorithms and Data Structures Symposium. | (2011) |
| Latin American Theoretical Informatics Symposium. | (2010) |
| International Colloquium on Automata, Languages, and Programming. | (2008) |
| International Conference on 3-D Digital Imaging and Modeling. | (2007) |
| Workshop on Algorithms in Bioinformatics. | (2006) |
| ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. | (2006) |
| Canadian Conference on Computational Geometry. | (2006) |
| AIAA Guidance, Navigation, and Control. | (2003) |
| AAS/AIAA Space Flight Mechanics Meeting. | (2003) |
| IEEE International Conference on Robotics and Automation. | (2002,2004) |

Invited Talks

| | |
|--|-----------------------|
| Hausdorff School for Advanced Math. Summer School on <i>Algorithmic Data Analysis</i> . Bonn, Germany. | (2022) |
| BYU Computer Science Distinguished Speakers Series, Provo, UT. | (2022) |
| University of Michigan, Ann Arbor, MI. | (2021) |
| Dixie Forum, St. George, UT. | (2021) |
| JMM: AMS Session on Foundations of Data Science, Denver, CO. | (2021) |
| MADALGO, University of Aarhus, Denmark. | (2008,2010,2012,2018) |
| Simons Worksop on Sublinear Algorithms and Nearest-Neighbor Search, Berkeley, CA. | (2018) |
| Visa Research, Palo Alto, CA. | (2018) |
| Sichuan University, Chengdu, China | (2017) |
| Workshop on Geometric Computing on Uncertain Data (CG Week), Boston, MA. | (2016) |
| Oregon State University, Corvallis, OR. | (2016) |
| University of Houston, Houston, TX. | (2016) |
| Arizona State University, Phoenix, AZ. | (2016) |
| Dagstuhl–Computational Geometry, Germany (invited plenary on Geometric Data Analysis). | (2015) |
| The Ohio State University, Columbus, OH. | (2010,2014) |
| CUHK Workshop on Theoretical Aspects of Big Data, Hong Kong, China. | (2013) |
| Tulane University, New Orleans, LA. | (2013) |
| INRIA–Saclay, Saclay, France. | (2012) |

| | |
|---|-------------|
| Rice University, Houston, TX. | (2012) |
| NII Shonan Meeting, Japan (on Large-scale Distributed Computation). | (2012) |
| JMM: SIAM Minisymposium on Computational Geometry, Boston, MA. | (2012) |
| Peking University, Beijing, China. | (2011) |
| Yahoo! Labs, Santa Clara, CA. | (2011) |
| University of Utah, Salt Lake City, UT. | (2009,2011) |
| Texas A&M University, College Station, TX. | (2011) |
| Duke University, Durham, NC. | (2010) |
| Institute for Science and Technology: Austria. | (2009) |
| Statistical and Applied Mathematical Sciences Institute, RTP, NC. | (2006) |
| AT&T: Shannon Labs, Florham Park, NJ. | (2005) |
| NASA, Johnson Space Center, TX. | (2003) |
| Draper Laboratories, Cambridge, MA. | (2002) |

Students

Current Students Supervising

| | |
|--|----------------------------|
| Mingxuan Han (PhD). | (expected 2022) |
| Tao Yang (PhD). co-advised by Qingyao Ai | (expected 2023) |
| Benwei Shi (PhD). | (entered 2018) |
| Hasan Pourmahmood (PhD). | (entered 2019) |
| Prince Osei Aboagye (PhD). | (joined group 2020) |
| Peter Jacobs (PhD) | (joined group Fall 2020) |
| Meysam Alishahi (PhD). co-advised by Bei Wang Phillips | (joined group Spring 2022) |
| Jack Shunn (BS). | (joined group Spring 2022) |
| Anuja Garg (MS). | (joined group Spring 2022) |
| Anna Bell (BS). | (joined group Summer 2022) |

Graduated Students

| | |
|---|--------|
| Yanqing Peng (PhD co-advised by Feifei Li) – first job: Facebook “ <i>Toward Designing Efficient and Secure Systems for Big Data</i> ” | (2022) |
| Yuwei Wang (MS co-advised by Feifei Li) – first job: Facebook | (2021) |
| Zhuoyue Zhao (PhD co-advised by Feifei Li) – first job: Assistant Prof at U Buffalo “ <i>Approximate Query Processing Via Random Sampling</i> ” | (2021) |
| Zhao Chang (PhD co-advised by Feifei Li) – first job: faculty at Xidan University “ <i>Scalable and Secure Data Analysis in Cloud</i> ” | (2021) |
| Sunipa Dev (PhD). – first job: NSF CI Fellow at UCLA “ <i>The Geometry of Distributed Representations for Better Alignment, Attenuated Bias, and Improved Interpretability</i> ” | (2020) |
| Austin Watkins (BS Thesis) – first job: PhD student at Johns Hopkins CS. “ <i>Using Existential Theory of the Reals to Bound VC-Dimension</i> ” | (2020) |
| WaiMing Tai (PhD). – first job: postdoc at U Chicago “ <i>Geometry of Kernel Density Estimate</i> ” | (2020) |
| Jiahui (Karen) Chen (BS Thesis). – first job: Facebook “ <i>Practical and Configurable Network Traffic Classification Using Probabilistic Machine Learning</i> ” | (2020) |
| Michael Matheny (PhD). – first job: Amazon “ <i>Approximate Statistical Discrepancy</i> ” | (2019) |
| Pingfan Tang (PhD). – first job: Google. “ <i>Robust Estimation and Sketching of Points, Lines, Trajectories and other Shapes</i> ” | (2019) |
| Jian Ying (MS Thesis). “ <i>Corrected Moran’s I Statistic</i> ” | (2019) |

- Benwei Shi (BS Thesis) - first job: PhD student UofU. (2018)
"Analyzing Simultaneous Iterations"
- Giorgi Kvernadze (BS Thesis) - first job: PhD student UofU. (2018)
"Data-Driven Secret Santa"
- Zahra Fahimfar (MS Thesis). (2018)
"Detecting Potential Lensed Galaxies Behind Foreground Galaxy Targets Using Machine Learning Techniques"
- Roy Wong (MS Thesis). (2017)
"Moran's I Spatial Auto-Correlation and Anomaly Detection Utilizing PCA and High Dimensional Feature Vectors"
- Saffia Hassan (BS Thesis). - first job: MS student UofU (2017)
"Evaluating Relationships Between Vector Spaces and Word Embeddings"
- Yan Zheng (PhD) - first job : Visa Research. (2017)
"Algorithms and Coresets for Large Scale Kernel Smoothing"
- Mina Ghashami (PhD) - first job : postdoc @ Rutgers. (2017)
"On FrequentDirections : A Streaming Matrix Sketching Algorithm"
- Drew McClelland (BS Thesis) - first job : Qualtrics. (2017)
"Analysis of Mapping Techniques on Spatial Scan Statistic"
- Yi Oi (MS Project) - first job : Expedia. (2017)
- Sierra Allred (Bachelors in Undergraduate Studies in *Data Science*) - first job : data scientist @ VidAngle. (2016)
"College Cost Analyzer: Pick What You Can Pay For"
- Kaiqiang Wang (MS Project). - first job : Google. (2016)
- Liang Zhang (MS Project). - first job : Microsoft. (2015)
- Raghvendra Singh (MS Thesis). - first job : InsideSales. (2015)
"Scalable Spatial Scan Statistics"
- Jamie Iong (BS Thesis). - first job : EMC. (2015)
"Solving K-depth Coverage problem using Sweep Line Algorithm and Red Black Tree"
- Tami Y. Porter-Jones (BS Thesis). - first job : Myriad Genetics. (2015)
"Detecting Large DNA Rearrangements Using NGS Data"
- Amey Desai (MS Thesis) - first job: UrbanEngines (Bay Area startup). (2014)
"Streaming Algorithms for Matrix Approximation"
- Sitaram Gautum (Bachelors in Undergraduate Studies in *Data Science*) (2014)
- Shashank Krishnaswamy (MS Project) - first job: Amazon. (2013)
"Quality Control in Weather Data with Quantiles"
- Alex Clemmer (BS Thesis) - first job: Microsoft. (2013)
"Streaming LDA"
- Yuan Fang (MS) - first job: Zillow. (2013)
- Supraja Jayakumar (MS) - first job: Cerner Systems. (2012)
- Lingbing Jiang (MS) - first job: Microsoft. (2012)

Funding

Total: **\$3,573,135** (roughly)

Algorithmic Problems in Geometric Statistical Problems on Spatial Datasets. NSF Computing Innovations Post-doctoral Fellow (Sep. 2009-Aug 2011) **\$246,250**.

- NSF 0937060 to CRA, subaward CIF-32 to the University of Utah (Sep 2009 - Aug 2010) **\$140,000**
- NSF 1019343 to CRA, subaward CIF-A-32 to the University of Utah (Sep 2010 - Aug 2011) **\$106,250**

Synopsis Data Structures for Data Analysis in Shape Space.

NSF-CCF 1115677 (Sep. 2011-Aug 2014) **\$127,682** (out of **\$347,716**).

Principal Investigator: Suresh Venkatasubramanian, Role: Senior Personnel

Building a Mergeable and Interactive Distributed Data Layer for Big Data Summarization Systems.

NSF-BIGDATA 1251019 (Sep 2013 - Aug 2016). *about* **\$308,421** (out of **\$685,380**).

Principal Investigator: Feifei Li, Role: co-PI

CAREER: Foundations for Geometric Analysis of Noisy Data.

NSF-CCF CAREER 1350888: (May 2014 - May 2019). **\$522,135**

Role: PI

STORM: Spatio-Temporal Online Reasoning and Management of Large Data.

NSF-ACI 1443046 (CIF21 DIBBs) (Sep 2014 - Aug 2018). *about* **\$289,493** (out of **\$1,157,975**).

Principal Investigator: Feifei Li, Role: co-PI

Seal: Secure Engine for AnaLytics - From Secure Similarity Search to Secure Data Analytics.

NSF-TWC 1514520 (Medium: Collaborative Research): (July 2015 - June 2019). *about* **\$300,003** (out of **\$600,007**).

Principal Investigator: Feifei Li, Role: co-PI

NeTS: Detecting and Localizing Spectrum Offenders Using Crowdsourcing.

NSF-CNS 1564287 (Medium: Collaborative Research): (August 2016 - July 2019). **\$283,991** (out of **\$951,975**).

Principal Investigator: Sneha Kasera, Role: co-PI

Persistent Data Summaries: Temporal Analytics on Big Data Histories.

NSF-III 1816149 (Small): (November 2018 - November 2021; NCE to Nov 2023). **\$246,666** (out of **\$499,934**).

Principal Investigator: Feifei Li, Role: co-PI

CDS&E: Extracting Models from Data - A Novel Data-Driven Simulation Strategy for Reacting Flows.

NSF-CFS 1953350 : (November 2020 - September 2023). *about* **\$226,610** (out of **\$453,220**).

Principal Investigator: James Sutherland, Role: co-PI

Disentangling Merchant Embedding Spaces.

Visa Research : (January 2020 - December 2020). **\$47,500**

Visa Research : (January 2021 - August 2021). **\$25,000**

Visa Research : (August 2021 - August 2022). **\$75,000**

Visa Research : (August 2022 - August 2023). **\$75,000**

Role: PI

III: Small: Towards a Database Engine for Interactive and Online Sampling and Analytics.

NSF-III 1619287 : (September 2020 - August 2021). *about* **\$300,000** (out of **\$500,000**)

Role: PI (transferred support), original PI: Feifei Li (September 2016 - September 2020)

AF: Small: The Geometry of Learning on Structured Data Objects.

NSF-CCF 2115677 : (November 2021 - September 2024). **\$499,384**

Role: PI

TEACHING

Teaching

(instructor average / SoC instructor average || course average / SoC course average)

Ethics in Data Science (cs3390/ds3390)

[adapted to large undergrad setting]

Spring 2021: 69 undergrad students. (instructor: 100/0/0 || course: 93/7/0) [*as % Yes/Neutral/No*]

Data Mining (cs5140/6140)

[self-developed]

Spring 2020: 125 | 43 undergrad + 82 graduate students. (instructor: 5.79/5.26 || course: 5.71/5.17)

Spring 2019: 117 | 39 undergrad + 78 graduate students. (instructor: 5.49/5.32 || course: 5.39/5.20)

Spring 2018: 103 | 30 undergrad + 73 graduate students. (instructor: 5.53/5.26 || course: 5.4/5.17)

Spring 2017: 129 | 40 undergrad + 89 graduate students. (instructor: 5.48/5.16 || course: 5.4/5.15)

Spring 2016: 70 | 23 undergrad + 47 graduate students. (instructor: 5.32/5.24 || course: 5.23/5.13)

Spring 2015: 98 | 29 undergrad + 69 graduate students. (instructor: 5.39/5.25 || course: 5.30/5.17)

Spring 2014: 69 | 14 undergrad + 55 graduate students. (instructor: 5.27/5.10 || course: 5.14/5.04)

Spring 2013: 40 | 10 undergrad + 30 graduate students. (instructor: 5.51/5.30 || course: 5.51/5.18)

Spring 2012: 35 | 8 undergrad + 27 graduate students. (instructor: 5.76/5.23 || course: 5.52/5.12)

Foundations of Data Analysis (Math for Data) (cs3190/ds3190/comp5960)

[self-developed]

Fall 2022: 183 | 166 undergrad students + 17 graduate students. (in progress)

Fall 2021: 128 | 123 undergrad students + 5 graduate students. (instructor: 5.59/– || course: 5.65/–)

Fall 2020: 87 undergrad students. (instructor: 5.44/5.26 || course: 5.37/5.17)

Fall 2019: 57 undergrad students. (instructor: 5.26/5.26 || course: 4.71/5.17)

Fall 2017: 46 | 42 undergrad + 4 graduate students. (instructor: 5.2/5.26 || course: 5.16/5.15)

Fall 2016: 35 undergrad students. (instructor: 5.32/5.32 || course: 5.29/5.14)

Probability and Statistics for Engineers (cs3130/ece3530)

Fall 2014: 93 undergrad students. (instructor: 5.23/5.36 || course: 5.12/5.21)

Models of Computation for Massive Data (cs7960)

[self-developed]

Fall 2013: 27 students. (instructor: 5.78/5.31 || course: 5.59/5.18)

Note: *officially taught as cs7931, a seminar version, while on parental leave*

Fall 2011: 29 students. (instructor: 5.64/5.38 || course: 5.24/5.12)

Data Mining Seminar (cs7931)

Spring 2015: *Matrix Sketching*. 15 students (instructor: 5.38/5.25 || course: 5.31/5.17)

[self-developed]

Note: *cross-listed as cs6961, co-taught with Mina Ghashami*

Fall 2012: *Sampling*. 16 students (instructor: 5.95/5.27 || course 5.86/5.15)

[self-developed]

Fall 2010: *Modeling Data with Uncertainty*. (taught while a postdoc at the U)

[self-developed]

Data Science Seminar / Data Reading Group (cs7941)

Fall 2022: *Data Science Seminar*. 11 students. (in progress)

Spring 2022: *Data Science Seminar*. 10 students. (instructor: 6.0/– || course: 6.0/–)

Fall 2021: *Data Science Seminar*. 13 students. (instructor: 5.83/– || course: 5.67/–)

Fall 2020: *Data Science Seminar*. 11 students. (instructor: 6.00/5.26 || course: 6.00/5.17)

Spring 2020: *Data Science Seminar*. 10 students. (instructor: 5.75/5.26 || course: 5.75/5.17)

Spring 2019: *Data Science Seminar*. 8 students. (instructor: 5.29/5.26 || course: 5.07/5.17)
 Spring 2019: *Data Science Seminar*. 8 students. (instructor: 5.29/5.26 || course: 5.07/5.17)
 Spring 2018: *Data Group Meeting*. 5 students. (instructor: 5.74/5.16 || course: 5.66/5.15)
 Spring 2017: *Data Group Meeting*. 5 students. (instructor: 5.5/5.16 || course: 5.5/5.15)
 Spring 2016: *Data Group Meeting*. 6 students. (instructor: 5.67/5.24 || course: 5.60/5.13)
 Fall 2014: *Data Group Meeting*. 6 students. (instructor: 5.75/5.36 || course: 5.67/5.21)
 Spring 2014: *Data Group Meeting*. 5 students. (instructor: 6/5.10 || course 5.96/5.04)
 Spring 2012: *Data Group Meeting*. 11 students. (instructor: 5.89/5.38 || course 5.89/5.12)

SERVICE

Internal Service

Utah Center for Data Science (Director) (2019-2022)

Data Science Program (Director) (2019-2022)

Oversees all data science educational programs in School of Computing

Manages BS in Data Science, and Undergraduate Certificates

Renamed/Updated Graduate Certificate in Data Science (2020)

Created Undergraduate Certificate in Data Science (2020)

Created Undergraduate Certificate in Data Fluency (2020)

Curriculum Committee

Data Science Representative (2019-2021)

Faculty Hiring Committees

sub-chair, Data Management (2022)

Lecturing Faculty (2021)

Data / Theory (2019)

Machine Learning (2018)

Statistics Cluster (with Math + BioEng) (2014-17, 19-20)

Foundations of Data Science (2015)

Big Data Analytics (2014)

ML+NLP (2013)

Info Visualization (2013-15)

College of Engineering Graduate Curriculum Committee (2021-)

Bachelors of Science in Data Science

Led effort to formulate and create this new major (2017-2019)

Big Data Graduate Program (Director) (2013-18)

Created new *Graduate Certificate in Big Data* (2014)

Outreach: advertise certificate to companies and around campus, develop advertising materials (2014-19)

Webpage: Designed and maintain content for big data webpage, including admissions form (2014-18)

<http://www.cs.utah.edu/bigdata>

Video: Arranged for live-streaming to students at Adobe, Ancestry.com, Overstock.com, Microsoft (2014-20)

<https://www.youtube.com/channel/UCDUS80bdunpmvWVPyFRPqFQ>

Admissions: Created, maintained, and ran admissions process (2014-18)

Academic Advising: meet with most students before they apply, guide them through classes (2014-20)

Data Track

Track Director (2014-18)

Created "Data Science Option," coordinating with Math MStat (2016)

Redesigned Thesis Proposal Guidelines (2011, 2019)

Utah Data Science Day

Co-organizing campus-wide events on Data Science: 161, 250+, and 150+ registered attendees. (2016-19)
 Webpage: Organizing this event and advertising all Utah data science programs (2016-19)
<http://datascience.utah.edu>

Utah Data Science Club

Faculty Mentor. (<http://datascience.utah.edu/club.html>) (2017-19, 21-)
 Arranged Data Scientist Panel (about 100 attending) (2017)

Poster Chair, Graduate Recruitment

Recruited posters, arranged for prizes, advertised event, coordinated judging (2012,2013)
 Brought in "poster expert" for review session (2013)

Admissions Committee (Data, Algorithms) (2012-15)
 SoC Scholarship Committee (2015-17)

External Service

Editorial Boards

Editorial Board for Computing in Geometry and Topology. (2021-)
Associate Editor for IEEE Transactions on Knowledge and Data Engineering (TKDE). (2016-2020)
SIAM Journal of Scientific Computing, Special Section for CSE15 on CSE Software and Big Data in CSE. (2014-16)

Program Chair

International Symposium on Computational Geometry (co-chair). (2024)
 International Conference on Data Engineering, Demo Track (co-chair). (2022)

(Senior) Program Committees (manages 20+ papers)

International Symposium on Computational Geometry (SOCG). (2016,2021)
European Symposium on Algorithms (ESA). (2013,2017,2021)
ACM-SIAM Symposium on Discrete Algorithms (SODA). (2015)
ACM Symposium on Principles of Database Systems (PODS). (2015,2017)
International Conference on Database Theory (ICDT). (2017)
Fall Workshop in Computational Geometry (FWCG). (2012,2014,2021)
ACM International Conference on Information and Knowledge Management (CIKM). (2013)

Program Committees / Listed Reviewer (manages 5-10 papers)

International Conferences on Artificial Intelligence and Statistics (AISTats). (2018–2022)
Conference on Neural Information Processing Systems (NeurIPS). (2018–2021)
ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL). (2017–2020)
International Conference on Machine Learning (ICML). (2019)
IEEE International Conference on Data Engineering (ICDE). (2014,2016)
Workshop on Massive Data Algorithmics (MASSIVE). (2014,2015,2016)
International Conference on Database Systems for Advanced Applications. (2014)
ACM IKDD Conference on Data Science. (2014)
International Workshop on Big Dynamic Distributed Data, a VLDB Workshop. (2013)
ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). (2012)
Robotics: Science and Systems (RSS). (2006)

Journal Reviewing (for many several times)

Computational Geometry: Theory and Applications (CGTA).
Computational Statistics and Data Analysis (CSDA).
Computational Statistics (COST).

Discrete and Computational Geometry (DCG).
Distributed and Parallel Databases (DaPD).
International Journal of Computational Geometry (IJCGA).
Journal of Computational Geometry (JoCG).
Journal of Discrete Algorithms (JDA).
ACM Journal of Experimental Algorithms (JEA).
Journal of Machine Learning Research (JMLR).
The London Mathematical Society (LMS).
Wiley Journal on Statistical Analysis and Data Mining (SADM).
SIAM Journal of Computing (SICOMP).
SIAM Journal of Discrete Mathematics (SIDMA).
SIAM Journal of Scientific Computing (SISC).
ACM Transactions on Algorithms (TALG).
ACM Transactions on Database Systems (TODS).
IEEE Transactions on Knowledge Discovery and Data Engineering (TKDE).
Theory of Computing (ToC).
Theory of Computing Systems (TOCS).
The Visual Computer (VisComp).
IEEE Transactions on Multimedia.
IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI).
IEEE Signal Processing Letters.
Numerical Algorithms.
BMC Systems Biology.

Conference Sub-Reviewing (for most several times)

ACM Symposium on Computational Geometry (SoCG).
ACM-SIAM Symposium on Discrete Algorithms (SoDA).
ACM Symposium on Theory of Computing (SToC).
IEEE Symposium on Foundations of Computer Science (FoCS).
European Symposium on Algorithms (ESA).
International Colloquium on Automata, Languages, and Programming (ICALP).
International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX).
ACM Symposium on Principles of Database Systems (PODS).
Algorithms and Data Structures Symposium (WADS).
Scandinavian Symposium and Workshops on Algorithm Theory (SWAT).
Latin American Symposium on Theoretical Informatics (LATIN).
ACM International Conference on Knowledge Discovery and Data Mining (KDD).
Annual Conference on Neural Information Processing Systems (NIPS).
ACM International Conference on Advanced in Geographic Information Systems (ACM SIGSPATIAL).
Conference on Information and Knowledge Mining (CIKM).
ACM International Conference for High Performance Computing, Networking, Storage, and Analysis (SC).

Proposal Panelist

NSF (10 times since 2011).
International external reviewer (6 times).
Israel (4 times)
Netherlands (1 time)
European Research Foundation (1 time)

Workshop/Conference Organization

6th Workshop on *Geometry and Machine Learning* at SOCG, co-organizer.
 Dagstuhl Workshop on *Computational Geometry* (accepted), co-organizer.

(2022)

(May 2023)

| | |
|---|-------------|
| 5th Workshop on <i>Geometry and Machine Learning</i> at SOCG, co-organizer. | (2021) |
| Dagstuhl Workshop on <i>Computational Geometry</i> 21181, co-organizer. | (May 2021) |
| CGWeek Workshop Committee, chair. | (2020) |
| 4th Workshop on <i>Geometry and Machine Learning</i> at SOCG, co-organizer. | (2019) |
| CGWeek Young Researcher Forum, program committee. | (2016,2019) |
| IEEE ICDE 2019 PhD Symposium, program committee. | (2019) |
| 3rd Workshop on <i>Geometry and Machine Learning</i> at SOCG, co-organizer. | (2018) |
| 2nd Workshop on <i>Geometry and Machine Learning</i> at SOCG, co-organizer. | (2017) |
| Workshop on <i>Geometry and Machine Learning</i> at jointly <i>ACM STOC</i> and SOCG, organizer. | (2016) |
| <i>SIAM CSE</i> co-organized and served on panel on "Data Science: What is It and How to Teach It." | (2015) |
| <i>ACM SIGMOD</i> demo and workshop chair / <i>ACM PODS</i> local arrangements chair. | (2014) |
| <i>ACM Symposium on Computational Geometry</i> , Workshop on Computational Geometry, program committee. | (2014) |
| <i>Workshop on Computational Geometry in the Field (8F-CG)</i> at SoCG 2012, organizer. | (2012) |
| <i>Symposium on Computational Geometry (SoCG)</i> , local arrangements. | (2010) |
| Duke CS Graduate Student Retreat, lobbied for, developed, and chaired first ever (now biennial event). | (2008) |

Community Engagement

| | |
|---|------------------|
| Created booth for "BS in Data Science" for Utah Discover Engineering traveling exhibit. | (2019-2021) |
| Speaker on "What is Data Science?" at University of Utah Engineering Day (for highschoolers). | (2019,2020,2021) |
| Speaker at MIT Alumni (and friends) Club. | (2017) |
| Speaker at Goldman Sachs Tech Expo. | (2017) |
| Speaker at Utah Biomedical Data Science Boot Camp. | (2016) |
| Speaker on "Big Data and Data Science" at University of Utah Engineering Day (for highschoolers). | (2015) |
| Participant in NSF <i>Data Science Education Workshop</i> . | (2015) |
| Invited Speaker at Lambda Lounge Utah "Sketch Data Structures." | (2015) |
| Invited presenter at VOLCOM IT Exchange "Big Data at U of U." | (2015) |
| SIAM Connect video on "What is Data Science?" https://www.youtube.com/watch?v=KUByrBWS2HU | (2015) |
| Interview guest on KPCW <i>Cool Science Radio</i> http://kpcw.org/post/cool-science-radio-february-26-2015 | (2015) |
| Invited presenter in Big Data Session at Chinese Association for Science and Technology, Utah Chapter | (2015) |
| Panelist at Utah Data & Analytics Summit | (2015) |
| Mentor for Daniel Liu, first place in Mathematics at SLVSEF | (2013) |
| Judge at Salt Lake Valley Science & Engineering Fair (SLVSEF) | (2012) |
| Guest lecturer at Duke Talent Identification Program (for middle schoolers in CS) | (2008,2009) |
| Advised creation of CSters (Rice Women in Computer Science club) | (2003) |

Last updated: January 29, 2023