Center for Evidence-based Medicine
Brown University
Providence, RI 02912 USA
✉ byron_wallace@brown.edu
http://www.cebm.brown.edu/byron

Byron C. Wallace

Name, Position, Academic Departments

Name **Byron C. Wallace**

Position Assistant Professor (Research), Dept. of Health Services and Public Policy, Brown University

Research Interests

My research is in machine learning/data mining and natural language processing with an emphasis on applications in health informatics. I am particularly interested in using machine learning to mine, organize and filter clinical/biomedical texts, especially in the context of evidence-based medicine. There are too few experts to make sense of the torrents of published clinical data. I am interested in mitigating this problem by developing novel learning algorithms to induce models that semi-automate the clinical evidence synthesis process, thereby reducing workload. More broadly, I am interested in core machine learning issues: e.g., structured and unstructured classification techniques; semi-supervised learning methods; learning with imbalanced data; and learning from alternative forms of supervision. I am also broadly interested in computational methods for evidence-synthesis.

Education

2012 **Ph.D. Computer Science**, Tufts University, Medford, MA

Advisor Carla E. Brodley

Ph.D. Thesis *Machine Learning in Health Informatics: Making Better use of Domain Experts*

Committee Members Carla Brodley (Tufts), Roni Khardon (Tufts), Anselm Blumer (Tufts), Thomas A. Trikalinos (Brown) and Jaime Carbonell (CMU)

2006 **B.S. Computer Science**, University of Massachusetts at Amherst, Amherst, MA

Minor in Philosophy

Cumulative GPA 3.7 (*Baystate Scholar*)

Professional Experience

2012–*onwards* **Assistant Professor (Research)**, Center for Evidence-based Medicine, Brown University Providence, RI

2008–2012 **Research Computer Scientist**, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA

2007–2008 **Research Assistant (in the *machine learning group*)**, Department of Computer Science, Tufts University, Medford, MA

Summer 2006 **Software Engineer**, IBM, Westford, MA

2002–2006 **Research Assistant (in the *RIPPLEs group*)**, Department of Computer Science, UMass, Amherset, MA

Publications

Refereed Conference Publications

A preamble (especially for non-computer scientists): for historical reasons, conferences are the main venue of research dissemination in computer science. Conference proceedings comprise full-length articles and are peer-reviewed (usually by 2-3 reviewers). Acceptance rates for top-tier conference range from 10-25%.

1. Byron C Wallace and Issa J Dahabreh. Class probability estimates are unreliable for imbalanced data (and how to fix them). In *IEEE 12th International Conference on Data Mining (ICDM)*, pages 695–704. IEEE, 2012. Selected as one of the 'best of ICDM-2012'.
2. Byron C Wallace. Multiple narrative disentanglement: Unraveling infinite jest. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 1–10. ACL, 2012.
3. Byron C. Wallace, Kevin Small, Carla E. Brodley, Joseph Lau, and Thomas A. Trikalinos. Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium (IHI)*, pages 819–824. ACM, 2012.
4. Kevin Small, Byron C. Wallace, Carla E. Brodley, and Thomas A. Trikalinos. The constrained weight space SVM: Learning with ranked features. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 754–763, 2011.
5. Byron C. Wallace, K. Small, C.E. Brodley, and T.A. Trikalinos. Class imbalance, redux. In *IEEE 11th International Conference on Data Mining (ICDM)*, pages 754–763. IEEE, 2011.
6. Byron C. Wallace, Kevin Small, Carla E. Brodley, and Thomas A. Trikalinos. Who should label what? instance allocation in multiple expert active learning. In *Siam International Conference on Data Mining (SDM)*, pages 176–187. SIAM, 2011.
7. Byron C. Wallace, Kevin Small, Carla E. Brodley, Joseph Lau, and Thomas A. Trikalinos. Modeling annotation time to reduce workload in comparative effectiveness reviews. In *Proceedings of the 1st ACM International Health Informatics Symposium (IHI)*, pages 28–35. ACM, 2010.
8. Byron C. Wallace, Kevin Small, Carla E. Brodley, and Thomas A. Trikalinos. Active learning for biomedical citation screening. In *Proceedings of the 16th international conference on Knowledge discovery and data mining (KDD)*, pages 173–182. ACM, 2010.

Journal Articles

9. Byron C Wallace. Computational irony: A survey and new perspectives. *Artificial Intelligence Review*, pages 1–17, 2013.
10. Byron C Wallace and Issa J Dahabreh. Improving class probability estimates for imbalanced data. *Knowledge and Information Synthesis (KAIS)*, pages 1–20, 2013.
11. Byron C Wallace, Issa J Dahabreh, Chistopher H Schmid, Joseph Lau, and Thomas A Trikalinos. Modernizing the systematic review process to inform comparative effectiveness: tools and methods. *Journal of Comparative Effectiveness Review*, 2(3):273–282, 2013.
12. Carla E. Brodley, Umaa Rebbapragada, Kevin Small, and Byron C. Wallace. Challenges and opportunities in applied machine learning. *Artificial Intelligence Magazine*, 33(1):11–24, 2012.
13. Byron C. Wallace, Issa J. Dahabreh, Thomas A. Trikalinos, Joseph Lau, Paul Trow, and Christopher H. Schmid. Closing the gap between methodologists and end-users: R as a computational back-end. *Journal of Statistical Software*, 49(5):1–15, 6 2012.
14. Byron C Wallace, Kevin Small, Carla E Brodley, Joseph Lau, Christopher H Schmid, Lars Bertram, Christina M Lill, Joshua T Cohen, and Thomas A Trikalinos. Toward modernizing the systematic review pipeline in genetics: efficient updating via data mining. *Genetics in Medicine*, 14(7):663–669, 2012.
15. R.D. Whitaker, S. Pember, B.C. Wallace, C.E. Brodley, and D.R. Walt. Single cell time-resolved quorum responses reveal dependence on cell density and configuration. *Journal of Biological Chemistry*, 286(24):21623–21632, 2011.

16. PJ Castaldi, MH Cho, M Cohn, F Langerman, S Moran, N Tarragona, H Moukhachen, R Venugopal, D Hasimja, E Kao, BC Wallace, CP Hersh, S Bagade, L Bertram, EK Silverman, and TA Trikalinos. The copd genetic association compendium: a comprehensive online database of copd genetic associations. *Human molecular genetics*, 19(3):526–534, 2010.
17. Byron C. Wallace, Thomas A. Trikalinos, Joseph Lau, Carla E. Brodley, and Christopher H. Schmid. Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*, 11(1):55+, 2010.
18. Byron C. Wallace, Christopher H. Schmid, Joseph Lau, and Thomas A. Trikalinos. Meta-analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC medical research methodology*, 9(1):80+, 2009.

Refereed Workshop/Symposium Publications

1. Michael J. Paul, Byron C. Wallace, and Mark Dredze. What affects patient (dis)satisfaction? analyzing online doctor ratings with a joint topic-sentiment model. In *Proceedings of the AAAI Workshop on Expanding the Boundaries of Health Informatics Using AI (HIAI)*. AAAI, 2013.
2. Byron C. Wallace, Issa J. Dahabreh, Kelly H. Moran, Carla E. Brodley, and Thomas A. Trikalinos. Active literature discovery for scoping evidence reviews: How many needles are there? In *Proceedings of the KDD Workshop on Data Mining for Healthcare (KDD-DMH)*, 2013.
3. Byron C. Wallace, Kevin Small, Carla E. Brodley, and Thomas A. Trikalinos. Active learning for biomedical citation screening. In *Proceedings of the 2010 Northeastern Student Conference on Artificial Intelligence (NESCAI)*, 2010.

Selected Talks

Note that this is non-exhaustive and that these are in addition to the talks accompanying the conference publications above.

1. Byron C. Wallace. Semi-Automating Systematic Reviews: Text mining, NLP and machine learning. *#CochraneTech Symposium: Technology and the future of the systematic review*, upcoming, 2013.
2. Byron C. Wallace. Statistical models of patient-doctor communication. *Meaningful Use of Complex Medical Data*, upcoming, 2013.
3. Byron C. Wallace. Better models, less effort: Active learning and dual supervision. *Department of Computer Science at UMass, Boston*, 03/07/2012, 2012.
4. Byron C. Wallace. Machine learning in systematic reviews: Making better use of domain expertise. *Statistical Analysis of "Big Data" Group, Brown University*, 11/09/2012, 2012.
5. Byron C. Wallace. Open meta-analyst: open-source, cross-platform software for advanced meta-analysis. *Joint Colloquium of the Cochrane and Campbell Collaborations, Keystone, Colorado*, 2010.
6. Byron C. Wallace. Using machine learning to reduce the systematic review workload. *Joint Colloquium of the Cochrane and Campbell Collaborations, Keystone, Colorado*, 2010.
7. Byron C. Wallace, Christopher H. Schmid, Joseph Lau, Carla Brodley, and Thomas A. Trikalinos. Semi-automated screening of biomedical citations for systematic reviews. *4th Annual Meeting of the Society for Research Synthesis Methodology, Seattle, WA*, 2009.

Commentaries

1. Byron C. Wallace and Thomas A. Trikalinos. In Response to: Applications of text mining within systematic reviews. *Cochrane Methods*, 2012.

Peer Reviewing

Conference program committees on which I have served (or am serving): *2013 European Conference on Machine Learning (ECML), 2013 Association for the Advancement of Artificial Intelligence (AAAI), 2013 meeting of the American Medical Informatics Association (AMIA), 2013 IEEE Conference on Big Data, 2012 Knowledge Discovery and Databases (SIGKDD) Workshop on Health Informatics, 2012 ACM International Conference on Information and Knowledge Management (CIKM)*

Other conferences for which I have peer-reviewed (as an external reviewer) *2011 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)*

Some journals for which I regularly peer-review: *Agency for Healthcare Research and Quality (AHRQ), Bioinformatics, BMC Medical Informatics and Decision Making, BMC Medical Research Methodology, BMC Research Notes, Bioinformatics, Current Bioinformatics, Data Mining and Knowledge Discovery, IEEE Transactions on Knowledge and Data Engineering, Journal of Machine Learning Research (JMLR), Knowledge and Information Systems (KAIS), Research Synthesis Methods (RSM)*

Academic Honors & Professional Awards

Recipient of the *Outstanding Graduate Researcher at the Doctoral Level* award. Tufts University. 2012.

Recipient of the *Outstanding Student in the Area of Systems in Computer Science* award. UMass, Amherst. 2006.

Recipient of the 2005-2006 *Jonathon Edwards Philosophy Essay Prize*. UMass, Amherst. 2006.

Recipient of the 2005-2006 *Gerald F. Scanlon Student Employee of the Year* Award. UMass, Amherst. 2006.

Teaching

I was co-head instructor of CS150 AIH: *Artificial Intelligence in Health Informatics* alongside Dr. Kevin Small in the Computer Science department at Tufts University. Fall 2011. http://www.cs.tufts.edu/comp/150AIH/

Research Grants & Contracts

Grant Title	R01: Making Advanced Statistical Tools Accessible for Quantitative Research Synthesis and Discovery in Ecology and Evolutionary Biology.
Funder	National Science Foundation (NSF)
Role	Principal Investigator
Program	Collaborative research: ABI Development: Making Advanced Statistical Tools Accessible for Quantitative Research Synthesis and Discovery in Ecology and Evolutionary Biology.
Amount	~$500,000
Collaborators	Co-Principal Investigator Jessica Gurevitch (SUNY), Investigator Thomas Trikalinos, Investigator Christopher H. Schmid

Grant Title	R01 HS 018494: Semi-automating Citation Screening for Systematic Reviews
Funder	AHRQ
Amount	$1.2 million
Period	12/01/2009–11/30/2012
Role	Key Investigator
Collaborators	Principal Investigator Thomas A. Trikalinos, Investigator Joseph Lau

Grant Title	R01 HS 018574: Modernizing Meta-Analysis to Facilitate Comparative Effectiveness Reviews
Funder	AHRQ
Amount	$1.2 million
Period	12/01/2009–11/30/2012
Role	Key Investigator
Collaborators	Principal Investigator Christopher H. Schmid, Investigator Thomas A. Trikalinos, Investigator Joseph Lau

Proposals Submitted / Under Review

Grant Title Computational Methods to Optimize Evidence Identification for Systematic and Scoping Reviews. Submitted 3/2013.

Funder Patient Centered Outcomes Research Institute (PCORI).

Role Principal Investigator

Skills

Fluent (or at least was at one time...) in: Python, Java, C#

Experience with: R, Matlab/Octave, JAGS/(Win)BUGS, Scheme, C/C++, javascript