

**DAVID M. BLEI**  
35 Olden Street  
Princeton, NJ  
blei@cs.princeton.edu

## Academic Positions

**Assistant Professor** **January 2006 – present**  
**Department of Computer Science, Princeton University, Princeton NJ**

**Postdoctoral Fellow** **August 2004 – December 2005**  
**Machine Learning Department, Carnegie Mellon University, Pittsburgh PA**  
*Advisor:* Prof. John Lafferty

## Education

**University of California, Berkeley, Berkeley CA** **September 1999 – August 2004**  
**Ph.D., Computer Science**  
*Dissertation:* “Probabilistic models of text and images”  
*Committee:* Prof. Michael Jordan (advisor), Prof. Stuart Russell, Prof. Peter Bickel

**Brown University, Providence RI** **January, 1993 – May, 1997**  
**B.Sc. (Honors), Computer Science and Mathematics**  
*Honors thesis:* “Shortest paths in a dynamic uncertain domain”  
*Advisor:* Prof. Leslie Kaelbling

## Awards

<i>E. L. Keyes, Jr. Emerson Electric Co. Faculty Award</i>	2008
<i>National Science Foundation CAREER Award</i>	2008
<i>Microsoft New Faculty Fellowship Finalist</i>	2007
<i>Microsoft Research Award</i>	2007
<i>Google Research Award</i>	2006, 2007
<i>U.C. Berkeley C.V. Ramamoorthy Distinguished Research Award</i>	2006
<i>NSF International Postdoctoral Fellowship (declined)</i>	2004
<i>Best Student Paper Award, Neural Information Processing Systems</i>	2003
<i>Microsoft Research Fellowship</i>	2002
<i>Berkeley Micro-Electronics Fellowship</i>	1999
<i>Sigma Xi Scientific Honor Society</i>	1997

## Refereed Conference Publications

- J. Boyd-Graber and D. Blei. Multilingual topic models for unaligned text. In *Uncertainty in Artificial Intelligence*, 2009.
- J. Chang, J. Boyd-Graber, and D. Blei. Connections between the lines: Augmenting social networks with text. In *Knowledge Discovery and Data Mining*, 2009.
- C. Wang, D. Blei, and F. Li. Simultaneous image classification and annotation. In *Computer Vision and Pattern Recognition*, 2009.
- J. Chang and D. Blei. Relational topic models for document networks. In *Artificial Intelligence and Statistics*, 2009.
- C. Wang, B. Thiesson, C. Meek, and D. Blei. Markov topic models. In *Artificial Intelligence and Statistics*, 2009.
- J. Boyd-Graber and D. Blei. Syntactic topic models. In *Advances in Neural Information Processing Systems*, 2008.
- E. Airoldi, D. Blei, S. Fienberg, and E. Xing. Mixed membership stochastic blockmodels. In *Advances in Neural Information Processing Systems*, 2008.
- I. Mukherjee and D. Blei. Relative performance guarantees for approximate inference in latent Dirichlet allocation. In *Advances in Neural Information Processing Systems*, 2008.
- D. Blei and J. McAuliffe. Supervised topic models. In *Advances in Neural Information Processing Systems 20*, 2007.
- J. Boyd-Graber, D. Blei, and X. Zhu. A topic model for word sense disambiguation. In *Empirical Methods in Natural Language Processing*, 2007.
- M. Dudik, D. Blei, and R. Schapire. Hierarchical maximum entropy density estimation. In *Proceedings of the 28th International Conference on Machine Learning*, 2007.
- D. Kaplan and D. Blei. A computational approach to style in American poetry. In *IEEE Conference on Data Mining*, 2007.
- W. Li, A. McCallum, and D. Blei. Nonparametric Bayes pachinko allocation. In *The 23rd Conference on Uncertainty in Artificial Intelligence*, 2007.
- D. Blei and J. Lafferty. Dynamic topic models. In *Proceedings of the 23rd International Conference on Machine Learning*, 2006.
- D. Blei and J. Lafferty. Correlated topic models. In *Advances in Neural Information Processing Systems 18 (NIPS) 18*, pages 147–154. MIT Press, 2005.
- T. Griffiths, M. Steyvers, D. Blei, and J. Tenenbaum. Integrating topics and syntax. In *Advances in Neural Information Processing Systems 17*, 2004.
- D. Blei and M. Jordan. Variational methods for the Dirichlet process. In *21st International Conference on Machine Learning*, 2004.
- D. Blei, T. Griffiths, M. Jordan, and J. Tenenbaum. Hierarchical topic models and the nested Chinese restaurant process. In *Neural Information Processing Systems (NIPS) 16*, 2003. (Best paper award.).
- D. Blei and M. Jordan. Modeling annotated data. In *Proceedings of the 26th annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 127–134. ACM Press, 2003.
- D. Blei, A. Ng, and M. Jordan. Latent Dirichlet allocation. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems (NIPS) 12*, Cambridge, MA, 2002. MIT Press.

D. Blei, J. Bagnell, and A. McCallum. Learning with scope, with application to information extraction and classification. In *Uncertainty in Artificial Intelligence: Proceedings of the Eighteenth Conference (UAI-2002)*, pages 53–60, San Francisco, CA, 2002. Morgan Kaufmann Publishers.

D. Blei and P. Moreno. Topic segmentation with an aspect hidden markov model. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 343–348. ACM Press, 2001.

## Refereed Journal Publications

E. Airoldi, D. Blei, S. Fienberg, and E. Xing. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, 9:1981–2014, 2008.

D. Blei and J. Lafferty. A correlated topic model of *Science*. *Annals of Applied Statistics*, 1(1), 2007.

D. Blei and S. Fienberg. Discussion of model-based clustering for social networks. *Journal of the Royal Statistical Society, Series A*, 170:332, 2007.

Y. Teh, M. Jordan, M. Beal, and D. Blei. Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2007.

J. McAuliffe, D. Blei, and M. Jordan. Nonparametric empirical Bayes for the Dirichlet process mixture model. *Statistics and Computing*, 16(1):5–14, 2006.

D. Blei, K. Franks, M. Jordan, and S. Mian. Statistical modeling of biomedical corpora: mining the Caenorhabditis Genetic Center Bibliography for genes related to life span. *BMC Bioinformatics*, 7(250), 2006.

D. Blei and M. Jordan. Variational inference for Dirichlet process mixtures. *Journal of Bayesian Analysis*, 1(1):121–144, 2005.

D. Blei, A. Ng, and M. Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, January 2003.

K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3:1107–1135, 2003.

## Book Chapters

D. Blei. Discussion. In *Statistical Network Analysis: Models, Issues and New Directions*, Lecture Notes in Computer Science, pages 186–187. Springer-Verlag, 2007.

E. Airoldi, D. Blei, S. Fienberg, and E. Xing. Combining stochastic block models and mixed membership for statistical network analysis. In *Statistical Network Analysis: Models, Issues and New Directions*, Lecture Notes in Computer Science, pages 57–74. Springer-Verlag, 2007.

D. Blei, A. Ng, and M. Jordan. Hierarchical Bayesian models for applications in information retrieval. In J. Bernardo, J. Berger, A. Dawid, D. Heckerman, A. Smith, and M. West, editors, *Bayesian Statistics 7*, volume 7, pages 25–44. Oxford University Press, 2003.

## Edited Volume

E. Airoldi, D. Blei, S. Fienberg, A. Goldenberg, E. Xing, and A. Zheng, editors. *Statistical Network Analysis: Models, Issues and New Directions*. Lecture Notes in Computer Science. Springer-Verlag, 2007.

## Professional Activities

**Senior program committee:** Neural Information Processing Systems (2009), International Conference for Machine Learning (ICML) (2006, 2007, 2008), AI and Statistics (2008), Bayesian Case Studies (2009)

**Co-organizer:** 2008 NIPS Workshop on Analyzing Graphs, 2006 ICML Workshop on the Statistics of Social Networks; 2003 NIPS Workshop on Syntax, Semantics, and Statistics

**Book reviewing:** Springer Statistics Series

**Journal reviewing:** Journal of Machine Learning Research, Machine Learning Journal, Journal of Bayesian Analysis, Journal of the American Statistical Association, Journal of the Royal Statistical Society, Annals of Applied Statistics, IEEE Transactions on Pattern Analysis and Machine Intelligence, Proceedings of the National Academy of Science, IEEE Transactions on Neural Networks, International Journal on Very Large Data Bases, IEEE Transactions on Image Processing, IEEE Transactions on Audio, Speech, and Language Processing, ACM Transactions on Knowledge Discovery from Data

**Conference reviewing:** Neural Information Processing Systems (NIPS), Artificial Intelligence and Statistics (AI-STATS), International Joint Conference on Artificial Intelligence (IJCAI), American Association of Artificial Intelligence (AAAI), North American Association of Computational Linguistics (NAACL), Association of Computational Linguistics (ACL), International Conference of Machine Learning (ICML), Uncertainty in Artificial Intelligence (UAI), Empirical Methods in Natural Language Processing (EMNLP), ACM SIGIR Conference on Research and Development in Information Retrieval

**Ph.D. thesis committees:** Jordan Boyd-Graber (Princeton, expected 2009), Jonathan Chang (Princeton, expected 2009), Lauren Hanna (Princeton, expected 2009), Vasileios Kandylas (University of Pennsylvania, expected 2009), Emily Fox (MIT, expected 2009), Chenwei Zhu (Princeton, 2008), Zafer Barutcuoglu (Princeton, 2008), Katherine Heller (University College London, 2008), Suhrid Balakrishnan (Rutgers, 2007), Wei Li (UMass Amherst, 2007), Miroslav Dudik (Princeton, 2007)

**Editorial Board:** Journal of Machine Learning Research, Journal of Artificial Intelligence Research, Information Retrieval (2006-2008), Chapman Hall Series on Computer Science and Data Analysis

**Princeton:** Executive Committee for the Committee on Statistical Studies, Program in Applied and Computational Mathematics, Affiliated Faculty: Center for Information Technology Policy, Affiliated Faculty: Neuroscience, Advisor to Computer Science A.B. classes of 2009 and 2010

**Professional Memberships:** Institute of Mathematical Statistics, Bernoulli Society, DIMACS

**Invited Talks:** Carnegie Mellon University (2003, 2005, 2009), Harvard University (2009), Massachusetts Institute of Technology (2003, 2007, 2008, 2009), Rutgers University (2007, 2009), Nature Publishing Group (2008), University of Chicago Econometrics Colloquium (2008), BAE Systems (2008), University of Cambridge (2008), University College London (2008), Cornell University (2007), Columbia University (2007), U.C. Irvine (2007), Google (2004, 2006, 2007), Microsoft Research (2007), University of Connecticut (2006, 2007), University of Pennsylvania (2006), Joint Statistics Meetings (2006), Educational Testing Service (2006), IPAM (2006, 2007), Duke University (2006), Johns Hopkins University (2006), UIUC (2006), U.C. Santa Cruz (2005), U.C. San Diego (2005), U.C.L.A. (2005), Princeton University (2005), Brown University (2005), New York University (2005), UMass Amherst (2004), University of Toronto (2003), AAAI Spring Symposium (2002), Xerox PARC (2002), IBM Almaden (2002)

**Grant reviewing:** NSF Panel (IIS)