

Daniel B. Neill

1432 South Negley Avenue
Pittsburgh, PA 15217, U.S.A.
Phone: (412) 268-3885

E-mail: neill@cs.cmu.edu
<http://www.cs.cmu.edu/~neill>
U.S. Citizen

Employment

2006-present: Assistant Professor of Information Systems, H. John Heinz III College, School of Public Policy and Management, School of Information Systems and Management, Carnegie Mellon University.

2006-present: Courtesy Assistant Professor, Machine Learning Department, School of Computer Science, Carnegie Mellon University.

2006-present: Courtesy Assistant Professor, Robotics Institute, School of Computer Science, Carnegie Mellon University.

2007-present: Adjunct Assistant Professor, Department of Biomedical Informatics, School of Medicine, University of Pittsburgh.

Education

Ph.D. (Computer Science), **Carnegie Mellon University**, 2006.

Thesis: "Detection of Spatial and Spatio-Temporal Clusters"

Advisor: Andrew W. Moore

M.S. (Computer Science), **Carnegie Mellon University**, 2004.

Advisor: Andrew W. Moore

M.Phil. (Computer Speech), **Cambridge University**, 2002.

Thesis: "Fully Automatic Word Sense Induction by Semantic Clustering"

Advisor: Karen Sparck-Jones

B.S.E. (Electrical Engineering/Computer Science), **Duke University**, 2001.

Graduated summa cum laude, 1st in class, with highest distinction.

Thesis: "Optimality Under Noise"

Advisor: David Kraines

Publications

Book Chapters

1. D. B. Neill, G. F. Cooper, K. Das, X. Jiang, and J. Schneider. Bayesian network scan statistics for multivariate pattern detection. In J. Glaz, V. Pozdnyakov, and S. Wallenstein, eds., *Scan Statistics: Methods and Applications*, 221-250, 2009.
2. D. B. Neill and A. W. Moore. Methods for detecting spatial and spatio-temporal clusters. In M. Wagner, A. Moore, and R. Aryel, eds., *Handbook of Biosurveillance*, 243-254, 2006.

3. D. B. Neill and A. W. Moore. Efficient scan statistic computations. In A. Lawson and K. Kleinman, eds., *Spatial and Syndromic Surveillance for Public Health*. Chichester, UK: Wiley, 189-202, 2005.

Refereed Journal Articles

4. D. B. Neill. Fast subset scan for spatial pattern detection. *Journal of the Royal Statistical Society (Series B: Statistical Methodology)*, accepted for publication, 2011.
5. S. Hasan, G. T. Duncan, D. B. Neill, and R. Padman. Automatic detection of omissions in medication lists. *Journal of the American Medical Informatics Association*, 18(4): 449-458, 2011.
6. D. B. Neill. Fast Bayesian scan statistics for multivariate event detection and visualization. *Statistics in Medicine*, 30: 455-469, 2011.
7. R. J. Leeman-Neill, R. R. Seethala, S. V. Singh, M. L. Freilino, J. S. Bednash, S. M. Thomas, M. C. Panahandeh, W. E. Gooding, S. C. Joyce, M. W. Lingen, D. B. Neill, and J. R. Grandis. Inhibition of EGFR-STAT3 signaling with erlotinib prevents carcinogenesis in a chemically-induced mouse model of oral squamous cell carcinoma. *Cancer Prevention Research*, 4(2): 230-237, 2011.
8. D. Oliveira, D. B. Neill, J. H. Garrett Jr., and L. Soibelman. Detection of patterns in water distribution pipe breakage using spatial scan statistics for point events in a physical network. *Journal of Computing in Civil Engineering*, 25(1): 21-30, 2011.
9. D. B. Neill and G. F. Cooper. A multivariate Bayesian scan statistic for early event detection and characterization. *Machine Learning* 79: 261-282, 2010.
10. X. Jiang, D. B. Neill, and G. F. Cooper. A Bayesian network model for spatial event surveillance. *International Journal of Approximate Reasoning* 51: 224-239, 2010.
11. R. J. Leeman-Neill, Q. Cai, S. C. Joyce, S. M. Thomas, N. E. Bhola, D. B. Neill, J. L. Arbiser, and J. R. Grandis. Honokiol inhibits epidermal growth factor receptor signaling and enhances the antitumor effects of epidermal growth factor receptor inhibitors. *Clinical Cancer Research* 16(9): 2571-2579, 2010.
12. R. J. Leeman-Neill, S. E. Wheeler, S. V. Singh, S. M. Thomas, R. R. Seethala, D. B. Neill, M. C. Panahandeh, E.-R. Hahm, S. C. Joyce, M. Sen, Q. Cai, M. L. Freilino, C. Li, D. E. Johnson, and J. R. Grandis. Guggulsterone enhances head and neck cancer therapies via inhibition of signal transducer and activator of transcription-3. *Carcinogenesis* 30(11): 1848-1856, 2009.
13. D. B. Neill. Expectation-based scan statistics for monitoring spatial time series data. *International Journal of Forecasting* 25: 498-517, 2009.
14. D. B. Neill. An empirical comparison of spatial scan statistics for outbreak detection. *International Journal of Health Geographics* 8: 20, 2009.
15. D. B. Neill. Cascade effects in heterogeneous populations. *Rationality and Society* 17(2): 191-241, 2005.

16. D. B. Neill. Evolutionary stability for large populations. *Journal of Theoretical Biology* 227(3): 397-401, 2004.
17. M. M. Wagner, F.-C. Tsui, J. Espino, W. Hogan, J. Hutman, J. Hersh, D. Neill, A. Moore, G. Parks, C. Lewis, and R. Aller. A national retail data monitor for public health surveillance. *Morbidity and Mortality Weekly Report, Supplement on Syndromic Surveillance* 53: 40-42, 2004.
18. D. B. Neill. Optimality under noise: higher memory strategies for the Alternating Prisoner's Dilemma. *Journal of Theoretical Biology* 211(2): 159-180, 2001.

Refereed Conference Proceedings

19. X. Jiang, D. B. Neill, and G. F. Cooper. Generalized AMOC curves for evaluation and improvement of event surveillance. *Proceedings of the American Medical Informatics Association Annual Symposium*, 281-285, 2009.
20. S. Hasan, G. T. Duncan, D. B. Neill, and R. Padman. Towards a collaborative filtering approach to medication reconciliation. *Proceedings of the American Medical Informatics Association Annual Symposium*, 288-292, 2008.
21. K. Das, J. Schneider, and D. B. Neill. Anomaly pattern detection in categorical datasets. *Proceedings of the 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 169-176, 2008.
22. C. A. Harle, D. B. Neill, and R. Padman. An information visualization approach to classification and assessment of diabetes risk in primary care. *Proceedings of the 3rd INFORMS Workshop on Data Mining and Health Informatics*, 2008.
23. M. Makatchev and D. B. Neill. Learning outbreak regions in Bayesian spatial scan statistics. *Proceedings of the ICML/UAI/COLT 2008 Workshop on Machine Learning for Health Care Applications*, 2008.
24. D. B. Neill, A. W. Moore, and G. F. Cooper. A Bayesian spatial scan statistic. In Y. Weiss, *et al.*, eds. *Advances in Neural Information Processing Systems 18*, 1003-1010, 2006.
25. M. R. Sabhnani, D. B. Neill, A. W. Moore, A. Dubrawski, and W.-K. Wong. Efficient analytics for effective monitoring of biomedical security. *Proceedings of the International Conference on Information and Automation*, 2005.
26. D. B. Neill, A. W. Moore, M. R. Sabhnani, and K. Daniel. Detection of emerging space-time clusters. *Proceedings of the 11th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 218-227, 2005.
27. D. B. Neill and A. W. Moore. Anomalous spatial cluster detection. *Proceedings of the KDD 2005 Workshop on Data Mining Methods for Anomaly Detection*, 2005.
28. M. R. Sabhnani, D. B. Neill, A. W. Moore, F.-C. Tsui, M. M. Wagner, and J. U. Espino. Detecting anomalous patterns in pharmacy retail data. *Proceedings of the KDD 2005 Workshop on Data Mining Methods for Anomaly Detection*, 2005.

29. P. Hsiung, A. Moore, D. Neill, and J. Schneider. Alias detection in link data sets. *Proceedings of the First International Conference on Intelligence Analysis*, 2005.
30. D. B. Neill, A. W. Moore, F. Pereira, and T. Mitchell. Detecting significant multidimensional spatial clusters. In L. K. Saul, *et al.*, eds., *Advances in Neural Information Processing Systems 17*, 969-976, 2005.
31. D. B. Neill and A. W. Moore. Rapid detection of significant spatial clusters. *Proceedings of the 10th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 256-265, 2004.
32. D. B. Neill and A. W. Moore. A fast multi-resolution method for detection of significant spatial disease clusters. In S. Thrun, *et al.*, eds., *Advances in Neural Information Processing Systems 16*, 651-658, 2004.
33. D. B. Neill. Cooperation and coordination in the Turn-Taking Dilemma. *Proceedings of the Ninth Conference on Theoretical Aspects of Rationality and Knowledge*, 231-244, 2003.

Refereed Conference and Journal Abstracts

34. D. B. Neill, C. Harle, R. Padman, F. Solano, and J. Zgibor. Clinician information needs for data visualization based diabetes risk assessment and guideline compliance. *Proceedings of the American Medical Informatics Association Annual Symposium*, 2011.
35. D. B. Neill, E. McFowland III, and H. Zheng. Fast subset scan for multivariate spatial biosurveillance. *Emerging Health Threats Journal*, 4:s42, 2011.
36. D. B. Neill and Y. Liu. Generalized fast subset sums for Bayesian detection and visualization. *Emerging Health Threats Journal*, 4:s43, 2011.
37. H. Zheng, R. Padman, D. B. Neill, and S. Hasan. A comparison of collaborative filtering methods for medication reconciliation. *Proceedings of the 13th International Congress on Medical Informatics*, 2010.
38. D. Gartner, R. Kolisch, R. Padman, and D. B. Neill. Early DRG classification of inpatients in hospitals. *Proceedings of the 35th Conference on Operational Research Applied to Health Services*, 2009.
39. D. B. Neill. Fast subset sums for multivariate Bayesian scan statistics. *Proceedings of the International Society for Disease Surveillance Annual Conference*, available online at www.syndromic.org, 2009.
40. S. Speakman and D. B. Neill. Fast graph scan for scalable detection of arbitrary connected clusters. *Proceedings of the International Society for Disease Surveillance Annual Conference*, available online at www.syndromic.org, 2009.
41. K. Das, J. Schneider, and D. B. Neill. Anomaly pattern detection for biosurveillance. *Advances in Disease Surveillance* 5: 19, 2008.
42. M. Makatchev and D. B. Neill. Learning outbreak regions for Bayesian spatial biosurveillance. *Advances in Disease Surveillance* 5: 45, 2008.

43. D. B. Neill. Fast and flexible outbreak detection by linear-time subset scanning. *Advances in Disease Surveillance* 5: 48, 2008.
44. D. B. Neill and W. L. Gorr. Detecting and preventing emerging epidemics of crime. *Advances in Disease Surveillance* 4: 13, 2007.
45. D. B. Neill and J. Lingwall. A nonparametric scan statistic for multivariate disease surveillance. *Advances in Disease Surveillance* 4: 106, 2007.
46. D. B. Neill. Incorporating learning into disease surveillance systems. *Advances in Disease Surveillance* 4: 107, 2007.
47. D. B. Neill. An empirical comparison of spatial scan statistics for outbreak detection. *Advances in Disease Surveillance* 4: 259, 2007.
48. D. B. Neill, A. W. Moore, and G. F. Cooper. A multivariate Bayesian scan statistic. *Advances in Disease Surveillance* 2: 60, 2007.
49. D. B. Neill and M. R. Sabhnani. A robust expectation-based spatial scan statistic. *Advances in Disease Surveillance* 2: 61, 2007.
50. D. B. Neill, A. W. Moore, and G. F. Cooper. A Bayesian scan statistic for spatial cluster detection. *Advances in Disease Surveillance* 1: 55, 2006. **Received “Best Research Presentation” award, National Syndromic Surveillance Conference, 2005.**
51. D. B. Neill, A. W. Moore, M. R. Sabhnani, and K. Daniel. An expectation-based scan statistic for detection of space-time clusters. *Advances in Disease Surveillance* 1: 56, 2006.
52. M. R. Sabhnani, D. B. Neill, A. W. Moore, F.-C. Tsui, M. M. Wagner, and J. U. Espino. Monitoring pharmacy retail data for anomalous space-time clusters. *Advances in Disease Surveillance* 1: 62, 2006.
53. D. B. Neill, A. W. Moore, and M. R. Sabhnani. Detecting elongated disease clusters. *Morbidity and Mortality Weekly Report, Supplement on Syndromic Surveillance* 54: 197, 2005.
54. D. B. Neill and A. W. Moore. A fast grid-based scan statistic for detection of significant spatial disease clusters. *Morbidity and Mortality Weekly Report, Supplement on Syndromic Surveillance* 53: 255, 2004.

Other Journal Articles

55. S. W. Malone, C. A. Miller, and D. B. Neill. Traffic flow models and the evacuation problem. *Undergraduate Journal of Mathematics and its Applications* 22(3): 273-292, 2001. **Winner of the 2001 Mathematical Contest in Modeling.**
56. S. W. Malone, J. A. Mermin, and D. B. Neill. Air traffic control. *Undergraduate Journal of Mathematics and its Applications* 21(3): 227-241, 2000. **Winner of the 2000 Mathematical Contest in Modeling.**

Technical reports and theses

57. D. B. Neill. Detection of spatial and spatio-temporal clusters. Ph.D. thesis, Carnegie Mellon University, Department of Computer Science, Technical Report CMU-CS-06-142, 2006.
58. D. B. Neill and A. W. Moore. Detecting space-time clusters: prior work and new directions. Carnegie Mellon University, Department of Computer Science, Technical Report CMU-CS-05-115, 2005.
59. D. B. Neill. Evolutionary dynamics with large aggregate shocks. Carnegie Mellon University, Department of Computer Science, Technical Report CMU-CS-03-197, 2003.
60. D. B. Neill. An evolutionary resolution to the Finitely Repeated Prisoner's Dilemma paradox. Carnegie Mellon University, Department of Computer Science, Technical Report CMU-CS-03-155, 2003.
61. D. B. Neill. Fully automatic word sense induction by semantic clustering. Cambridge University, M.Phil. thesis, 2002.
62. D. B. Neill. Optimality under noise: higher memory strategies for the Alternating Prisoner's Dilemma. Duke University, undergraduate honors thesis, 2000.

Submitted for publication

- D. B. Neill, E. McFowland III, and H. Zheng. "Fast subset scan for multivariate event detection," journal paper under review.
- S. Somanchi and D. B. Neill. "Fast graph structure learning from unlabeled data for event detection," conference paper under review.
- K. Shao, Y. Liu, and D. B. Neill. "A generalized fast subset sums framework for Bayesian event detection," conference paper under review.
- C. Harle, D. B. Neill, and R. Padman. "An information visualization approach to type 2 diabetes risk assessment in primary care," journal paper under revision.
- D. Gartner, R. Kolisch, D. B. Neill, and R. Padman. "Machine learning approaches for early DRG classification of inpatient data in hospitals," journal paper under revision.
- R. F. Davies, D. B. Neill, *et al.* "Detection of the Walkerton gastroenteritis outbreak by text mining of emergency room health records," journal paper under revision.
- K. Das, J. Schneider, and D. B. Neill. "Detecting anomalous groups in categorical datasets," conference paper under revision.

Invited Talks and Tutorials

- "Fast multivariate subset scanning for scalable cluster detection," *Joint Statistical Meetings 2011*, Miami, FL, August 2011.

- “Machine learning for population health and disease surveillance,” *Advanced Analytics Workshop*, Washington, DC, April 2011.
- “Spatial and subset scanning for multivariate health surveillance,” *Data Fusion Research Meeting*, Ottawa, ON, March 2011.
- “Spatial scanning tips and tricks for practical outbreak detection,” invited webinar for the International Society for Disease Surveillance, January 2011.
- “Research challenges for biosurveillance: the next ten years” (invited plenary), *International Society for Disease Surveillance Annual Conference*, Park City, UT, December 2010.
- “Fast generalized subset scan for anomalous pattern detection,” *INFORMS Annual Conference*, Austin, TX, November 2010.
- “Scalable detection of anomalous patterns with connectivity constraints,” *INFORMS Annual Conference*, Austin, TX, November 2010.
- “Fast subset sums for scalable Bayesian detection and visualization,” *Fifth International Workshop on Applied Probability*, Madrid, Spain, July 2010.
- “Fast subset scanning for multivariate event detection,” *ENAR 2010 Annual Meeting*, New Orleans, LA, March 2010.
- “Application of spatial scan statistic methods to crime hot spot analysis,” *Tenth Crime Mapping Research Conference*, New Orleans, LA, August 2009.
- “Event detection,” half-day tutorial (with Weng-Keen Wong). *15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Paris, France, June 2009.
- “Multivariate Bayesian scan statistics for event detection and characterization,” *Twelfth Biennial CDC/ATSDR Symposium on Statistical Methods*, Decatur, GA, April 2009.
- “A nonparametric scan statistic for multivariate spatial biosurveillance,” *Joint Statistical Meetings 2008*, Denver, CO, August 2008.
- “Linear-time subset scanning,” *Fourth International Workshop on Applied Probability*, Compiègne, France, July 2008.
- “Multivariate event detection and characterization,” *Washington Statistical Society Seminar*, Washington, DC, May 2008.
- “Multivariate outbreak detection and characterization,” *Donald A. B. Lindberg Lecture and Symposium*, Pittsburgh, PA, May 2008.
- “A multivariate Bayesian method for spatial biosurveillance,” *Joint Statistical Meetings 2007*, Salt Lake City, UT, July 2007.
- “Monitoring multivariate spatial time series data for disease outbreak detection,” *27th Annual International Symposium on Forecasting*, New York, NY, June 2007.

“Bayesian disease surveillance by detection of anomalous clusters,” *Third ECADS Syndromic Surveillance Conference*, Ottawa, ON, October 2006.

“Bayesian disease surveillance by detection of anomalous clusters,” *Third International Workshop on Applied Probability*, Storrs, CT, May 2006.

“Scaling up geographic disease surveillance,” *Second ECADS Syndromic Surveillance Conference*, Ottawa, ON, June 2005.

Grant Funding

UPMC Healthcare Innovation Grant, Neill (PI), unrestricted gift awarded 11/8/2010, funded by University of Pittsburgh Medical Center- Technology Development Center. “Anomalous Pattern Detection from Healthcare Data Streams.” This project will apply novel pattern detection methods to detect anomalous patterns of patient care. Total award: \$121,503. Role: PI.

UPMC Healthcare Innovation Grant, Padman (PI), unrestricted gift awarded 11/8/2010, funded by University of Pittsburgh Medical Center- Technology Development Center. “Information Visualization for Cognitively Guided Decision Making for Diabetes Risk Assessment and Guideline Compliance.” This project will develop and evaluate novel information visualization tools and methods for improving diabetes care. Total award: \$110,120. Role: co-PI.

NSF IIS-0953330, Neill (PI), 7/1/2010-6/30-2015, funded by National Science Foundation. “CA-REER: Machine Learning and Event Detection for the Public Good.” This project will create and explore novel methods for detection of emerging events in massive, complex, real-world datasets. This research will be integrated with a multi-pronged educational initiative to incorporate machine learning into the public policy curriculum. Total award: \$529,962. Role: PI.

NSF IIS-0916345, Neill (PI), 8/1/2009-7/31/2012, funded by National Science Foundation. “III: Small: Fast Subset Scan for Anomalous Pattern Detection”. This project will develop new, general subset scan methods for efficient pattern detection in massive datasets. Total award: \$499,991. Role: PI.

NSF IIS-0911032, Dubrawski (PI), 9/1/2009-8/31/2013, funded by National Science Foundation. “III: Large: Discovering Complex Anomalous Patterns”. This project will develop an integrated probabilistic framework for pattern discovery, incorporating detection, characterization, explanation, and learning from user feedback. Total award: \$2,598,153. Role: Co-PI.

CDC 8-R01-HK000020, Dubrawski (PI), 9/30/2006-9/29/2008, funded by Centers for Disease Control and Prevention. “Efficient, Scalable, Multisource Surveillance Algorithms for BioSense”. This project will develop multivariate Bayesian biosurveillance methods for inclusion in the BioSense system. Total award: \$1,198,409. Role: Co-PI.

NSF IIS-0325581, Cooper (PI), 9/1/2003-8/31/2008, funded by National Science Foundation. “ITR: Bayesian Modeling for Biosurveillance”. This project will develop novel Bayesian methodologies for the detection of disease outbreaks. CMU award: \$1,246,800. Role: senior personnel.

CRTI-08-190RD, Davies (PI), 7/2009-6/2013, funded by CRTI. “Data Fusion Solutions for Monitoring CBRNE Threats”. This project focuses on general solutions for integrating multiple data sources for public health surveillance and integrates these solutions into two specific applications,

detection of severe outbreaks in hospitalized patients and surveillance of events related to illicit substance abuse. Total award: \$3,000,000. Role: Technical team, expert in statistical detection methods and data mining.

CRTI-06-0234TA, Davies (PI), 7/2007-7/2010, funded by CRTI. “Advanced Syndromic Surveillance and Emergency Triage (ASSET)”. This project will develop and deploy a system for syndromic surveillance of Emergency Department data in Ottawa, Ontario, for earlier detection of disease outbreaks and bioterrorist attacks. Total award: \$2,000,000. CMU subcontract: \$25,475. Role: Technical team, expert in statistical detection methods and data mining.

Recent Awards and Honors

Named one of the “top ten artificial intelligence researchers to watch” by *IEEE Intelligent Systems* (“AI’s 10 to Watch”, Jan/Feb 2011).

NSF CAREER Award, 2010.

“Best Research Presentation” award, *National Syndromic Surveillance Conference*, 2005.

NSF Graduate Research Fellowship, 2002-2005.

Winston Churchill Scholarship, 2001-2002.

Barry M. Goldwater Scholarship, 1999-2001.

Two time winner of Mathematical Contest in Modeling, 2000 and 2001.

University Awards and Honors

Carnegie Mellon University Doctoral Fellowship, 2002-2006.

Duke University, Walter J. Seeley Scholastic Award in Engineering, 2001.

Duke University, George Sherrerd Award in Electrical Engineering, 2001.

Oxford University, Lord Rothermere Scholarship, 2000.

Duke University, PRUV Research Fellowship, 2000.

Duke University, Angier B. Duke Scholarship, 1997-2001.

Honor Societies: Phi Beta Kappa, Tau Beta Pi, Eta Kappa Nu.

Teaching

Spring 2010-present: Instructor for 90-921/10-831, Special Topics in Machine Learning and Policy, Carnegie Mellon University.

Spring 2009-present: Instructor for 90-904/10-830, Research Seminar in Machine Learning and Policy, Carnegie Mellon University.

Spring 2008-present: Instructor for 90-866, Large Scale Data Analysis for Public Policy, Carnegie Mellon University.

Fall 2006-present: Instructor for 95-796, Statistics for IT Managers, Carnegie Mellon University.

Spring 2005: Head teaching assistant for 15-781, Machine Learning, Carnegie Mellon University.

Spring 2004: Teaching assistant for 15-780, Advanced AI Concepts, Carnegie Mellon University.

Summer 1998: University of South Florida, Mathematics and Engineering Program. Taught undergraduate-level courses in mathematics and computer science to gifted high school students.

Students Advised and Thesis Committees

Advisor for Skyler Speakman, Edward McFowland III, and Sriram Somanchi, Ph.D. students, H. John Heinz III College, Carnegie Mellon University.

Advisor for Kan Shao, M.S. in Machine Learning student, Carnegie Mellon University.

Advisor for Yandong Liu, Amrut Nagasunder, and Tarun Kumar, M.S. students, School of Computer Science, Carnegie Mellon University.

Thesis committee member for Stephen Fancsali, Ph.D. student, Department of Philosophy, Carnegie Mellon University.

Thesis committee member for Sharique Hasan, Ph.D. student, H. John Heinz III College, Carnegie Mellon University.

Thesis committee member for Kaustav Das, Ph.D. student, Machine Learning Department, School of Computer Science, Carnegie Mellon University.

Thesis committee member for Xia Jiang, Ph.D. student, Department of Biomedical Informatics, School of Medicine, University of Pittsburgh.

Thesis committee member for Sean Green, Ph.D. student, Department of Engineering and Public Policy, Carnegie Mellon University.

Heinz First Research Paper advisor for Chris Harle, Ph.D. student, H. John Heinz III College, Carnegie Mellon University.

Work-study advisor for Sayantan (Daemon) Das, undergraduate/MISM student, Carnegie Mellon University.

Professional Activities

Active participant in six large-scale efforts for development and deployment of disease surveillance systems, including the CDC BioSense project, the National Biosurveillance Integration System, the National Retail Data Monitor, and three Canadian disease surveillance projects funded by CRTI (ECADS, ASSET, and Data Fusion).

Established a new PhD program in Machine Learning and Public Policy at Carnegie Mellon University. This joint program between the Machine Learning Department and Heinz College (School of Public Policy) is the first of its kind. The program will build bridges and encourage collaborations between researchers in machine learning, computer science, public policy, and management. It will also attract and develop a highly competitive group of students with unique skills in developing new machine learning tools and applying them to real-world policy domains.

Developed a new course, “Large Scale Data Analysis for Public Policy,” for the Heinz College at Carnegie Mellon University. This master’s level course enables students to tackle a wide scope of policy problems using state-of-the-art machine learning methods.

Developed a new course, “Research Seminar in Machine Learning and Policy,” for the Heinz College and Machine Learning Department at Carnegie Mellon University. This Ph.D. level seminar prepares students for cutting-edge research at the intersection of machine learning and public policy through in-depth discussion of current research articles, essential topics, and ongoing projects.

Developed a new course series, “Special Topics in Machine Learning and Policy,” for the Heinz College and Machine Learning Department at Carnegie Mellon University. Topics covered include Event and Pattern Detection (2010), Machine Learning for the Developing World (2011), and Harnessing the Wisdom of Crowds (2012).

Scientific Program Chair for the *International Society for Disease Surveillance Annual Conference*, to be held in Atlanta, GA in December 2011. Also served as track chair (2007), program committee member (2007-2009), and abstract reviewer (2006-2011) for this conference.

Program committee member for the *23rd AAAI Conference on Artificial Intelligence* (2008), *25th International Conference on Machine Learning* (2008), *BioSecure Workshop* (2009), and *International Symposium on System Informatics and Engineering* (2011). Served on National Science Foundation grant review panel (CCF Division), 2009. “Artificial Intelligence and Health” Department Editor for *IEEE Intelligent Systems*, 2011.

Journal paper reviews for *Machine Learning*, *Applied Mathematical Modelling*, *IEEE Transactions on Evolutionary Computation*, *IEEE Transactions on Knowledge and Data Engineering*, *ACM Transactions on Knowledge Discovery in Data*, *Information Systems Research*, *PLoS Medicine*, *Management Science*, *Statistics in Medicine*, *Knowledge and Information Systems*, *Journal of Machine Learning Research*, *Artificial Intelligence and Law*, *PLoS Computational Biology*, *Theoretical Population Biology*, *Journal of the American Statistical Association*, *Environmetrics*, *Biometrics*, *International Journal of Health Geographics*, and *Systems, Man, and Cybernetics B*.

Conference paper reviews for *Neural Information Processing Systems* (2006), *American Medical Informatics Association Annual Symposium* (2007), *23rd AAAI Conference on Artificial Intelligence* (2008), *25th International Conference on Machine Learning* (2008), and *14th International Conference on Artificial Intelligence and Statistics* (2011).

Service on Carnegie Mellon committees: Heinz Scientific Computing Committee, Heinz Information Systems Management Faculty Hiring Committee, Heinz Ph.D. Committee. Co-organizer of Heinz College Faculty Research Seminars, 2011-2012.

Member of Sigma Xi, American Statistical Association, International Institute of Forecasters, and International Society for Disease Surveillance.

References

Andrew W. Moore

Professor (on LOA, serving as director of Google Pittsburgh)

School of Computer Science, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA 15213

Phone: (412) 268-7599, E-mail: awm@cs.cmu.edu

Tom Mitchell

Fredkin Professor and Department Head, Machine Learning Department

School of Computer Science, Carnegie Mellon University

5000 Forbes Avenue, Pittsburgh, PA 15213

Phone: (412) 268-2611, E-mail: tom.mitchell@cs.cmu.edu

Jeff Schneider

Associate Research Professor
School of Computer Science, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
Phone: (412) 268-2339, E-mail: schneide@cs.cmu.edu

Gregory F. Cooper
Associate Professor and Vice Chair of Biomedical Informatics
Department of Biomedical Informatics, University of Pittsburgh
M-183 Parkvale Building, 200 Meyran Avenue, Pittsburgh, PA 15260.
Phone: (412) 647-7113, E-mail: gfc@cbmi.pitt.edu

Artur Dubrawski
Systems Scientist and Director of the Auton Laboratory
School of Computer Science, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
Phone: (412) 268-6233, E-mail: awd@cs.cmu.edu