GUY LEBANON

1308 Klaus Building, 266 Ferst Drive
Georgia Institute of Technology
Atlanta GA 30332-0765 USA
www.cc.gatech.edu/~lebanon
lebanon (a) cc.gatech.edu
1-404-889-9838

A. ACADEMIC EDUCATION

1.	PhD	Carnegie Mellon University	School of Computer Science	2005
2.	MS	Carnegie Mellon University	School of Computer Science	2002
3.	MS	Technion - Israel Institute of Technology	Computer Science Department	2000
4.	BA	Technion - Israel Institute of Technology Summa-Cum Laude	Computer Science Department	1999

PhD Dissertation: Riemannian Geometry and Statistical Machine Learning
PhD Advisor: John Lafferty

B. RESEARCH APPOINTMENTS

1.	Georgia Institute of Technology, Atlanta GA USA *Associate Professor of Computing (with tenure)*	2011-
2.	Google, Mountain View CA USA *Visiting Scientist*	2012-2013
3.	Yahoo, Sunnyvale CA USA *Research Scientist (part time)*	2011-2012
4.	Georgia Institute of Technology, Atlanta GA USA *Assistant Professor of Computing*	2008-2011
5.	Purdue University, West Lafayette IN USA *Assistant Professor of Statistics (75%)* *Assistant Professor of Electrical and Computer Engineering (25%)*	2005-2008
6.	Carnegie Mellon University, Pittsburgh PA USA *Postdoctoral Researcher*	2005-2005

C. SELECTED AWARDS AND HONORS

1.	Program Chair, ACM CIKM Conference	2012
2.	Yahoo! Faculty Research and Engagement Award	2011
3.	Raytheon Faculty Fellowship Award	2010
4.	NSF Faculty Early Career (CAREER) Award	2008
5.	Class of 1969 Teaching Fellow, Georgia Institute of Technology	2008
6.	Teaching for Tomorrow Award, Purdue University	2007
7.	Best Presentation Award, LTI Student Research Symposium	2004
8.	Siebel Scholar, The Siebel Scholars Foundation	2004
9.	Bachelor Degree Awarded Summa Cum Laude, Technion	1999
10.	President's Award and Scholarship, Technion	1998

D. PROGRAMMING EXPERIENCE

1.	Technion Research Institute, Haifa, Israel	1998-1999
2.	RAFAEL Advanced Defense Systems, Israel	1997-1998

E. ADVISORY BOARDS

1.	Pave Inc., Atlanta GA, USA	2010-2012
2.	Reactive Search S.R.L., Trento, Italy	2010-2012
3.	Chatterspike Inc., Indianapolis, IN USA	2008-2009
4.	Vynate Inc., Indianapolis, IN USA	2006-2008

F. SPONSORED RESEARCH AND GIFTS

Lead Principal Investigator

1. *CAREER: Multi-Resolution Representations of Documents.* National Science Foundation (CISE/IIS/III). $405,548, 2008-2012.
2. *IPS: Decision Theoretic Approaches to Measuring and Minimizing Customized Privacy Risk*. National Science Foundation (CISE/IIS/IPS). $371,625, 2007-2010.
3. *Statistical Inference for Censored Preference Data.* National Science Foundation (MPS/DMS/Statistics). $175,881, 2009-2011.
4. *Machine Learning for Tied and Incomplete Preference Data*. US-Israel Binational Science Foundation (BSF). $150,000, 2011-2013.
5. *Assessing the Readability of Documents and Statistical Tools for non-Euclidean Data*. National Science Foundation (MPS/DMS/Statistics). $113,000, 2006-2008.
6. *Machine Learning and Visualization for Computational Journalism*. Raytheon Faculty Fellowship Award. Raytheon Corporation through the Georgia Institute of Technology. $53,805, 2010-2011.
7. *Supplementary CAREER Award Funding*. Office of the Provost, Georgia Institute of Technology. $20,000, 2008-2012.
8. *Isotonic Conditional Random Fields and Local Sentiment Flow.* Purdue Research Foundation. $14,627, *2007-2008.*
9. *The Mood Manifold and Mood Prediction.* Yahoo! Faculty Research and Engagement Award. $4,000 gift, 2011-2012.
10. *Class of 1969 Teaching Fellowship. Georgia Institute of Technology.* $1,000 gift, 2008-2009.
11. *Purdue Teaching for Tomorrow Award.* Purdue University. $1,000, 2006-2007.

Co-Principal Investigator

1. *Purdue Regional Visualization and Analytics Center*. US Department of Homeland Security. $750,000, 2006-2008. David Ebert (PI) with Guy Lebanon (co-PI) and other co-PIs.
2. *Systematic Control and Management of Data Integrity, Quality and Provenance for Command and Control Applications*. Air Force Office of Scientific Research, $300,000, 2007-2010. Elisa Bertino (PI) with Guy Lebanon (co-PI).

Senior Personnel

1. *FODAVA-Lead: Foundations of Data Analysis and Visual Analytics.* National Science Foundation and US Department of Homeland Security. $3,000,000, 2008-2012. Haesun Park (PI) with co-PIs and Guy Lebanon (senior personnel: associate director of FODAVA for education).
2. *Social Satellite Seedling*. Defense Advanced Research Projects Agency (DARPA). 2011-2011. Lora

Weiss (PI) with David Bader (co-PI) and G. Lebanon (senior personnel). Dr. Lebanon's part was $45,000

3. *SBIR Phase I: Software to Aggregate, Correlate, Analyze and Trend data for Knowledge Management in Decision Making*. National Science Foundation (IIP). $150,000. 2007-2008. Benjamin Ranck (PI) with Guy Lebanon (senior personnel).

Other Activities Related to Sponsored Research

1. Grant Proposal Reviewer and Panelist for the US Department of Energy
2. Grant Proposal Reviewer and Panelist for the US National Science Foundation
3. Grant Proposal Reviewer for the US National Security Agency
4. Grant Proposal Reviewer for the Maryland Technology Transfer Fund
5. Grant Proposal Reviewer for the Israel Science Foundation
6. Participated in an MSRI Workshop titled *The Mathematics of Visual Analysis* that produced recommendations for a joint call for proposals by the National Science Foundation and the Department of Homeland Security.

G. PROFESSIONAL SERVICE

Program Chair or Co-Chair

1. ACM International Conference on Information and Knowledge Management (CIKM), Maui HI USA 2012
2. Visual Analytics Education Workshop, collocated with the Visual Analytics Community Consortium Meeting. Hyattsville, MD USA 2010.
3. IEEE VisWeek Workshop on Scale and Complexity in Data and Visual Analytics. Salt Lake City, UT USA 2010
4. NIPS Workshop on Statistical Machine Learning for Visual Analytics. Whistler BC Canada 2009.
5. NIPS Workshop on Learning with Orderings. Whistler BC Canada 2009.
6. NIPS Workshop on Algebraic Methods in Machine Learning. Whistler BC Canada 2008.
7. NIPS Mini-Symposium on Algebraic Methods in Machine Learning. Vancouver BC Canada 2008.

Journal Editing

1. Guest Editor of *Data Mining and Knowledge Discovery* for Special Issue on Intelligent Interactive Data Visualization.

Area Chair or Senior Program Committee Member

1. Neural Information Processing Systems (NIPS) 2011
2. International Conference on Machine Learning (ICML) 2011
3. Twenty-Fifth Conference on Artificial Intelligence (AAAI) 2011
4. International Joint Conference on Artificial Intelligence (IJCAI) 2011
5. Uncertainty in Artificial Intelligence (UAI) 2009

Program Committee Member

1. AI & Statistics 2010
2. American Statistical Association Conference on Nonparametric Statistics and Statistical Learning 2010
3. Annual ACM SIGIR Conference 2010
4. Best Paper Award Selection Committee, American Statistical Association Section on Statistical Learning and Data Mining 2010.
5. Empirical Methods in Natural Language Processing (EMNLP) 2010
6. International Conference on Computational Linguistics (COLING) 2010

7. International Conference on Machine Learning (ICML) 2010
8. Learning on Cores, Clusters, and Clouds NIPS workshop 2010.
9. AI & Statistics 2009
10. Annual ACM SIGIR Conference 2009
11. International Conference on Machine Learning (ICML) 2009
12. ACM Conference on Information and Knowledge Management (CIKM) 2009
13. SIAM VizMining Workshop 2009
14. The 2nd Midwest Statistics Research Colloquium 2009
15. SIGIR Workshop Learning to Rank for Information Retrieval 2009
16. SIGIR Workshop on Redundancy, Diversity, and Interdependent Document Relevance 2009
17. ACM Conference on Information and Knowledge Management (CIKM) 2008
18. Empirical Methods in Natural Language Processing (EMNLP) 2008
19. International Conference on Machine Learning (ICML) 2008
20. SIGIR Workshop Learning to Rank for Information Retrieval 2008
21. SIGIR Workshop Beyond Relevance Feedback 2008
22. Uncertainty in Artificial Intelligence (UAI) 2008
23. The 23rd AAAI Conference on Artificial Intelligence 2008
24. International Symposium on Artificial Intelligence and Mathematics 2008
25. AI & Statistics 2007
26. Combined European Conference on Machine Learning (ECML) and European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD) 2007
27. International Conference on Data Mining (ICDM) 2007
28. International Conference on Machine Learning (ICML) 2007
29. Joint Conference on Empirical Methods in Natural Language Processing (EMNLP) and Conference on Computational Natural Language Learning (CONLL) 2007
30. SIGIR Workshop Learning to Rank for Information Retrieval 2007
31. Uncertainty in Artificial Intelligence (UAI) 2007
32. International Conference on Data Mining (ICDM) 2006
33. Joint Conference of the Association for Computational Linguistics (ACL) and the International Conference on Computational Linguistics (COLING) 2006
34. Learning to Compare Examples (NIPS workshop) 2006
35. Uncertainty in Artificial Intelligence (UAI) 2006
36. AI & Statistics 2005
37. Uncertainty in Artificial Intelligence (UAI) 2005

Conference Reviewing (excluding program committees listed above)

ACM SIGIR	2007
ACM SIGMOD	2008

AMS-IMS-SIAM Conference on Machine and Statistical Learning	2006
ECML	2002
IEEE Information Visualization Conference	2009
IEEE Symposium on Visual Analytics Science and Technology	2009, 2010
NIPS	2002, 2003, 2004, 2005, 2006, 2007, 2008
NIPS "power reviewer"	2009, 2010
USENIX Workshop on Hot Topics in Parallelism	2011

Journal Reviewing

ACM Journal of Data and Information Quality, Artificial Intelligence Journal, Applied Optics, Computational Statistics and Data Analysis, IEEE Transactions on Evolutionary Computation, IEEE Transactions on Information Theory, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Secure and Dependable Computing, IEEE Transactions on Signal Processing, IEEE Transactions on Visualization and Computer Graphics, Journal of Artificial Intelligence Research, Journal of Machine Learning Research, Journal of the American Statistical Association, Journal of the Optical Society of America, Knowledge and Information Systems, Machine Learning, Mathematical Programming, Neural Computation, Neurocomputing, Optics Letters.

Textbook Reviewing

Cambridge University Press
McGraw Hill
Morgan and Claypool
Springer

H. TEACHING

Evidence of Teaching Effectiveness

1. Received the Class of 1969 Teaching Fellowship and participated in a year long program for enhancing teaching and learning. Georgia Institute of Technology, 2008-2009.
2. Received the 2007 Teaching for Tomorrow Award from the Provost of Purdue University
3. Average instructor rating by students greater than 4.00/5.00

Instructor

1. Computing and Society, Georgia Institute of Technology, Spring 2012
2. Machine Learning / Computational Data Analysis, Georgia Institute of Technology, Fall 2011
3. Data and Visual Analytics, Georgia Institute of Technology, Spring 2011
4. Computational Data Analysis, Georgia Institute of Technology, Fall 2010
5. Data and Visual Analytics, Georgia Institute of Technology, Spring 2010
6. Combinatorial Methods in Non-Parametric Density Estimation, Georgia Institute of Technology, Spring 2009
7. Data Structures and Algorithms, Georgia Institute of Technology, Fall 2008
8. Graphical Models in Machine Learning. Georgia Institute of Technology, Fall 2008
9. Introduction to Computational Statistics, Purdue University, Spring 2008
10. Wavelets and Multiresolution Analysis, Spring 2008
11. Signals and Systems, Purdue University, Fall 2007
12. Information Theory, Machine Learning, and Statistics, Purdue University, Spring 2007
13. Introduction to Computational Statistics, Purdue University, Fall 2006
14. Statistical Machine Learning, Purdue University, Fall 2006
15. Statistical Theory, Purdue University, Spring 2006
16. Statistical Machine Learning, Purdue University, Spring 2006
17. Probabilistic Methods in Electrical and Computer Engineering, Purdue University, Fall 2005

Teaching Assistant

1. Probability and Statistics for Computer Science, Carnegie Mellon University, Fall 2002
2. Language and Statistics, Carnegie Mellon University, Spring 2001
3. Computer Vision, Technion, Spring 2000
4. Digital Image Processing, Technion, Spring 2000
5. Introduction to Digital Signal and Image Processing, Technion, Fall 1999

Other Teaching Activities

1. Participated in a workshop on curriculum design for the new discipline of data and visual analytics.
2. Designed a new core course in computational statistics aimed at statistics graduate students
3. Designed a new PhD qualifying examination in computational statistics
4. Participated in a MSRI workshop on the future of computing in statistics education
5. Attended the following teaching seminars at the Eberly Center for Teaching Excellence at Carnegie Mellon University

    Course and Syllabus Design, Creating a Teaching Portfolio, Overview of Student Cognition, Overview of Student Motivation, Creating Effective Assignments and Exams, Promoting Meaningful and Engaged Knowledge through Service Learning, Encouraging Intellectual Development and Critical Thinking, Conducting Productive and Engaging Discussions.

I. CURRENT GRADUATE STUDENTS AND POSTDOCS

Post-Doc

1. Fuxin Li

PhD Students

1. Krishnakumar Balasubramanian
2. Seungyeon Kim
3. Joonseok Lee
4. Mingxuan Sun

MS Students

5. Sanjeet Hajarnis
6. Kaushik Rangadurai

J. FORMER GRADUATE STUDENTS

Former PhD Students

1. Joshua Dillon (Google)
2. Paul Kidwell (Lawrence Livermore National Lab)
3. Yi Mao (Microsoft)
4. Yang Zhao (Google)

Former MS Students

1. Krishnakumar Balasubramanian (Georgia Institute of Technology)
2. Joshua Dillon (Google)
3. Seungyeon Kim (Georgia Institute of Technology)
4. Yi Mao (Microsoft)
5. Yanjun Zhao (Citadel Investment Group)

K. INVITED TALKS AT CONFERENCES, SYMPOSIA, AND WORKSHOPS

1. Machine learning for Information Visualization. Refereed tutorial at *IEEE VisWeek*. Salt Lake City UT. October 2010.
2. Experiences and Conclusions from Teaching Visual Analytics at Georgia Tech. *Visual Analytics Curriculum Workshop, collocated with the annual VAC consortium meeting*. Hyattsville, MD. August 2010 (also workshop organizer).
3. Visualizing Similarities of Search Engines using the Weighted Hoeffding Distance on Permutations. *ASA Conference on Non-Parametric Statistics and Statistical Learning*. Columbus OH. May 2010 (also session organizer).
4. Visualizing Similarities of Search Engines using the Weighted Hoeffding Distance on Permutations. *NIPS Workshop on Learning with Ordering*. Whistler, BC. December 2009 (also workshop organizer)
5. New Directions in Text Visualization. *NIPS Workshop on Statistical Machine Learning for Visual Analytics.* Whistler BC. December 2009 (also workshop organizer).
6. Modeling and Visualization of Missing Preference Data. *The Joint Statistical Meeting*. Washington DC. August 2009.
7. Models on Permutations and Censored Preference Data. *Neural Information Processing Systems Mini-Symposium*. Vancouver BC. December 2008 (also mini-symposium organizer).
8. Riemannian Metrics for Image Spaces. *IEEE Statistical Signal Processing Workshop*. Madison, WI. August 2007.
9. Efficient and Coherent Framework for Aggregating Ranking Data. *39th Symposium on the Interface of Statistics, Computing Science, and Applications*. Philadelphia, PA. May 2007.
10. Expected geometry and statistical translation in text analysis. *SAMSI Workshop on Geometry, Random Matrices, and Statistical Inference.* Research Triangle Park, NC. January 2007.
11. Visualizing Heterogeneous Data. *MSRI Workshop on Mathematics of Visual Analytics*. Berkeley, CA. October 2006.
12. Conditional Models on the Ranking Poset. *38th Symposium on the Interface of Statistics, Computing Science, and Applications*. Pasadena, CA. May 2006.
13. Information Geometry and Classification of Text Documents. *2nd Symposium on Information Geometry and its Applications*. Tokyo University, Japan. December 2005.
14. A Unifying View of Classification and Ranking. *Workshop on Reductions in Machine Learning*. Toyota Technological Institute. Chicago, IL. September 2003.
15. Conditional Models on the Ranking Poset. *NIPS Beyond Classification and Regression Workshop*. Whistler, BC Canada. December 2002.

L. INVITED TALKS AT COLLOQUIA AND SEMINARS

1. Estimating Probabilities in Recommendation Systems. *Yahoo! Research*. New York City USA, August 2011.
2. Unsupervised Supervised Learning: Who Needs Labels Anyway? *The Hebrew University*. Jerusalem Israel, May 2011.
3. Unsupervised Supervised Learning: Who Needs Labels Anyway? *Technion – Israel Institute of Technology*. Haifa Israel, May 2011.
4. Unsupervised Supervised Learning: Who Needs Labels Anyway? *University of California*. Irvine CA, April 2010.
5. Unsupervised Supervised Learning: Who Needs Labels Anyway? *Google Research*. Mountain View CA, April 2010.
6. Non-Parametric Modeling and Visualization of Partially Ranked Data. School of Computer Science and Engineering, *Hebrew University*. Jerusalem, Israel, June 2008.
7. Non-Parametric Modeling of Partially Ranked Data. Statistics Department, *The Ohio State University*. Columbus, OH, April 2008.
8. The Locally Weighted Bag of Words Representation for Documents. Electrical Engineering Department, *University of Washington*. Seattle, WA, November 2007.
9. The Locally Weighted Bag of Words Representation for Documents. Computer Science Department, *University of California*, Berkeley, CA, November 2007.
10. Beyond k-Anonymity: A Decision Theoretic Approach to Privacy Preservation. *Accenture Technology Labs*. Chicago, IL, November 2007.
11. The Locally Weighted Bag of Words Representation for Documents. Language Technologies Institute, *Carnegie Mellon University*. Pittsburgh, PA, October 2007.
12. Non-Parametric Modeling of Partially Ranked Data. Intelligence Seminar, *Carnegie Mellon University*. Pittsburgh, PA, October 2007.
13. Sequential Document Visualization. Computer Science Department, *University of Wisconsin*. Madison, WI, August 2007.
14. Visualizing Text using the Lowbow Representation. *Google*. Mountain View, CA, October 2006.
15. Sequential Representation of Documents and Simplicial Curves. *Toyota Technological Institute*. Chicago, IL, May 2006.
16. Information Geometry and Classification of Text Documents. Computer Science and Engineering Department, *Michigan State University*. November 2005.
17. Riemannian Geometry and Text Classification. Information Retrieval Group, *IBM Research Lab*. Haifa Israel, January 2005.
18. Boosting and Maximum Likelihood for Exponential Models. Computer Science Department, *Technion – Israel Institute of Technology*. July 2003.
19. Boosting and Maximum Likelihood for Exponential Models. School of Mathematical Science, *Tel-Aviv University*, Israel. July 2003.

20. Boosting and Maximum Likelihood for Exponential Models. School of Computer Science and Engineering, *Hebrew University*. Jerusalem, Israel, July 2003.
21. Boosting and Maximum Likelihood for Exponential Models. Machine Learning and Applied Statistics Group, *Microsoft Researc*h. Redmond, WA, August 2002.
22. Conditional Models for Ranked Data. Statistics Department, *The Ohio State University*. Columbus, OH, 2002.

M. PUBLICATIONS

Theses

[M1] G. Lebanon. *Riemannian Geometry and Statistical Machine Learning*. PhD Dissertation. Carnegie Mellon University, Technical Report CMU-LTI-05-189, 2005.

[M2] G. Lebanon and J. Lafferty. *Boosting and Maximum Likelihood for Exponential Models.* MS Project Report. Carnegie Mellon University, Technical Report CMU-CS-01-144, 2001.

[M3] G. Lebanon. *A Variational Approach to Moiré Pattern Synthesis*. MS Thesis. Technion – Israel Institute of Technology, 2000.

Book

[L1] G. Lebanon. *The Analysis of Data: A Comprehensive Introduction to Probability, Statistics, and Machine Learning*. Expected publication date: August 2012.

Refereed Journal Papers

[J1] M. Sun, G. Lebanon, and P. Kidwell. Estimating Probabilities in Recommendation Systems. *Journal of the Royal Statistical Society* 2012.

[J2] K. Balasubramanian, P. Donmez, and G. Lebanon. Unsupervised Supervised Learning II: Margin-Based Classification without Labels. *Journal of Machine Learning Research* **12**(Dec) 2011.

[J3] P. Kidwell, G. Lebanon, and K. Collins-Thompson. Statistical Estimation of Word Acquisition with Application to Readability Prediction. *Journal of the American Statistical Association* **106**(493):21-30, 2011.

[J4] J. V. Dillon and G. Lebanon. Stochastic Composite Likelihood. *Journal of Machine Learning Research* **11**(Oct):2597-2633, 2010.

[J5] P. Donmez, K. Balasubramanian, and G. Lebanon. Unsupervised Supervised Learning I: Estimating Classification and Regression Errors without Labels. *Journal of Machine Learning Research* **11**(April):1323-1351, 2010.

[J6] G. Lebanon, Y. Zhao, and Y. Zhao. Modeling Temporal Text Streams using the Local Multinomial Model. *Electronic Journal of Statistics* **4**:566-584, 2010.

[J7] Y. Mao and G. Lebanon. Generalized Isotonic Conditional Random Fields. *Machine Learning* **77**(2-3):225-248, 2009.

[J8] D. J. Kasik, D. Ebert, G. Lebanon, H. Park, and W. M. Pottenger. Data Transformations and Representations for Information Generation. *Information Visualization* **8**(4):275-285, 2009.

[J9] G. Lebanon, M. Scannapieco, M. R. Fouad, and E. Bertino. Beyond *k*-Anonymity: A Decision Theoretic Framework for Assessing Privacy Risk. *Transactions on Data Privacy* **2**(3):153-183, 2009.

[J10] E. Greenshtein, J. Park, and G. Lebanon. Regularization through Variable Selection and

Conditional MLE with Application to Classification in High Dimensions. *Journal of Statistical Planning and Inference* **139**(2):385-395, 2009.

[J11] P. Kidwell, G. Lebanon and W. S. Cleveland. Visualizing Incomplete and Partially Ranked Data. *IEEE Transactions on Visualization and Computer Graphics* (Proc. INFOVIS) **14**(6):1356-1363, 2008.

[J12] G. Lebanon and Y. Mao. Non-Parametric Modeling of Partially Ranked Data. *Journal of Machine Learning Research* **9**(Oct): 2401-2429, 2008.

[J13] G. Lebanon, Y. Mao, and J. Dillon. The Locally Weighted Bag of Words Framework for Document Representation. *Journal of Machine Learning Research* **8**(Oct):2405-2441, 2007.

[J14] Y. Mao, J. Dillon, and G. Lebanon. Sequential Document Visualization. *IEEE Transactions on Visualization and Computer Graphics* (Proc. INFOVIS) **13**(6):1208-1215, 2007.

[J15] G. Lebanon. Metric Learning for Text Documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(4):497-508, 2006.

[J16] G. Lebanon. Axiomatic Geometry of Conditional Models. *IEEE Transactions on Information Theory* **51**(4):1283-1294, 2005.

[J17] J. Lafferty and G. Lebanon. Diffusion Kernels on Statistical Manifolds. *Journal of Machine Learning Research* **6**(Jan):129-163, 2005.

[J18] G. Lebanon and A. Bruckstein. Variational Approach to Moiré Pattern Synthesis. *Journal of the Optical Society of America A* **18**(6):1371-1382, 2001.

Refereed Book Chapters and Papers in Special Volumes

[B1] P. Kidwell and G. Lebanon. Kernel Smoothing for Preference Data using Generating Functions. In M. Vienna and H. Wynn (eds). *Algebraic Methods in Statistics and Probability II.* Contemporary Mathematics Series, The American Mathematical Society 2010.

[B2] G. Lebanon. Axiomatic Geometries for Text Documents. In P. Gibilisco, E. Riccomagno, M.-P. Rogantin, and H. P. Wynn (eds). *Algebraic and Geometric Methods in Statistics.* Cambridge University Press 2009.

[B3] G. Lebanon and A. Bruckstein. On Designing Moiré Patterns. In V. Cantoni, M. Marinaro, and A. Petrosino (eds), *Visual Attention Mechanism*, pages 205-219, Springer 2002.

Refereed Conference Papers

[C1] M. Sun, G. Lebanon, and P. Kidwell. Estimating Probabilities in Recommendation Systems. *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics* (AISTATS) JMLR W&CP 2011.

[C2] K. Balasubramanian, P. Donmez, and G. Lebanon. Unsupervised Supervised Learning II: Margin-Based Classification without Labels. *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics* (AISTATS) JMLR W&CP 2011

[C3] J. V. Dillon, K. Balasubramanian, and G. Lebanon. Asymptotic Analysis of Generative Semi-

Supervised Learning. *Proceedings of the 27th International Conference on Machine Learning* (ICML) 2010.

[C4] M. Sun, G. Lebanon, and K. Collins-Thompson. Visualizing Search Strategies in the World Wide Web. *Proceedings of the 19th International World Wide Web Conference* (WWW) 2010.

[C5] S. Kim and G. Lebanon. Local Space-Time Smoothing for Version Controlled Documents. *Proceedings of the International Conference on Computational Linguistics* (COLING) 2010.

[C6] Y. Mao, K. Balasubramanian, and G. Lebanon. Linguistic Geometries for Unsupervised Dimensionality Reduction. *Proceedings of the International Conference on Computational Linguistics* (COLING) 2010.

[C7] Y. Mao and G. Lebanon. Domain Knowledge Uncertainty and Probabilistic Parameter Constraints. *Proceedings of the 25rd Conference on Uncertainty in Artificial Intelligence* (UAI) AUAI Press 2009.

[C8] P. Kidwell, G. Lebanon, and K. Collins-Thompson. Statistical Estimation of Word Acquisition with Application to Readability Prediction. *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP) 2009.

[C9] J. V. Dillon and G. Lebanon. Statistical and Computational Tradeoffs in Stochastic Composite Likelihood. *Proceedings of the 12th International Conference on Artificial Intelligence and Statistics.* (AISTATS) JMLR W&CP 5:129-139, 2009.

[C10] G. Lebanon and Y. Zhao. Local Likelihood Modeling of the Concept Drift Phenomenon. *Proceedings of the 25th International Conference in Machine Learning* (ICML) 2008.

[C11] G. Lebanon and Y. Mao. Non-Parametric Modeling of Partially Ranked Data. *Advances in Neural Information Processing Systems* 20 (NIPS). The MIT Press 2008.

[C12] G. M. Howard, S. Bagchi, and G. Lebanon. Determining placement of intrusion detectors for a distributed application through Bayesian network modeling. *Proceedings of the 11th International Symposium on Recent Advances in Intrusion Detection* (RAID). Lecture Notes in Computer Science, volume 5230, pages 271-290. Springer, 2008.

[C13] Y. Mao and G. Lebanon. Isotonic Conditional Random Fields and Local Sentiment Flow. *Advances in Neural Information Processing Systems* 19 (NIPS), pages 961-968. The MIT Press 2007.

[C14] J. V. Dillon, Y. Mao, G. Lebanon and J. Zhang. Statistical Translation, Heat Kernels, and Expected Distances. *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence* (UAI) pages 93-100. AUAI Press 2007.

[C15] G. Lebanon. Sequential Document Representations and Simplicial Curves. *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence* (UAI) pages 273-280. AUAI Press 2006.

[C16] G. Lebanon, M. Scannapieco, M. R. Fouad, and E. Bertino. Beyond *k*-Anonymity: A Decision Theoretic Framework for Assessing Privacy Risk. *Privacy in statistical databases.* Lecture Notes in Computer Science, volume 4302, pages 217-232. Springer, 2006.

[C17] G. Lebanon. Information Geometry, the Embedding Principle, and Document Classification. *Proceedings of the 2nd International Symposium on Information Geometry and its Applications*, pages 101-108, 2005.

[C18] G. Lebanon. An Extended Čencov-Campbell Characterization of Conditional Information Geometry. *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence* (UAI) AUAI Press 2004.

[C19] G. Lebanon and J. Lafferty. Hyperplane Margin Classifiers on the Multinomial Manifold. *Proceedings of the 21st International Conference on Machine Learning* (ICML). Morgan Kaufmann Press 2004.

[C20] G. Lebanon and J. Lafferty. Conditional Models on the Ranking Poset. *Advances in Neural Information Processing Systems* 15 (NIPS), pages 431-438. The MIT Press 2003.

[C21] J. Lafferty and G. Lebanon. Information Diffusion Kernels. *Advances in Neural Information Processing Systems* 15 (NIPS), pages 391-399. The MIT Press 2003.

[C22] G. Lebanon. Learning Riemannian Metrics. *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence* (UAI). AUAI Press 2003.

[C23] G. Lebanon and J. Lafferty. Cranking: Combining Rankings using Conditional Probability Models on Permutations. *Proceedings of the 19th International Conference on Machine Learning* (ICML). Morgan Kaufmann Press 2002.

[C24] G. Lebanon and J. Lafferty. Boosting and Maximum Likelihood for Exponential Models. *Advances in Neural Information Processing Systems* 14 (NIPS), pages 447-454. The MIT Press 2002.

[C25] G. Lebanon and A. Bruckstein. Designing Moiré Patterns. *Energy Minimization Methods in Computer Vision and Pattern Recognition* (EMMCVPR). Lecture Notes in Computer Science, Volume 2134, pages 185-201. Springer 2001.

Refereed Workshop Papers and Unrefereed Technical Reports

[W1] M. Sun, G. Lebanon, and K. Collins-Thompson. Visualizing Spatial Proximity of Search Algorithms. *NIPS Workshop on Learning with Ordering*, 2009.

[W2] M. Fouad, G. Lebanon, and E. Bertino. ARUBA: A Risk-Utility-Based Algorithm for Data Disclosure. *Proceedings of the 5th VLDB Workshop on Secure Data Management* (SDM). Lecture Notes in Computer Science, Springer 2008.

[W3] J. V. Dillon, Y. Mao, G. Lebanon, and J. Zhang. Statistical Translation, Heat Kernels, and Expected Distances. *NIPS workshop on Learning to Compare Examples*, 2006.

[W4] Y. Mao and G. Lebanon. Sequential Models for Sentiment Prediction. *ICML workshop on Learning in Structured Output Spaces*, 2006.

[W5] Z. Zhang, M. Gupta, S. Yang, G. Lebanon, Y. C. Hu, and S. Midkiff. *Extracting Source Level Program Similarities from Dynamic Behavior*. Technical Report TR-ECE-08-08, Purdue University, 2008.

[W6] G. Lebanon. *Computing the Volume Element of a Family of Transformations on the Multinomial Simplex*. Technical Report CMU-CS-03-145, Carnegie Mellon University 2003.