Jordan, M. I. (2011). Message From the President: The era of Big Data. *The ISBA Bulletin*, *18*(2), 1-3.

Jordan, M. I. (2011). Message From the President: What are the open problems in Bayesian statistics? *The ISBA Bulletin*, *18*(1), 1-4.

Jordan, M. I. (2010). Leo Breiman. *Annals of Applied Statistics*, *4*, 1642-1643.

Bartlett, P., Jordan, M. I., & McAuliffe, J. D. (2006). Discussion of "Support vector machines with applications." *Statistical Science*, *21*, 341-346.

Bartlett, P., Jordan, M. I., & McAuliffe, J. (2004). Discussion of boosting papers. *Annals of Statistics*, *32*, 85-91.

REFEREED CONFERENCE PROCEEDINGS

Lindsten, F., Schön, T., & Jordan, M. I. (2012). A semiparametric Bayesian approach to Wiener system identification. 16th IFAC Symposium on System Identification (SYSID 2012), Brussels, Belgium.

Paisley, J., Blei, D. & Jordan, M. I. (2012). Stick-breaking beta processes and the Poisson process. In N. Lawrence and M. Girolami (Eds.), *Proceedings of the Fifteenth Conference on Artificial Intelligence and Statistics (AISTATS)*, Canary Islands.

Wauthier, F. & Jordan, M. I. (2012). Bayesian bias mitigation for crowdsourcing. In J. Shawe-Taylor, R. Zemel, P. Bartlett and F. Pereira (Eds.), *Advances in Neural Information Processing (NIPS) 24*, Red Hook, NY: Curran Associates.

Mackey, L., Talwalkar, A. & Jordan, M. I. (2012). Divide-and-conquer matrix factorization. In J. Shawe-Taylor, R. Zemel, P. Bartlett and F. Pereira (Eds.), *Advances in Neural Information Processing (NIPS) 24*, Red Hook, NY: Curran Associates.

Duchi, J., Agarwal, A., Johansson, M., & Jordan, M. I. (2011). Ergodic subgradient descent. *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL.

Liang, P., Jordan, M. I. & Klein, D. (2011). Learning dependency-based compositional semantics. *49th Annual Meeting of the Association for Computational Linguistics (ACL)*, Portland, OR.

Trushkowsky, B., Bodik, P., Fox, A., Franklin, M., Jordan, M. I., & Patterson, D. (2011). Scaling a distributed storage system under stringent performance requirements. *9th USENIX Conference on File and Storage Technologies (FAST '11)*, San Jose, CA.

Shyr, A., Darrell, T., Jordan, M. I. & Urtasun, R. (2011). Supervised hierarchical Pitman-Yor process for natural scene segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, CO.

Bouchard, A. & Jordan, M. I. (2011). Variational inference over combinatorial spaces. In J. Lafferty, C. Williams, J. Shawe-Taylor and R. Zemel (Eds.), *Advances in Neural Information Processing (NIPS) 23*, Red Hook, NY: Curran Associates.

Kleiner, A., Rahimi, A. & Jordan, M. I. (2011). Random conic pursuit for semidefinite programming. In J. Lafferty, C. Williams, J. Shawe-Taylor and R. Zemel (Eds.), *Advances in Neural Information Processing (NIPS) 23*, Red Hook, NY: Curran Associates.

Wauthier, F. & Jordan, M. I. (2011). Heavy-tailed processes for selective shrinkage. In J. Lafferty, C. Williams, J. Shawe-Taylor and R. Zemel (Eds.), *Advances in Neural Information Processing (NIPS) 23*, Red Hook, NY: Curran Associates.

Adams, R., Ghahramani, Z. & Jordan, M. I. (2011). Tree-structured stick breaking for hierarchical data. In J. Lafferty, C. Williams, J. Shawe-Taylor and R. Zemel (Eds.), *Advances in Neural Information Processing (NIPS) 23*, Red Hook, NY: Curran Associates.

Wang, M., Sha, F. & Jordan, M. I. (2011). Unsupervised kernel dimension reduction. In J. Lafferty, C. Williams, J. Shawe-Taylor and R. Zemel (Eds.), *Advances in Neural Information Processing (NIPS) 23*, Red Hook, NY: Curran Associates.

Wauthier, F., Jordan, M. I., & Jojic, N. (2011). Nonparametric combinatorial sequence models. *15th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, Springer Lecture Notes in Bioinformatics.

Niu, D., Dy, J. & Jordan, M. I., (2011). Dimensionality reduction for spectral clustering. G. Gordon and D. Dunson (Eds.), *Proceedings of the Fourteenth Conference on Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL.

Wang, P., Laskey, K. B., Domeniconi, C., & Jordan, M. I. (2011). Nonparametric Bayesian co-clustering ensembles. *SIAM International Conference on Data Mining (SDM)*, Phoenix, AZ.

Guan, Y., Dy, J., & Jordan, M. I. (2011). A unified probabilistic model for global and local unsupervised feature selection. In L. Getoor and T. Scheffer (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Chowdhury, M., Zaharia, M., Ma, J., Jordan, M. I. & Stoica, I. (2011). Managing data transfers in computer clusters with Orchestra. *ACM SIGCOMM*, Toronto, CA.

Huang, M.-Y., Mackey, L., Keranen, S., Weber, G., Jordan, M. I., Knowles, D., Biggin, M., & Hamann, B. (2011). Visually relating gene expression and in vivo DNA binding data. *IEEE International Conference on Bioinformatics and Biomedicine (IEEE BIBM)*, Atlanta, GA.

Liang, P., Jordan, M. I. & Klein, D. (2010). Learning programs: A hierarchical Bayesian approach. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Duchi, J., Mackey, L. & Jordan, M. I. (2010). On the consistency of ranking algorithms. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Mackey, L., Weiss, D. & Jordan, M. I. (2010). Mixed membership matrix factorization. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Niu, D., Dy, J., & Jordan, M. I. (2010). Multiple non-redundant spectral clustering views. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Ting, D., Huang, L. & Jordan, M. I. (2010). An analysis of the convergence of graph Laplacians. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Xu, W., Huang, L., Fox, A., Patterson, D., & Jordan, M. I. (2010). Detecting large-scale system problems by mining console logs. In T. Joachims and J. Fuernkranz (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Simma, A. & Jordan, M. I. (2010). Modeling events with cascades of Poisson processes. In P. Grunwald and P. Spirtes (Eds.), *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence (UAI)*, AUAI Press.

Shyr, A., Urtasun, R., & Jordan, M. I. (2010). Sufficient dimension reduction for visual sequence classification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, CA.

Bodik, P., Fox, A., Franklin, M., Jordan, M. I., & Patterson, D. (2010). Characterizing, modeling, and generating workload spikes for stateful services. *First ACM Symposium on Cloud Computing (SOCC)*, Indianapolis, IN.

Sutton, C. A. & Jordan, M. I., (2010). Inference and learning in networks of queues. In Y. W. Teh and M. Titterington (Eds.), *Proceedings of the Thirteenth Conference on Artificial Intelligence and Statistics (AISTATS)*, Sardinia, Italy.

Zhang, Z., Dai, G., Wang, D., & Jordan, M. I., (2010). Bayesian generalized kernel models. In Y. W. Teh and M. Titterington (Eds.), *Proceedings of the Thirteenth Conference on Artificial Intelligence and Statistics (AISTATS)*, Sardinia, Italy.

Zhang, Z., Dai, G., & Jordan, M. I., (2010). Matrix-variate Dirichlet process mixture models. In Y. W. Teh and M. Titterington (Eds.), *Proceedings of the Thirteenth Conference on Artificial Intelligence and Statistics (AISTATS)*, Sardinia, Italy.

Xu, W., Huang, L., Fox, A., Patterson, D., & Jordan, M. I. (2010). Experience mining Google's production console logs. In *Proceedings of the 2010 Workshop on Managing Systems via Log Analysis and Machine Learning Techniques (SLAML)*, Vancouver, BC.

Fox, E. B., Sudderth, E., Jordan, M. I., & Willsky, A. S. (2010). Sharing features among dynamical systems with beta processes. In Y. Bengio, D. Schuurmans, J. Lafferty and C. Williams (Eds.), *Advances in Neural Information Processing (NIPS) 22*, Red Hook, NY: Curran Associates.

Miller, K., Griffiths, T., & Jordan, M. I. (2010). Nonparametric latent feature models for link prediction. In Y. Bengio, D. Schuurmans, J. Lafferty and C. Williams (Eds.), *Advances in Neural Information Processing (NIPS) 22*, Red Hook, NY: Curran Associates.

Liang, P., Bach, F., Bouchard, G., & Jordan, M. I. (2010). An asymptotic analysis of smooth regularizers. In Y. Bengio, D. Schuurmans, J. Lafferty and C. Williams (Eds.), *Advances in Neural Information Processing (NIPS) 22*, Red Hook, NY: Curran Associates.

Xu, W., Huang, L., Fox, A., Patterson, D., & Jordan, M. I. (2009). Large-scale system problems detection by mining console logs. In *22nd ACM Symposium on Operating Systems Principles (SOSP)*, Big Sky, MT.

Liang, P., Jordan, M. I., & Klein, D. (2009). Learning semantic correspondences with less supervision. *47th Annual Meeting of the Association for Computational Linguistics (ACL)*, Singapore.

Fox, E. B., Sudderth, E., Jordan, M. I., & Willsky, A. S. (2009). Nonparametric Bayesian identification of jump systems with sparse dependencies. *Proceedings of the 15th IFAC Symposium on System Identification*, Saint-Malo, France.

Yan, D., Huang, L., & Jordan, M. I. (2009). Fast approximate spectral clustering. *15th ACM Conference on Knowledge Discovery and Data Mining (SIGKDD)*, Paris, France.

Bouchard-Cote, A. & Jordan, M. I. (2009). Optimization of structured mean field objectives. In J. Bilmes and A. Ng (Eds.), *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI)*, AUAI Press.

Liang, P., Klein, D., & Jordan, M. I. (2009). Learning from measurements in exponential families. In L. Bottou and M. Littman (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Sudderth, E. & Jordan, M. I. Shared segmentation of natural scenes using dependent Pitman-Yor processes. (2009). In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Bouchard-Cote, A., Jordan, M. I., & Klein, D. (2009). Efficient inference in phylogenetic InDel trees. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Obozinski, G., Wainwright, M. J., & Jordan, M. I. (2009). High-dimensional union support recovery in multivariate regression. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Fox, E. B., Sudderth, E., Jordan, M. I., & Willsky, A. S. (2009). Nonparametric Bayesian learning of switching linear dynamical systems. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Zhang, Z. & Jordan, M. I. (2009). Posterior consistency of the Silverman g-prior in Bayesian model choice. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Huang, L., Yan, D., Jordan, M. I., & Taft, N. (2009). Spectral clustering with perturbed data. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Lacoste-Julien, S., Sha, F., & Jordan, M. I. (2009). DiscLDA: Discriminative learning for dimensionality reduction and classification. In D. Koller, Y. Bengio, D. Schuurmans and L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 21*, Red Hook, NY: Curran Associates.

Xu, W., Huang, L., Fox, A., Patterson, D. and Jordan, M. I. (2009). Online system problem detection by mining patterns of console logs. *IEEE International Conference on Data Mining (ICDM)*, Miami, FL.

Zhang, Z., Dai, G., & Jordan, M. I. (2009). A flexible and efficient algorithm for regularized Fisher discriminant analysis. In W. Buntine, M. Grobelnik, D. Mladenic, J. Shawe-Taylor (Eds.), *Machine Learning and Knowledge Discovery in Databases: European Conference (ECML PKDD)*, Bled, Slovenia.

Ganapathi, A., Kuno, H., Dayal, U., Wiener, J., Fox, A., Jordan, M. I., and Patterson, D. (2009). Predicting multiple metrics for queries: Better decisions enabled by machine learning. In *25th IEEE International Conference on Data Engineering (ICDE)*, Shanghai, China.

Zhang, Z., Jordan, M. I., Li, W-J., & Yeung, D-Y. (2009). Coherence functions for multicategory margin-based classification methods. In D. van Dyk and M. Welling (Eds.), *Proceedings of the Twelth Conference on Artificial Intelligence and Statistics (AISTATS)*, Clearwater Beach, FL.

Bodik, P., Griffith, R., Sutton, C., Fox, A., Jordan, M. I., and Patterson, D. (2009). Statistical machine learning makes automatic control practical for Internet datacenters. *Workshop on Hot Topics in Cloud Computing (HotCloud)*, San Diego, CA.

Engelhardt, B., Jordan, M. I., Repo, S., & Brenner, S. (2009). Phylogenetic molecular function annotation. *Journal of Physics: Conference Series, 180*, 012024.

Zhang, Z. & Jordan, M. I. (2009). Latent variable models for dimensionality reduction. In D. van Dyk and M. Welling (Eds.), *Proceedings of the Twelth Conference on Artificial Intelligence and Statistics (AISTATS)*, Clearwater Beach, FL.

Bodik, P., Griffith, R., Sutton, C., Fox, A., Jordan, M. I., and Patterson, D. (2009). Automatic exploration of datacenter performance regimes. *Proceedings of the First Workshop on Automated Control for Datacenters and Clouds (ACDC)*, Barcelona, Spain.

Miller, K., Griffiths, T. & Jordan, M. I. (2008). The phylogenetic Indian buffet process: A non-exchangeable nonparametric prior for latent features. In D. McAllister (Ed.), *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence (UAI)*, AUAI Press.

Liang, P. & Jordan, M. I. (2008). An analysis of generative, discriminative, and pseudo-likelihood estimators. In *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Fox, E., Sudderth, E., Jordan, M. I. & Willsky, A. (2008). An HDP-HMM for systems with state persistence. In *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Blum, B., Jordan, M. I., Kim, D., Das, R. Bradley, P., Das, R. & Baker, D. (2008). Feature selection methods for improving protein structure prediction with Rosetta. In J. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), *Advances in Neural Information Processing (NIPS) 20*, 137–144, Red Hook, NY: Curran Associates.

Liang, P., Klein, D., & Jordan, M. I. (2008). Agreement-based learning. In J. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), *Advances in Neural Information Processing (NIPS) 20*, 913–920, Red Hook, NY: Curran Associates.

Nguyen, X., Wainwright, M., & Jordan, M. I. (2008). Estimating divergence functionals and the likelihood ratio by penalized convex risk minimization. In J. Platt, D. Koller, Y. Singer, and S. Roweis (Eds.), *Advances in Neural Information Processing (NIPS) 20*, 1089–1096, Red Hook, NY: Curran Associates.

Sankararaman, S., Kimmel, G., Halperin, E. & Jordan, M. I. (2008). On the inference of ancestries in admixed populations. *12th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, Springer Lecture Notes in Bioinformatics.

Obozinski, G., Wainwright, M., & Jordan, M. I. (2008). Union support recovery in high-dimensional multivariate regression. *Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL.

Ding, C., Li, T., & Jordan, M. I. (2008). Nonnegative matrix factorization for combinatorial optimization: Spectral clustering, graph matching, and clique finding. *IEEE International Conference on Data Mining*, Pisa, Italy.

Sutton, C. & Jordan, M. I. (2008). Probabilistic inference in queueing networks. In *Workshop on Tackling Computer Systems Problems with Machine Learning Techniques (SYSML)*, San Diego, CA.

Xu, W., Huang, L. Fox, A., Patterson, D., & Jordan, M. I. (2008). Mining console logs for large-scale system problem detection. In *Workshop on Tackling Computer Systems Problems with Machine Learning Techniques (SYSML)*, San Diego, CA.

Thibaux, R. & Jordan, M. I. (2007). Hierarchical beta processes and the Indian buffet process. In M. Meila and X. Shen (Eds.), *Proceedings of the Eleventh Conference on Artificial Intelligence and Statistics (AISTATS)*, Puerto Rico.

Kivinen, J., Sudderth, E., & Jordan, M. I. (2007). Learning multiscale representations of natural scenes using Dirichlet processes. *IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brazil.

Liang, P., Petrov, S., Klein, D., & Jordan, M. I. (2007). The infinite PCFG using Dirichlet processes. *Empirical Methods in Natural Language Processing (EMNLP)*. Prague, Czech Republic.

Liang, P., Jordan, M. I., & Taskar, B. (2007). A permutation-augmented sampler for DP mixture models. In Z. Ghahramani (Ed.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Nilsson, J., Sha, F., & Jordan, M. I. (2007). Regression on manifolds using kernel dimension reduction. In Z. Ghahramani (Ed.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Nguyen, X., Wainwright, M., & Jordan, M. I. (2007). Nonparametric estimation of the likelihood ratio and divergence functionals. In *International Symposium on Information Theory (ISIT)*, Nice, France.

Kivinen, J., Sudderth, E., & Jordan, M. I. (2007). Image denoising with nonparametric hidden Markov trees. *IEEE International Conference on Image Processing (ICIP)*, San Antonio, TX.

Huang, L., Nguyen, X., Garofalakis, M., Hellerstein, J., Jordan, M. I., Joseph, A., & Taft, N. (2007). Communication-efficient online detection of network-wide anomalies. In *26th Annual IEEE Conference on Computer Communications (INFOCOM)*, Anchorage, AS.

Huang, L., Nguyen, X., Garofalakis, M., Jordan, M. I., Joseph, A., & Taft, N. (2007). In-network PCA and network anomaly detection. In T. Hofmann, J. Platt and B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 19*, Cambridge, MA: MIT Press.

Li, T., Ding, C., & Jordan, M. I. (2007). Solving consensus and semi-supervised clustering problems using nonnegative matrix factorization *IEEE International Conference on Data Mining (ICDM)*, Omaha, NE.

Bodik, P., Sutton, C., Fox, A., Jordan, M. I., & Patterson, D. (2007). Response-time modeling for resource allocation and energy-informed SLAs. *Workshop on Statistical Learning Techniques for Solving Systems Problems (SYSML)*, Whistler, BC.

Nguyen, X., Wainwright, M., & Jordan, M. I. (2006). On optimal quantization rules for sequential decision problems. In *International Symposium on Information Theory (ISIT)*, Seattle, WA.

Zhang, Z., & Jordan, M. I. (2006). Bayesian multicategory support vector machines. In R. Dechter & T. Richardson (Eds.), *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI)*. AUAI Press.

Xing, E. P., Sohn, K-A., Jordan, M. I., & Teh, Y. W. (2006). Bayesian multi-population haplotype inference via a hierarchical Dirichlet process mixture. In W. Cohen & A. Moore (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Engelhardt, B., Jordan, M. I., & Brenner, S. (2006). A statistical graphical model for predicting protein molecular function. In W. Cohen & A. Moore (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Zheng, A., Jordan, M. I., Liblit, B., Naik, M., & Aiken, A. (2006). Statistical debugging: Simultaneous identification of multiple bugs. In W. Cohen & A. Moore (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Lacoste-Julien, S., Taskar, B., Klein, D. & Jordan, M. I. (2006). Word alignment via quadratic assignment. In J. Bilmes, J. Chu-Carroll, & M. Sanderson (Eds.), *Proceedings of the North American Chapter of the Association for Computational Linguistics Annual Meeting (HLT-NAACL).*

Flaherty, P., Jordan, M. I., & Arkin, A. P. (2006). Robust design of biological experiments. In Y. Weiss and B. Schölkopf and J. Platt (Eds.), *Advances in Neural Information Processing (NIPS) 18*, Cambridge, MA: MIT Press.

Nguyen, X., Wainwright, M., & Jordan, M. I. (2006). Divergence measures, surrogate loss functions and experimental design. In Y. Weiss and B. Schölkopf and J. Platt (Eds.), *Advances in Neural Information Processing (NIPS) 18*, Cambridge, MA: MIT Press.

Taskar, B., Lacoste-Julien, S., & Jordan, M. I. (2006). Structured prediction via the extragradient method. In Y. Weiss and B. Schölkopf and J. Platt (Eds.), *Advances in Neural Information Processing (NIPS) 18*, Cambridge, MA: MIT Press.

Bodik, P., Fox, A., Jordan, M. I., Patterson, D., Banerjee, A., Jagannathan, R., Su, T., Tenginakai, S., Turner, B., & Ingalls, J. (2006). Advanced tools for operators at Amazon.com. *Workshop on Hot Topics in Autonomic Computing*, Dublin, Ireland.

Rosen-Zvi, M., Jordan, M. I., & Yuille, A. (2005). The DLR hierarchy of approximate inference. In M. Chickering, F. Bacchus, & T. Jaakkola (Eds.), *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence (UAI).* San Mateo, CA: Morgan Kaufmann.

Bach, F. R., & Jordan, M. I. (2005). Predictive low-rank decomposition for kernel methods. In S. Dzeroski, L. De Raedt, & S. Wrobel (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Bodik, P., Friedman, G., Biewald, L., Levine, H., Candea, G., Patel, K., Tolle, G., Hui, J., Fox, A., Jordan, M. I., & Patterson, D. (2005). Combining visualization and statistical analysis to improve operator confidence and efficiency for failure detection and localization. In *International Conference on Autonomic Computing (ICAC)*, Seattle, WA.

Liblit, B., Naik, M., Zheng, A. X., Aiken, A., & Jordan, M. I. (2005). Scalable statistical bug isolation. In V. Sarkar & M. W. Hall (Eds.), *ACM SIGPLAN 2005 Conference on Programming Language Design and Implementation (PLDI).*

Nguyen, X., Wainwright, M., & Jordan, M. I. (2005). On information divergence measures, surrogate loss functions and decentralized hypothesis testing. *Proceedings of the 43rd Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL.

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2005). Sharing clusters among related groups: Hierarchical Dirichlet processes. In L. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 17*, Cambridge, MA: MIT Press.

Lawrence, N. D. & Jordan, M. I. (2005). Semi-supervised learning via Gaussian processes. In L. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 17*, Cambridge, MA: MIT Press.

Bach, F. R., Thibaux, R., & Jordan, M. I. (2005). Computing regularization paths for learning multiple kernels. In L. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 17*, Cambridge, MA: MIT Press.

Bach, F. R., & Jordan, M. I. (2005). Blind one-microphone speech separation: A spectral learning approach. In L. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 17*, Cambridge, MA: MIT Press.

D'Aspremont, A., El Ghaoui, L., Jordan, M. I., & Lanckriet, G. R. G. (2005). A direct formulation for sparse PCA using semidefinite programming. In L. Saul, Y. Weiss & L. Bottou (Eds.), *Advances in Neural Information Processing (NIPS) 17*, Cambridge, MA: MIT Press.

Bach, F. R., & Jordan, M. I. (2005). Discriminative training of Hidden Markov Models for multiple pitch tracking. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE Press.

Vogel, B., Jordan, M. I., & Wessel, D. (2005). Multi-instrument musical transcription using a dynamic graphical model. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE Press.

Teh, Y. W., Seeger, M., & Jordan, M. I. (2005). Semiparametric latent factor models. In R. Cowell and Z. Ghahramani (Eds.), *Proceedings of the Ninth Conference on Artificial Intelligence and Statistics (AISTATS)*, Barbados.

Fox, A., Kiciman, E., Patterson, D., Jordan, M. I., & Katz, R. (2004). Combining statistical monitoring and predictable recovery for self-management. In D. Garlan, J. Kramer, A. Wolf (Eds.). *ACM SIGSOFT Proceedings of the Workshop on Self-Managed Systems (WOSS)*.

Bach, F. R., Lanckriet, G. R. G., & Jordan, M. I. (2004). Multiple kernel learning, conic duality, and the SMO algorithm. In C. E. Brodley (Ed.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Nguyen, X., Wainwright, M., & Jordan, M. I. (2004). Decentralized detection and classification using kernel methods. In C. E. Brodley (Ed.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Blei, D. M., & Jordan, M. I. (2004). Variational methods for the Dirichlet process. In C. E. Brodley (Ed.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.

Xing, E. P., Sharan, R., & Jordan, M. I. (2004). Haplotype modeling via the Dirichlet process. *RECOMB Workshop on Computational Methods for SNPs and Haplotypes*, Springer-Verlag: New York.

Xing, E. P., Jordan, M. I., & Russell, S. (2004). Graph partition strategies for generalized mean field inference. In M. Chickering & J. Halpern (Eds.), *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Chen, M. Zheng, A. X., Lloyd, J., Jordan, M. I., & Brewer, E. (2004). Failure diagnosis using decision trees. In *International Conference on Autonomic Computing (ICAC)*, New York, NY: IEEE Computer Society.

Liblit, B., Naik, M., Zheng, A. X., Aiken, A., & Jordan, M. I. (2004). Public deployment of cooperative bug isolation. *Workshop on Remote Analysis and Measurement of Software Systems (RAMSS)*, Edinburgh, UK.

Lanckriet, G. R. G., Deng, M., Cristianini, N., Jordan, M. I., & Noble, W. S. (2004). Kernel-based data fusion and its application to protein function prediction in yeast. In R. Altman, A. Dunker, L. Hunter, T. Jung, & T. Klein (Eds.), *Pacific Symposium on Biocomputing (PSB)*, Hawaii.

Bartlett, P., Jordan, M. I., & McAuliffe, J. (2004). Large margin classifiers: convex loss, low noise, and convergence rates. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Ng, A., Kim, H. J., Jordan, M. I., & Sastry, S. (2004). Autonomous helicopter flight via reinforcement learning. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Wainwright, M., & Jordan, M. I. (2004). Semidefinite relaxations for approximate inference on graphs with cycles. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Zheng, A. X., Jordan, M. I., Liblit, B., & Aiken, A. (2004). Statistical debugging of sampled programs. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Blei, D., Griffiths, T., Jordan, M. I., & Tenenbaum, J. (2004). Hierarchical topic models and the nested Chinese restaurant process. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Fukumizu, K., Bach, F. R., & Jordan, M. I. (2004). Kernel dimensionality reduction for supervised learning. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Bach, F. R., & Jordan, M. I. (2004). Learning spectral clustering. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Nguyen, X., & Jordan, M. I. (2004). On the concentration of expectation and approximate inference in layered Bayesian networks. In S. Thrun, L. Saul & B. Schölkopf (Eds.), *Advances in Neural Information Processing (NIPS) 16*, Cambridge, MA: MIT Press.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., & Sastry, S. (2003). Kalman filtering with intermittent observations. *Proceedings of the 42nd IEEE Conference on Decision and Control (CDC)*, Hawaii.

Wainwright, M., & Jordan, M. I. (2003). Variational inference in graphical models: The view from the marginal polytope. *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL.

Blei, D., Jordan, M. I., & Ng, A. (2003). Hierarchical Bayesian models for applications in information retrieval. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith & M. West (Eds.), *Bayesian Statistics 7*, 25–43, Oxford: Oxford University Press.

Liblit, B., Aiken, A., Zheng, A. X., & Jordan, M. I. (2003). Bug isolation via remote program sampling. *ACM SIGPLAN 2003 Conference on Programming Language Design and Implementation (PLDI)*, San Diego, CA.

Xing, E. P., Jordan, M. I., & Russell, S. (2003). A generalized mean field algorithm for variational inference in exponential families. In C. Meek and U. Kjaerulff (Eds.), *Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Blei, D., & Jordan, M. I. (2003). Modeling annotated data. In E. N. Efthimiadis, S. T. Dumais, D. Hawking, K. Järvelin (Eds.), *Proceedings of the 26th International Conference on Research and Development in Information Retrieval (SIGIR)*. New York, NY: ACM Press.

Wainwright, M. Jordan, M. I. (2003). Semidefinite relaxations for approximate inference on graphs with cycles. In *International Symposium on Information Theory (ISIT)*, Tokyo, Japan.

Bach, F. R., & Jordan, M. I. (2003). Finding clusters in independent component analysis. In *Fourth International Symposium on Independent Component Analysis and Blind Source Separation (ICA)*, Tokyo, Japan.

Xing, E. P., Wu, W., Jordan, M. I., & Karp, R. M., (2003). LOGOS: A modular Bayesian model for *de novo* motif detection. In *IEEE Computer Society Bioinformatics Conference (CSB)*, Palo Alto, CA.

Liblit, B., Aiken, A., Zheng, A. X., & Jordan, M. I. (2003). Sampling user executions for bug isolation. *Workshop on Remote Analysis and Measurement of Software Systems (RAMSS)*, Portland, OR.

De Bernardinis, F., Jordan, M. I., & Sangiovanni Vincentelli, A. L. (2003). Support vector machines for analog circuit performance representation. In *Design Automation Conference (DAC)*, San Jose, CA.

Bach, F. R., & Jordan, M. I. (2003). Kernel ICA. In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE Press.

Todorov, E., & Jordan, M. I. (2003). A minimal intervention principle for coordinated movement. In S. Becker, S. Thrun & K. Obermayer (Eds.), *Advances in Neural Information Processing (NIPS) 15*, Cambridge, MA: MIT Press.

Lanckriet, G. R. G., El Ghaoui, L., & Jordan, M. I. (2003). Robust novelty detection with single-class MPM. In S. Becker, S. Thrun & K. Obermayer (Eds.), *Advances in Neural Information Processing (NIPS) 15*, Cambridge, MA: MIT Press.

Xing, E. P., Jordan, M. I., Karp, R. M., & Russell, S. (2003). A hierarchical Bayesian Markovian model for motifs in biopolymer sequences. In S. Becker, S. Thrun & K. Obermayer (Eds.), *Advances in Neural Information Processing (NIPS) 15*, Cambridge, MA: MIT Press.

Fukumizu, K., Bach, F. R., & Jordan, M. I. (2003). Feature extraction in regression with kernel Hilbert spaces. *Workshop on Information-Based Induction Sciences (IBIS)*, Kyoto, Japan.

Bach, F. R., & Jordan, M. I. (2003). Learning graphical models with Mercer kernels. In S. Becker, S. Thrun & K. Obermayer (Eds.), *Advances in Neural Information Processing (NIPS) 15*, Cambridge, MA: MIT Press.

Xing, E. P., Ng, A. Y., & Jordan, M. I. (2003). Distance metric learning, with application to clustering with side-information. In S. Becker, S. Thrun & K. Obermayer

(Eds.), *Advances in Neural Information Processing (NIPS) 15*, Cambridge, MA: MIT Press.

Bach, F. R., & Jordan, M. I. (2003). Analyse en composantes indépendantes et réseaux Bayésiens. *Dix-Neuvième Colloque GRETSI sur le Traitement du Signal et des Images*, Paris, France.

Tatikonda, S., & Jordan, M. I. (2002). Loopy belief propagation and Gibbs measures. *Proceedings of the 40th Annual Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL.

Micheli, M., & Jordan, M. I. (2002). Random sampling of a continuous-time stochastic dynamical system. *Proceedings of the Fifteenth International Symposium on Mathematical Theory of Networks and Systems.* Notre Dame, IN.

Bach, F. R., & Jordan, M. I. (2002). Tree-dependent component analysis. In D. Koller & N. Friedman (Eds.), *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI).* San Mateo, CA: Morgan Kaufmann.

Grate, L. R., Bhattacharyya, C., Jordan, M. I., & Mian, I. S. (2002). Simultaneous relevant feature identification and classification in high-dimensional spaces. *Workshop on Algorithms in Bioinformatics (WABI)*, Rome, Italy.

Ng, A., Jordan, M. I., & Weiss, Y. (2002). On Spectral Clustering: Analysis and an algorithm. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing (NIPS) 14*, Cambridge, MA: MIT Press.

Lanckriet, G. R. G., El Ghaoui, L., Bhattacharyya, C., & Jordan, M. I. (2002). Minimax probability machine. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing (NIPS) 14*, Cambridge, MA: MIT Press.

Blei, D., Ng, A., & Jordan, M. I. (2002). Latent Dirichlet allocation. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing (NIPS) 14*, Cambridge, MA: MIT Press.

Bach, F. R., & Jordan, M. I. (2002). Thin junction trees. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing (NIPS) 14*, Cambridge, MA: MIT Press.

Lanckriet, G. R. G., Cristianini, N., Bartlett, P., El Ghaoui, L., & Jordan, M. I. (2002). Learning the kernel matrix with semi-definite programming. *International Conference on Machine Learning (ICML)*, San Mateo, CA: Morgan Kaufmann.

Ng, A., & Jordan, M. I. (2002). On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes. In T. Dietterich, S. Becker

& Z. Ghahramani (Eds.), *Advances in Neural Information Processing (NIPS) 14*, Cambridge, MA: MIT Press.

Tatikonda, S., & Jordan, M. I. (2002). Loopy belief propagation and Gibbs measures. In D. Koller & N. Friedman (Eds.), *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Xing, E. P., Jordan, M. I., & Karp, R. M. (2001). Feature selection for high-dimensional genomic microarray data. In *Machine Learning: Proceedings of the Eighteenth International Conference (ICML)*, San Mateo, CA: Morgan Kaufmann.

Ng, A., Zheng, A., & Jordan, M. I. (2001). Stable algorithms for link analysis. *Proceedings of the 24th International Conference on Research and Development in Information Retrieval (SIGIR)*. New York, NY: ACM Press.

De Freitas, J., Hoejen-Soerensen, P., Jordan, M. I., & Russell, S. (2001). Variational MCMC. In J. Breese & D. Koller (Eds.), *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Ng, A., & Jordan, M. I. (2001). Convergence rates of the Voting Gibbs classifier, with application to Bayesian feature selection. In *Machine Learning: Proceedings of the Eighteenth International Conference (ICML)*, San Mateo, CA: Morgan Kaufmann.

Ng, A., Zheng, A., & Jordan, M. I. (2001). Link analysis, eigenvectors, and stability. *International Joint Conference on Artificial Intelligence (IJCAI)*, Seattle, WA.

Deshpande, A., Garofalakis, M. N., & Jordan, M. I. (2001). Efficient stepwise selection in decomposable models. In J. Breese & D. Koller (Eds.), *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Ng, A., & Jordan, M. I. (2000). Approximate inference algorithms for two-layer Bayesian networks. In S. Solla, & T. Leen (Eds.), *Advances in Neural Information Processing Systems (NIPS) 12*. Cambridge, MA: MIT Press.

Ng, A., & Jordan, M. I. (2000). PEGASUS: A policy search method for large MDPs and POMDPs. In C. Boutilier & M. Goldszmidt (Eds.), *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Hofmann, T., Puzicha, J., & Jordan. (1999). Learning from dyadic data. In Kearns, M. S., Solla, S. A., & Cohn, D. (Eds.), *Advances in Neural Information Processing Systems (NIPS) 11*, Cambridge, MA: MIT Press.

Murphy, K. P., Weiss, Y., & Jordan, M. I. (1999). Loopy belief propagation for approximate inference: An empirical study. In K. B. Laskey & H. Prade (Eds.), *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Bishop, C. M., Lawrence, N. D., Jaakkola, T. S., & Jordan, M. I. (1998). Approximating posterior distributions in belief networks using mixtures. In Jordan, M. I., Kearns, M. J. & Solla, S. A. (Eds.), *Advances in Neural Information Processing Systems (NIPS) 10*. Cambridge, MA: MIT Press.

Meila, M., & Jordan, M. I. (1998). Estimating dependency structure as a hidden variable. In Jordan, M. I., Kearns, M. J. & Solla, S. A. (Eds.), *Advances in Neural Information Processing Systems (NIPS) 10*. Cambridge, MA: MIT Press.

Houde, J. & Jordan, M. I. (1998). Adaptation in speech motor control. In Jordan, M. I., Kearns, M. J. & Solla, S. A. (Eds.), *Advances in Neural Information Processing Systems (NIPS) 10*. Cambridge, MA: MIT Press.

Lawrence, N. D., Bishop, C. M., Jordan, M. I., & Jaakkola, T. S. (1998). Mixture representations for inference and learning in Boltzmann machines. In G. F. Cooper & S. Moral (Eds.), *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI)*. San Mateo, CA: Morgan Kaufmann.

Jordan, M. I., Ghahramani, Z., & Saul, L. K. (1997). Hidden Markov decision trees. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), *Advances in Neural Information Processing Systems (NIPS) 9*. Cambridge, MA: MIT Press.

Jaakkola, T., & Jordan, M. I. (1997). Bayesian logistic regression: a variational approach. In D. Madigan & P. Smyth (Eds.), *Proceedings of the 1997 Conference on Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL.

Saul, L. K., & Jordan, M. I. (1997). Mixed memory Markov models. In D. Madigan & P. Smyth (Eds.), *Proceedings of the 1997 Conference on Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL.

Meila, M., & Jordan, M. I. (1997). An objective function for belief net triangulation. In D. Madigan & P. Smyth (Eds.), *Proceedings of the 1997 Conference on Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL.

Saul, L. K., & Jordan, M. I. (1997). A variational principle for model-based interpolation. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), *Advances in Neural Information Processing Systems (NIPS) 9*. Cambridge, MA: MIT Press.

Meila, M., & Jordan, M. I. (1997). Optimal triangulation with continuous cost functions. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), *Advances in Neural Information Processing Systems (NIPS) 9*. Cambridge, MA: MIT Press.

Jaakkola, T., & Jordan, M. I. (1997). Recursive algorithms for approximating probabilities in graphical models. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), *Advances in Neural Processing Systems (NIPS) 9.* Cambridge, MA: MIT Press.

Saul, L. K., & Jordan, M. I. (1996). Exploiting tractable substructures in intractable networks. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems (NIPS) 8.* Cambridge, MA: MIT Press.

Ghahramani, Z., & Jordan, M. I. (1996). Factorial Hidden Markov models. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems (NIPS) 8.* Cambridge, MA: MIT Press.

Jaakkola, T., Saul, L. K., & Jordan, M. I. (1996). Fast learning by bounding likelihoods in sigmoid belief networks. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems (NIPS) 8.* Cambridge, MA: MIT Press.

Sabes, P. N., & Jordan, M. I. (1996). Reinforcement learning by probability matching. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems (NIPS) 8.* Cambridge, MA: MIT Press.

Meila, M., & Jordan, M. I. (1996). Learning fine motion by Markov mixtures of experts. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems (NIPS) 8.* Cambridge, MA: MIT Press.

Jaakkola, T., & Jordan, M. I. (1996). Computing upper and lower bounds on likelihoods in intractable networks. In E. Horvitz (Ed.), *Workshop on Uncertainty in Artificial Intelligence (UAI)*, Portland, Oregon.

Jaakkola, T., Singh, S. P., & Jordan, M. I. (1995). Reinforcement learning algorithm for partially observable Markov decision problems. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7.* Cambridge, MA: MIT Press.

Saul, L. K., & Jordan, M. I. (1995). Boltzmann chains and hidden Markov models. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7.* Cambridge, MA: MIT Press.

Fun, W., & Jordan, M. I. (1995). The moving basin: Effective action-search in adaptive control. In *Proceedings of the World Conference on Neural Networks (WCNN).* Washington, DC.

Singh, S. P., Jaakkola, T., & Jordan, M. I. (1995). Reinforcement learning with soft state aggregation. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7.* Cambridge, MA: MIT Press.

Cohn, D., Ghahramani, Z., & Jordan, M. I. (1995). Active learning with statistical models. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7*. Cambridge, MA: MIT Press.

Xu, L., Jordan, M. I., & Hinton, G. E. (1995). An alternative model for mixtures of experts. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7*. Cambridge, MA: MIT Press.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). Neural forward dynamic models in human motor control: Psychophysical evidence. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7*. Cambridge, MA: MIT Press.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1995). Computational structure of coordinate transformations: A generalization study. In G. Tesauro, D. S. Touretzky & T. K. Leen, (Eds.), *Advances in Neural Information Processing Systems (NIPS) 7*. Cambridge, MA: MIT Press.

Ghahramani, Z., & Jordan, M. I. (1994). Supervised learning from incomplete data via the EM approach. In Cowan, J., Tesauro, G., & Alspector, J., (Eds.), *Neural Information Processing Systems 6*. San Mateo, CA: Morgan Kaufmann.

Jaakkola, T., Jordan, M. I., & Singh, S. P. (1994). Convergence of stochastic iterative dynamic programming algorithms. In Cowan, J., Tesauro, G., & Alspector, J., (Eds.), *Neural Information Processing Systems 6*. San Mateo, CA: Morgan Kaufmann.

Xu, L., & Jordan, M. I. (1994). Theoretical and experimental studies on convergence properties of EM algorithm based on finite Gaussian mixtures. In *Proceedings of the 1994 International Symposium on Artificial Neural Networks*, Tainan, Taiwan, pp. 380–385.

Xu, L., Jordan, M. I., & Hinton, G. E. (1994). A modified gating network for the mixtures of experts architecture. In *Proceedings of the 1994 World Congress on Neural Networks (WCNN)*, San Diego, CA, pp. 405–410.

Jordan, M. I. (1994). A statistical approach to decision tree modeling. In M. Warmuth (Ed.), *Proceedings of the Seventh Annual ACM Conference on Computational Learning Theory (COLT)*. New York: ACM Press.

Singh, S. P., Jaakkola, T., & Jordan, M. I. (1994). Learning without state estimation in partially observable Markovian decision processes. In *Machine Learning: Proceedings of the Eleventh International Conference (ICML)*, San Mateo, CA: Morgan Kaufmann. pp. 284–292.

Bavelier, D. & Jordan, M. I. (1993). A dynamical model of priming and repetition blindness. In Hanson, S. J., Cowan, J. D., & Giles, C. L., (Eds.), *Advances in Neural Information Processing Systems (NIPS) 5.* San Mateo, CA: Morgan Kaufmann.

Xu, L., & Jordan, M. (1993). EM learning of a generalized finite mixture model for combining multiple classifiers. *Proceedings of the World Conference on Neural Networks (WCNN).* Portland, OR, pp. 431–434.

Xu, L., & Jordan, M. (1993). Unsupervised learning by an EM algorithm based on finite mixture of Gaussians. *Proceedings of the World Conference on Neural Networks (WCNN).* Portland, OR, pp. 227–230.

Jordan, M. I. & Jacobs, R. A. (1993). Supervised learning and divide-and-conquer: A statistical approach. In P. E. Utgoff, (Ed.), *Machine Learning: Proceedings of the Tenth International Workshop (ICML).* San Mateo, CA: Morgan Kaufmann.

Jordan, M. I. & Jacobs, R. A. (1992). Hierarchies of adaptive experts. In J. Moody, S. Hanson, & R. Lippmann (Eds.), *Advances in Neural Information Processing Systems (NIPS) 4.* San Mateo, CA: Morgan Kaufmann. pp. 985–993.

Hirayama, M., Vatikiotis-Bateson, E., Kawato, M., & Jordan, M. I. (1992). Forward dynamics modeling of speech motor control using physiological data. In J. Moody, S. Hanson, & R. Lippmann (Eds.), *Advances in Neural Information Processing Systems (NIPS) 4.* San Mateo, CA: Morgan Kaufmann. pp. 191–199.

Hirayama, M., Vatikiotis-Bateson, E., Kawato, M., & Jordan, M. I. (1991). Speech motor control model using electromyography. *INCN Conference on Speech Communications*, 39–46.

Jordan, M. I. & Rumelhart, D. E. (1991). Internal world models and supervised learning. In L. Birnbaum and G. Collins, (Eds.), *Machine Learning: Proceedings of the Eighth International Workshop (ICML).* San Mateo, CA: Morgan Kaufmann. pp. 70–75.

Jacobs, R. A. & Jordan, M. I. (1991). A competitive modular connectionist architecture. In D. Touretzky (Ed.), *Advances in Neural Information Processing Systems (NIPS) 3.* San Mateo, CA: Morgan Kaufmann. pp. 767–773.

Jacobs, R. A. & Jordan, M. I. (1991). A modular connectionist architecture for learning piecewise control strategies. *Proceedings of the 1991 American Control Conference (ACC)*, Boston, MA. pp. 343–351.

Mazzoni P., Andersen R. A., & Jordan M. I. (1990). AR-P learning applied to a network model of cortical area 7a. *Proceedings of the International Joint Conference On Neural Networks (IJCNN)*, San Diego, CA, pp. 373–379.

Jordan, M. I. & Jacobs, R. A. (1990). Learning to control an unstable system with forward modeling. In D. Touretzky (Ed.), *Advances in Neural Information Processing Systems (NIPS) 2.* San Mateo, CA: Morgan Kaufmann. pp. 324–331.

Jordan, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society.* Englewood Cliffs, NJ: Erlbaum. pp. 531–546. [Reprinted in IEEE Tutorials Series, New York: IEEE Publishing Services, 1990].

BOOK CHAPTERS

Jordan, M. I. (2010). Hierarchical models, nested models and completely random measures. In M.-H. Chen, D. Dey, P. Müller, D. Sun, and K. Ye (Eds.), *Frontiers of Statistical Decision Making and Bayesian Analysis: In Honor of James O. Berger*, New York: Springer.

Jordan, M. I. (2010). Bayesian nonparametric learning: Expressive priors for intelligent systems. In R. Dechter, H. Geffner, and J. Halpern (Eds.), *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, College Publications.

Teh, Y. W. & Jordan, M. I. (2010). Hierarchical Bayesian nonparametric models with applications. In N. Hjort, C. Holmes, P. Müller, & S. Walker (Eds.), *Bayesian Nonparametrics: Principles and Practice*, Cambridge, UK: Cambridge University Press.

Liang, P., Jordan, M. I. & Klein, D. (2010). Probabilistic grammars and hierarchical Dirichlet processes. In T. O'Hagan & M. West (Eds.), *The Handbook of Applied Bayesian Analysis*, Oxford, UK: Oxford University Press.

Bach, F. R., & Jordan, M. I. (2008). Spectral clustering for speech separation. In J. Keshet & S. Bengio (Eds.), *Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods*, New York: John Wiley.

Wainwright, M., & Jordan, M. I. (2005). A variational principle for graphical models. In S. Haykin, J. Principe, T. Sejnowski & J. McWhirter (Eds.), *New Directions in Statistical Signal Processing: From Systems to Brain*, Cambridge, MA: MIT Press.

Lawrence, N. D. & Jordan, M. I. (2005). Gaussian processes and the null-category noise model. In O. Chapelle, A. Zien, and B. Schölkopf (Eds.), *Semi-Supervised Learning*, Cambridge, MA: MIT Press.

Lawrence, N. D., Platt, J. C. & Jordan, M. I. (2005). Extensions of the informative vector machine. In J. Winkler and N. D. Lawrence and M. Niranjan (Eds.), *Proceedings of the Sheffield Machine Learning Workshop*, Lecture Notes in Computer Science, New York: Springer.

Lanckriet, G. R. G., Cristianini, N., Jordan, M. I., & Noble, W. S. (2003). Kernel-based integration of genomic data using semidefinite programming. In B. Schölkopf, K. Tsuda & J-P. Vert (Eds.), *Kernel Methods in Computational Biology*, Cambridge, MA: MIT Press.

Jordan, M. I., & Weiss, Y. (2002). Graphical models: Probabilistic inference. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd edition.* Cambridge, MA: MIT Press.

Jordan, M. I., & Jacobs, R. A. (2002). Learning in modular and hierarchical systems. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd edition.* Cambridge, MA: MIT Press.

Jordan, M. I. (2001). Foreword. In D. Saad & M. Opper (Eds.), *Advanced Mean Field Theory Methods—Theory and Practice.* Cambridge, MA: MIT Press.

Jaakkola, T., & Jordan, M. I. (1999). Variational methods and the QMR-DT database. In C. M. Bishop (Ed.), *Neural Networks and Machine Learning.* Berlin: Springer-Verlag.

Jordan, M. I., & Wolpert, D. M. (1999). Computational motor control. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences, 2nd edition.* Cambridge, MA: MIT Press.

Jordan, M. I., & Russell, S. (1999). Computational intelligence. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences.* Cambridge, MA: MIT Press.

Jordan, M. I. (1999). Recurrent networks. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences.* Cambridge, MA: MIT Press.

Jordan, M. I. (1999). Neural networks. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences.* Cambridge, MA: MIT Press.

Jaakkola, T. S., & Jordan, M. I. (1998). Improving the mean field approximation via the use of mixture distributions. In M. I. Jordan (Ed.), *Learning in Graphical Models.* Dordrecht: Kluwer Academic Press.

Saul, L. K., & Jordan, M. I. (1998). A mean field learning algorithm for unsupervised neural networks. In M. I. Jordan (Ed.), *Learning in Graphical Models.* Dordrecht: Kluwer Academic Press.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1998). An introduction to variational methods for graphical models. In M. I. Jordan (Ed.), *Learning in Graphical Models.* Cambridge: MIT Press.

Ghahramani, Z., & Jordan, M. I. (1997). Mixture models for learning from incomplete data. In Greiner, R., Petsche, T., & Hanson, S. J. (Eds.), *Computational Learning Theory and Natural Learning Systems.* Cambridge, MA: MIT Press.

Meila, M., & Jordan, M. I. (1997). Markov mixtures of experts. In Murray-Smith, R., & Johansen, T. A. (Eds.), *Multiple Model Approaches to Modelling and Control.* London: Taylor and Francis.

Cohn, D., Ghahramani, Z., & Jordan, M. I. (1997). Active learning with statistical models. In Murray-Smith, R., & Johansen, T. A. (Eds.), *Multiple Model Approaches to Modelling and Control.* London: Taylor and Francis.

Jordan, M. I., & Bishop, C. (1997). Neural networks. In Tucker, A. B. (Ed.), *CRC Handbook of Computer Science*, Boca Raton, FL: CRC Press.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational models of sensorimotor organization. In P. Morasso & V. Sanguineti (Eds.), *Self-Organization Computational Maps and Motor Control.* Amsterdam: North-Holland.

Jordan, M. I. (1997). Serial order: A parallel, distributed processing approach. In J. W. Donahoe & V. P. Dorsel, (Eds.). *Neural-network Models of Cognition: Biobehavioral Foundations.* Amsterdam: Elsevier Science Press.

Jordan, M. I. (1996). Computational aspects of motor control and motor learning. In H. Heuer & S. Keele (Eds.), *Handbook of Perception and Action: Motor Skills.* New York: Academic Press.

Jordan, M. I., & Jacobs, R. A. (1995). Learning in modular and hierarchical systems. In M. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks.* Cambridge, MA: MIT Press.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995). Goal-based speech motor control: A theoretical framework and some preliminary data. In D. A. Robin, K. M. Yorkston, & D. R. Beukelman (Eds.), *Disorders of Motor Speech: Assessment, Treatment, and Clinical Characterization*, Baltimore, MD: Brookes Publishing Co.

Jordan, M. I. (1994). Computational motor control. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences.* Cambridge, MA: MIT Press.

Barto, A. G. & Jordan, M. I. (1992). Gradient following without backpropagation in layered networks. In S. M. Kosslyn and R. A. Andersen (Eds.), *Readings in Cognitive Neuroscience.* Cambridge, MA: MIT Press. [Originally appeared in *Proceedings of the IEEE First Annual International Conference on Neural Networks.* New York: IEEE Publishing Services, 1987].

Jordan, M. I. (1992). Supervised learning and excess degrees of freedom. In P. Mehra, & B. Wah, (Eds.), *Artificial Neural Networks: Concepts and Theory.* Los Alamitos, CA: IEEE Computer Society Press.

White, D. A. & Jordan, M. I. (1992). Optimal control: A foundation for intelligent control. In D. A. White, & D. A. Sofge (Eds.), *Handbook of Intelligent Control.* Amsterdam: Van Nostrand.

Jordan, M. I. (1992). Constraints on underspecified target trajectories. In P. Dario, G. Sandini, & P. Aebischer, (Eds.), *Robots and Biological Systems: Toward a New Bionics.* Heidelberg: Springer-Verlag.

Jordan, M. I. & Jacobs, R. A. (1991). Modularity, supervised learning, and unsupervised learning. In S. Davis (Ed.), *Connectionism: Theory and practice.* Oxford: Oxford University Press.

Jordan, M. I. (1990). Motor learning and the degrees of freedom problem. *Attention and Performance, XIII*, 796–836.

Jordan, M. I. (1990). Learning inverse mappings with forward models. In K. S. Narendra (Ed.), *Proceedings of the Sixth Yale Workshop on Adaptive and Learning Systems.* New York: Plenum Press.

Jordan, M. I. & Rosenbaum, D. A. (1989). Action. In M. I. Posner (Ed.), *Foundations of Cognitive Science.* Cambridge, MA: MIT Press.

Jordan, M. I. (1986). An introduction to linear algebra in parallel, distributed processing. In D. E. Rumelhart and J. L. McClelland, (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Cambridge, MA: MIT Press.

BOOKS

Domeniconi, C., & Jordan, M. I. (2001). *Discorsi sulle Reti Neurali e l'Apprendimento.* Milan: Franco Angeli Editore.

Jordan, M. I., LeCun, Y. & Solla, S. A. (Eds.). (2001). *Advances in Neural Information Processing Systems, Proceedings of the First Twelve Conferences on CD-ROM*, Cambridge MA: MIT Press.

Jordan, M. I., & Sejnowski, T. J. (Eds.). (2001). *Graphical Models: Foundations of Neural Computation.* Cambridge MA: MIT Press.

Jordan, M. I. (Ed.). (1999). *Learning in Graphical Models*, Cambridge, MA: MIT Press.

Jordan, M. I., Kearns, M. J. & Solla, S. A. (Eds.). (1998). *Advances in Neural Information Processing Systems 10*, Cambridge MA: MIT Press.

Mozer, M. C., Jordan, M. I., & Petsche, T. (Eds.). (1997). *Advances in Neural Information Processing Systems 9*, Cambridge MA: MIT Press.

EXHIBIT 3 A

Exhibit 3-A

Claim Chart of Dunja Mladenic, "Personal WebWatcher: design and implementation", *Technical Report IJS-DP-7472*, Department of Intelligent Systems, J. Sefan Institute, Slovenia (1996) ("MLADENIC").

as prior art to

Asserted Claims of U.S. Patent No. 6,9981,040 ("'040 Patent")
and
Asserted Claims of U.S. Patent No. 7,685,276 ("'276 Patent")

'040 Patent	MLADENIC
Claim 1	
A computer-implemented method for providing automatic, personalized information services to a user u, the method comprising:	Personal WebWatcher is based on the WebWatcher described in *WebWatcher: A Tour Guide for the World Wide Web*, by Joachims, Freitag, and Mitchell (1997); *WebWatcher: Machine Learning and Hypertext*, by Joachims, Freitag, Mitchell and Armstrong (1995); and *Web Watcher: A Learning Apprentice for the World Wide Web*, by Armstrong, Freitag, Joachims, and Mitchell (1995) ("Armstrong"). "[u]nlike WebWatcher, Personal WebWatcher (PWW) is structured to specialize for a particular user, modeling his/her interests." (Mladenic at 3). At a high level, Personal WebWatcher uses a corpus of documents and interactions that correspond to a specific user, rather than a set of documents and interactions that correspond to a web site, as WebWatcher does. As one would expect, using an individual user's documents for training data results in a system that is customized to that individual user. Personal WebWatcher then provides automatic, personalized services to the user by highlighting links that the user may be interested in. As seen in Figure 3 below, PWW places "eyeballs" next to selected links:

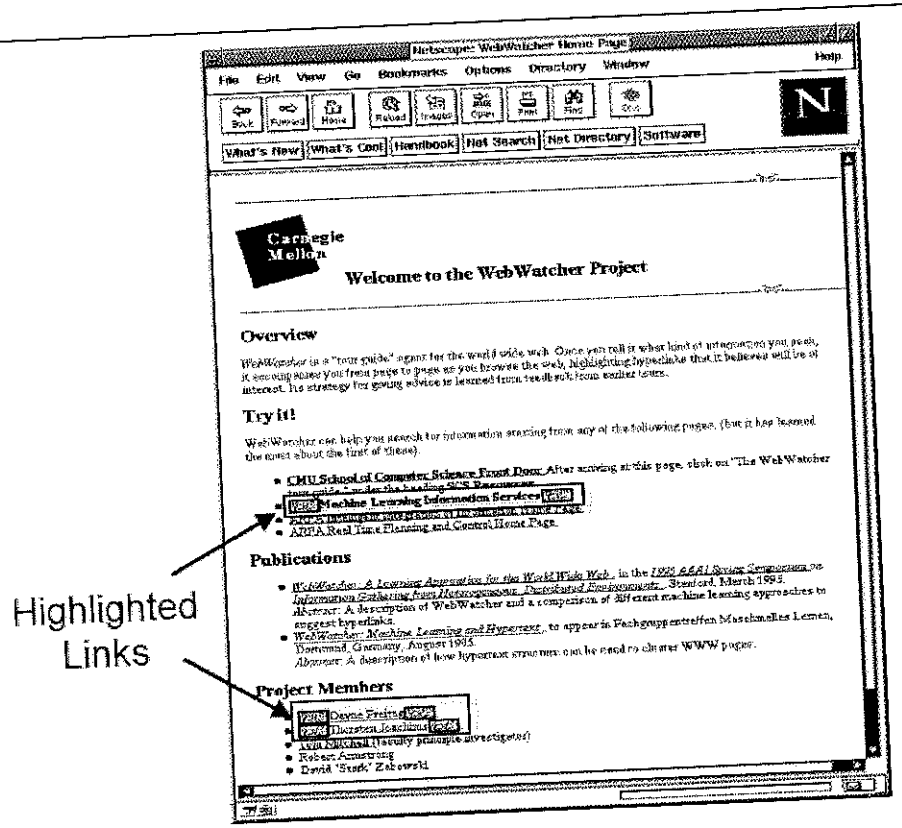
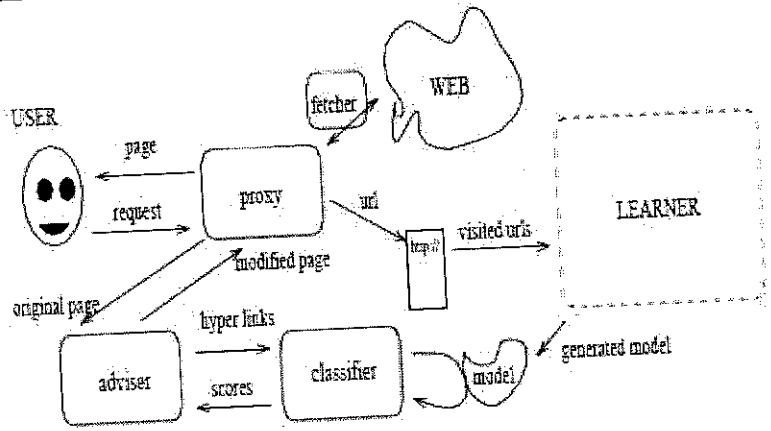
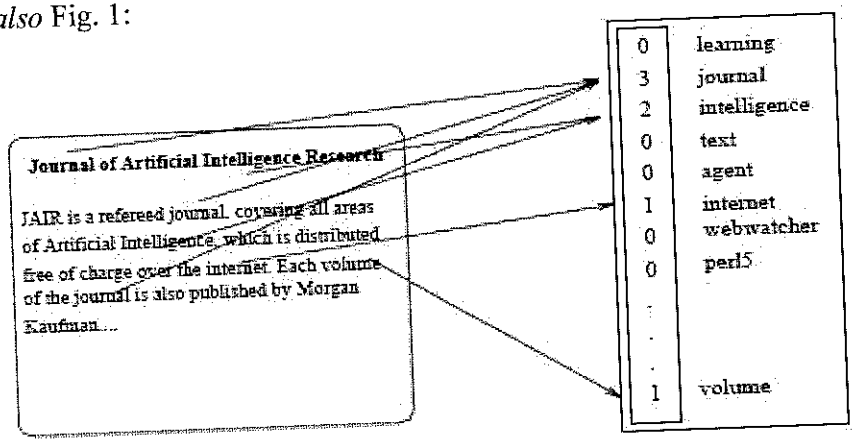


Figure 3: Example of HTML page presented to the user by PWW.

See also Mladenic at 2 ("Personal WebWatcher is a system that observes users of the WWW and suggests pages they might be interested in.... Web Watcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough."), Fig. 2:

	
a) transparently monitoring user interactions with data while the user is engaged in normal use of a computer;	MLADENIC, p. 2: "Personal WebWatcher is a system that observes users of the WWW and suggests pages they might be interested in. It learns user interests from the pages requested by the user." p. 3: "Unlike WebWatcher, Personal WebWatcher(PWW) is structured to specialize for a particular user, modeling her/his interests. It "watches over the user shoulder" the similar way WebWatcher does, but it avoids involving the user in its learning process." *See also*: **CULLISS** at 3:46-56; 5:18-21; 7:14-20. **REFUAH** at Abstract; 3:3-11; 5:34-50; 19:20-22. **SCHUETZE** at 5:36-40; 11:12-14; 18:11-17; 28:65 – 29:6. **WASFI** at Abstract, 60. Autonomy Press Release, at 2

	Autonomy Technology Whitepaper, AUT00069-70 Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 3.
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user;	MLADENIC, p. 3: "It solely records the addresses of pages requested by the user and highlights hyperlinks that it believes will be of interest." p. 8: "Both versions fetch visited documents and documents one step behind the hyperlinks of visited documents and store them as positive or negative examples of user interests, depending whether the user visited the document or not." p. 8: "Hyperlinks whose documents were visited by the user are considered to be positive examples, and all other to be negative examples of the user interests. The idea is that all hyperlinks were presented to the user and the user chose to visit some of them that meet her/his interests." *See also*: **CULLISS** at 3:13-35; 5:36-48; 7:14-42. **REFUAH** at Abstract; 5:34-50; 6:5-15; 8:31-39; 20:31-37. **SCHUETZE** at 10:14-18; 10:32-39; 11:12-17; 28:65 – 29:6; 34:34-37; *See generally* 17:47 – 18:27. **WASFI** at 58, 60, 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 **MONTEBELLO** at 3-4.

c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	Personal Web Watcher treats a web page as a "bag of words," meaning that "all words from the document are taken and no ordering of words or any structure of text is used." (Mladenic at 3-4.) *See also* Fig. 1:  Figure 1: Bag-of-words representation using frequency vector. MLADENIC, p. 7: PWW uses a learner module, which uses "a Naïve (Simple) Bayesian classifier on frequency vectors to generate a model of user interests, that is used for advising hyperlinks." p. 8: "[Personal WebWatcher] fetch[es] visited documents and documents one step behind the hyperlinks of visited documents and store[s] them as positive or negative examples of user interests, depending on whether the user visited the document or not." p. 9: "LEARNER transforms documents into examples in two phases : (1) (docs2exs and docs2addexs in Figure4) parsing each document, assigning an index to each word and representing it in three files as a line of word indices containing: all words, only headline words, only underlined words. (2) (exs2vec in Figure4) calculating score (e.g. information gain) for each word, selecting some top words and represent documents as bag-of-words keeping frequency for each of the top words." p. 10: "The model of user interests is designed to predict if some document is positive or negative example of user interests."

	pp. 10-11: "The model of user interests is generated "off-line," usually during the night and thus its generation is not so critical in time as its usage for prediction. One of the simplest idea for learning is to use hyperlinks that occurred on the documents presented to the user as training examples and learn to predict if a new hyperlink is positive or negative example of the user interests....But during the learning phase we can afford using more time than when adding advice, so why not retrieving document behind hyperlinks, instead of using the extended hyperlink representation? In that case, we can learn the model of user interests directly from documents whose interestingness we are trying to predict." *See also*: **CULLISS** at 3:57-65; 4:54 – 5:10. **REFUAH** at 2:9-35; 6:49-64; 8:30-58; *See generally* 14:21 – 15:45. **SCHUETZE** at 27:44-64. *See also* 27:65 – 28:14. **WASFI** at 58, 61, 63. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00069-70 Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3, 4.
d) analyzing a document d to identify properties of the document;	Personal Web Watcher treats a web page as a "bag of words," meaning that "all words from the document are taken and no ordering of words or any structure of text is used." (Mladenic at 3-4.) *See also* Fig. 1:

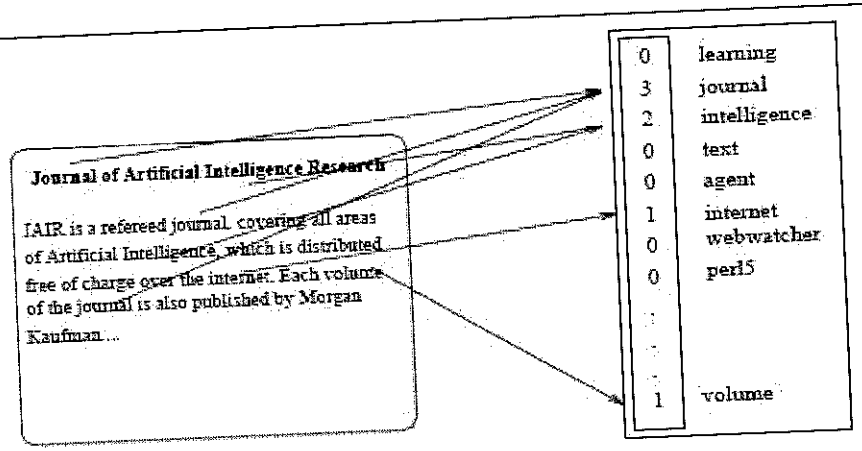


Figure 1: Bag-of-words representation using frequency vector.

MLADENIC, p. 4: "We decided to use the bag-of-words representation using frequency of word and observe success of given advice (whether user selected the advised hyperlink). In case of poor system performance, some additional information from HTML-structure could be added, for example, frequency of word in headlines of a given document."

p. 4: "Some later versions of the WebWatcher system change slightly the way of constructing text for learning, e.g., adding words in the document retrieved behind hyperlink."

p. 10: PWW is "actually predicting interestingness of document based on the hyperlink pointing to it, and not document itself" because "retrieving documents behind the requested hyperlinks is usually time consuming."

p. 12: "Documents are currently represented using the bag-of-words approach (see Section 3.1) and feature selection is performed using mutual information approach (see Section 3.2)."

See also:

CULLISS at 2:43-46; 5:36-48; 10:47-52.

	REFUAH at 7:53 – 8:6; 20:31-37; 21:6-30. **SCHUETZE** at 3:40-44; 5:59-63; 10:20-31; Fig.1. *See generally* 11:42 – 15:19. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068-69, AUT00071 Autonomy Agentware User Guide, at AUT00002, AUT00117 **MONTEBELLO** at 3, 4.
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	MLADENIC, p. 5: "Since we are more interested in positive class (interested documents) and we want to have words that are frequent, it might be better to include in the weighting formula the probability of a word occurring in the positive class or frequency of the word, (a formula) where w is a selected word, c is the positive class and TF(w) is the frequency of the word w." p. 7: "We decided to test different learning algorithms on PWW data (see Section 5), since it is not clear which algorithm is the most appropriate. The current version of PWW uses a Naive (Simple) Bayesian classifier on frequency vectors to generate a model of user interests, that is used for advising hyperlinks." p. 7: "The current version of PWW uses a Naïve(Simple) Bayesian classifier on frequency vectors to generate a model of user interests, that is used for advising hyperlinks." p. 8: "Both versions [of the document analyzer used in PWW] fetch visited documents and documents one step behind the hyperlinks of visited documents and store them as positive ***or negative*** examples of user interests, depending whether the user visited the document ***or not***." p. 10: "The model of user interests is designed to predict if some document is positive or negative

example of user interests."

p. 12, Table 2:

UserId and data source	probability of interestingness	number of examples	data entropy
usr150101			
Doc	0.094	1 333	0.449
HL	0.104	2 528	0.480
usr150202			
Doc	0.107	3 415	0.492
HL	0.053	4 798	0.301
usr150211			
Doc	0.089	2 038	0.436
HL	0.044	2 221	0.259
usr150502			
Doc	0.100	1 272	0.468
HL	0.100	2 198	0.468

See also:

CULLISS at 2:43-51; 5:36-62.

REFUAH at 3:56 – 4:4; 17:32 – 18:4.

SCHUETZE at 18:62 – 19:10.

WASFI at 60, 61.

Autonomy Press Release, at 1-2

	Autonomy Technology Whitepaper, at AUT00069, AUT00071 Autonomy Agentware User Guide, at AUT00119 **MONTEBELLO** at 4.
f) using the estimated probability to provide automatic, personalized information services to the user.	MLADENIC, p. 2: "Web Watcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough." pp. 7-8: "A limited number of hyperlinks that are scored above some threshold are recommended to the user, indicating their scores with graphical symbols placed around each advised hyperlinks. For example, in Figure 3 three hyperlinks are suggested by PWW: "Machine Learning Information Services" and two project members (Dayne Freitag, Thorsten Joachims)." p. 12: "Our experiments are performed on data collected for four users participating in the HOMENET project [14] with the data characteristics given in Table 2." *See also*: **CULLISS** at 2:39-51. **REFUAH** at Abstract, 2:63 – 3:11; 3:47-55; 17:21-43; 18:56-65; 23:11-28. **SCHUETZE** at 1:29-33; 7:54-60. **WASFI** at 61. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00068 Autonomy Agentware User Guide, at AUT00002, AUT00004, AUT00116-117

	MONTEBELLO at 4.
Claim 11	
11. The method of claim **1** further comprising estimating a posterior probability P(u|d,q) that the document d is of interest to the user u, given a query q submitted by the user.	MLADENIC, p. 2: "WebWatcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough. The idea is that the user provides a few keywords describing a search goal and WebWatcher highlights related hyperlinks on the current page and/or adds new hyperlinks to the current page." *See also*: **CULLISS** at 2:39-51, 9:41-49, 10:1-7, 10:47-52. **REFUAH** at 17:21-43. **SCHUETZE** at 21:57 – 22:16, 22:31-48, 30:58 – 31:13.
Claim 22	
22. The method of claim **1** wherein the monitored user interactions include a sequence of interaction times	*See also*: **MLADENIC** at 2. **REFUAH** at 5:57-58. **WASFI** at Abstract, 57. Autonomy Press Release, at 1 Autonomy Agentware User Guide, at AUT00124
Claim 32	
32. A program storage	*See* citations for claim 1[preamble].

device accessible by a central computer, tangibly embodying a program of instructions executable by the central computer to perform method steps for providing automatic, personalized information services to a user u, the method steps comprising:	
a) transparently monitoring user interactions with data while the user is engaged in normal use of a client computer in communication with the central computer;	*See* citations for claim 1[a].
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user;	*See* citations for claim 1[b].
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	*See* citations for claim 1[c].

d) analyzing a document d to identify properties of the document;	*See* citations for claim 1[d].
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	*See* citations for claim 1[e].
f) using the estimated probability to provide automatic, personalized information services to the user	*See* citations for claim 1[f].
Claim 34	
34. The program storage device of claim **32** wherein analyzing the document d provides for the analysis of documents having multiple distinct media types.	MLADENIC, p. 2: "WebWatcher can be describes as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough." p. 10: "The model of user interests is designed to predict if some document is positive or negative example of user interests. It is used to advice hyperlinks on the HTML document requested by the user." *See also*: **CULLISS** at 1:22-28, 2:19-24.

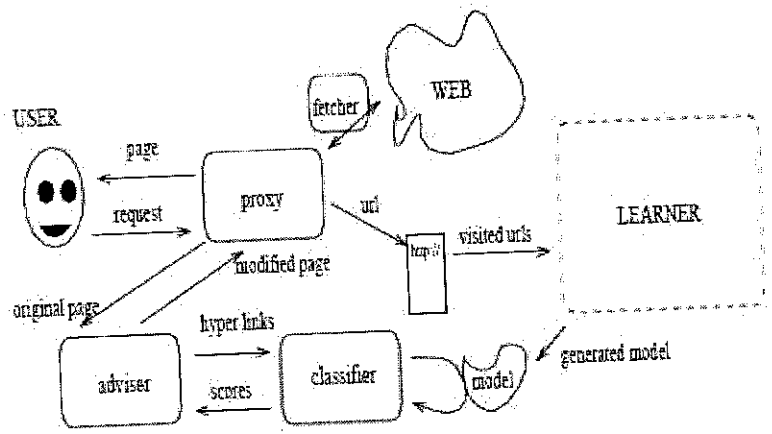
	REFUAH at 1:63 – 2:2. **SCHUETZE** at 4:12-35, 5:48-58. *See generally* "Summary of the Invention", 5:43 – 8:32. **WASFI** at 57, 58.
'276 Patent	MLADENIC
Claim 1	
A computer-implemented method for providing personalized information services to a user, the method comprising:	*See* citations for '040 Patent, claim 1[preamble].
a) transparently monitoring user interactions with data while the user is engaged in normal use of a browser program running on the computer;	*See* citations for '040 Patent, claim 1[a]. MLADENIC, p. 2: "Personal WebWatcher is a system that observes users of the WWW and suggests pages they might be interested in. It learns from the pages requested by the user. The learned model of user interests is then used to suggest hyperlinks on new HTML-pages requested by and presented to the user via Web browser that enables connection to [a] "proxy," e.g. Netscape." p. 7: "Personal WebWatcher consists of two main parts: a proxy server that interacts with the user via Web browser and a learner that provides the user-model to the server." p. 7: "Proxy waits in an infinite loop for a page request from the browser." *See also*: **REFUAH** at 5:57-58.

	WASFI at Abstract, 57.
b) analyzing the monitored data to determine documents of interest to the user;	*See* citations for '040 Patent, claim 1[b].
c) estimating parameters of a user-specific learning machine based at least in part on the documents of interest to the user;	*See* citations for '040 Patent, claim 1[c].
d) receiving a search query from the user;	MLADENIC, p. 1: "For example, Armstring et al. [2] developed WebWatcher, a system that assists user[]s in locating information on the World Wide Web [by] taking keywords from the user, suggesting hyperlinks and receiving evaluation." p. 2: "WebWatcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough. The idea is that the user provides a few keywords describing a search goal and WebWatcher highlights related hyperlinks on the current page and/or adds new hyperlinks to the current page." *See also*: **CULLISS** at 2:39-51. **REFUAH** at 1:63 – 2:2; 3:12-24. **SCHUETZE** at 21:57 – 22:16; 22:31-48. Autonomy Agentware User Guide, at AUT00002 **MONTEBELLO** at 3.

e) retrieving a plurality of documents based on the search query;	MLADENIC, p. 1: "For example, Armstring et al. [2] developed WebWatcher, a system that assists user[]s in locating information on the World Wide Web [by] taking keywords from the user, suggesting hyperlinks and receiving evaluation." p. 2: "WebWatcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough. The idea is that the user provides a few keywords describing a search goal and WebWatcher highlights related hyperlinks on the current page and/or adds new hyperlinks to the current page." *See also*: **CULLISS** at 2:39-51. **REFUAH**, 1:63 – 2:2; 3:12-24. **SCHUETZE**, 21:57 – 22:16; 22:31-48. **MONTEBELLO** at 3.
f) for each retrieved document of said plurality of retrieved documents: identifying properties of the retrieved document, and applying the identified properties of the retrieved document to the user-specific learning machine to estimate a probability that the retrieved document is of interest to the user; and	*See* citations for '040 Patent, claim 1[d, e].

g) using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user.	*See* citations for '040 Patent, claim 1[f]. MLADENIC, p. 2: "The idea is that the user provides a few keywords describing a search goal and WebWatcher highlights related hyperlinks on the current page and/or adds new hyperlinks to the current page."
Claim 3	
3. The method of claim **1**, wherein transparently monitoring user interactions with data comprises monitoring user interactions with data during multiple different modes of user interaction with network data.	MLADENIC, p. 3: "Unlike WebWatcher, Personal WebWatcher(PWW) is structured to specialize for a particular user, modeling her/his interests. It "watches over the user shoulder" the similar way WebWatcher does, but it avoids involving the user in its learning process." *See also*: **CULLISS** at 2:43-46, 3:29-35, 3:46-56. **REFUAH** at 5:34-50. **SCHUETZE** at 18:11-17. **WASFI** at 58, 61. Autonomy Press Release, at 1
Claim 5	
5. The method of claim **1**, further comprising analyzing the monitored data to determine documents not of interest to the user, and wherein estimating parameters of a	MLADENIC, p. 8: "Hyperlinks whose documents were visited by the user are considered to be positive examples, and all the other to be negative examples of the user interests. The idea is that all hyperlinks were presented to the user and the user chose to visit some of them that meet her/his interests." *See generally* pp. 10-11. *See also*:

user-specific learning machine further comprises estimating parameters of a user-specific learning machine based at least in part on the documents not of interest to the user.	**CULLISS** at 3:57-65, 4:61-64, 5:4-10. **REFUAH** at 22:6-14. **SCHUETZE** at 17:47-67. **WASFI** at 58. **TAN** at 4.2.1, 4.3.2. Autonomy Press Release, at 1 Autonomy Technology Whitepaper, at AUT00070 Autonomy Agentware User Guide, at AUT00123
Claim 6	
6. The method of claim **1**, wherein monitoring user interactions with data for a document comprises monitoring at least one type of data selected from the group consisting of information about the document, whether the user viewed the document, information about the user's interaction with the document, context information, the user's degree of interest in the	MLADENIC, p. 8: "Hyperlinks whose documents were visited by the user are considered to be positive examples, and all the other to be negative examples of the user interests. The idea is that all hyperlinks were presented to the user and the user chose to visit some of them that meet her/his interests." *See also*: **CULLISS** at 2:43-46, 3:27-35. **REFUAH** at 5:34-50, 14:54-59. **SCHUETZE** at 5:36-40, 11:12-14, 18:11-17. **WASFI** at Abstract, 60, 61.

document, time spent by the user viewing the document, whether the user followed at least one link contained in the document, and a number of links in the document followed by the user.	
Claim 7	
7. The method of claim **1**, wherein said plurality of retrieved documents correspond to a respective plurality of products.	MLADENIC, p. 2: "Personal WebWatcher is a system that observes users of the WWW and suggests pages they might be interested in." p. 2: "Web Watcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough." p. 8, Figure 2: 

	See also: **CULLISS** at 9:55- 10:13. *See generally* 9:55 – 11:33. **REFUAH** at 1:63 – 2:2, 3:56 – 4:4, 7:24-32, 18:35-39, 18:40-55. **SCHUETZE** at 35:66 – 36:8.
Claim 21	
21. The method of claim **1**, wherein using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user comprises presenting to the user at least said portion of the retrieved documents based on the estimated probability that the retrieved document is of interest to the user and the relevance of the retrieved document to the search query.	MLADENIC, p. 2: "WebWatcher can be described as an agent that assists users in locating information on the WWW. It learns by observing a user on her/his way through the WWW and suggests interesting hyperlinks whenever it is confident enough. The idea is that the user provides a few keywords describing a search goal and WebWatcher highlights related hyperlinks on the current page and/or adds new hyperlinks to the current page." *See* citations for claim 1[g].
Claim 22	
22. The method of claim **1**, wherein identifying properties of the retrieved	MLADENIC, p. 2: "[Personal WebWatcher] can also suggest pages related to the current page using information stored in the structure of hypertext without considering the text itself."

document comprises identifying properties selected from the properties consisting of a topic associated with the retrieved document, at least one product feature extracted from the retrieved document, an author of the retrieved document, an age of the retrieved document, a list of documents linked to the retrieved document, a number of users who have accessed the retrieved document, and a number of users who have saved the retrieved document in a favorite document list.	p. 3: "It solely records the addresses of pages requested by the user and highlights hyperlinks that it believes will be of interest." p. 4: "We decided to use the bag-of-words representation using frequency of word and observe success of given advice (whether user selected the advised hyperlink). In case of poor system performance, some additional information from HTML-structure could be added, for example, frequency of word in headlines of a given document." p. 10: "Since the prediction should be performed while the user is waiting for an HTML-document, we are actually predicting interestingness of [sic] document based on the hyperlink pointing to it, and not the document itself... [b]ut during the learning phase we can afford using more time than when adding advice, so why not retrieving [sic] documents behind hyperlinks, instead of using the extended hyperlink representation?" p. 12: "Documents are currently represented using the bag-of-words approach (see Section 3.1) and feature selection is performed using mutual information approach (see Section 3.2)." *See generally* 3-6. *See also*: **CULLISS** at 2:26-37. **REFUAH** at 7:53 – 8:6, 9:50-59, 20:19-30, 21:6-30. *See generally* 20:19- 21:36. **SCHUETZE** at 6:58 – 7:15, 10:40-56, Fig. 3. *See generally* 17:47 – 18:27. **WASFI** at 61.

EXHIBIT 3 B

Exhibit 3-B

Claim Chart of Autonomy's Agentware

as prior art to

Asserted Claim of U.S. Patent No. 6,9981,040 ("'040 Patent")
and
Asserted Claims of U.S. Patent No. 7,685,276 ("'276 Patent")

'040 Patent	AGENTWARE
Claim 1	
A computer-implemented method for providing automatic, personalized information services to a user u, the method comprising:	Autonomy's Agentware is a piece of software that "enables automatic personalization of Web content as well as flexible management of unstructured information." (Autonomy White Paper at 11.) Autonomy's Agentware employs "agents," or small programs that "locate information based on concepts and context, thereby selecting the most relevant information according to the individual's preferences." (Autonomy PR at 1). These personalized agents can learn about a user's interests in a variety of ways, including by "simply observing a user's actions." Accordingly, agents "do not require [the user] to fill out lengthy questionnaires or rate his likes and dislikes." Agents can "analyze a text and identify the key concepts within the document because [they] understand[] how the frequency and relationships of terms correlate with meaning." (Autonomy WP at 1). Once an agent has been trained, "it can compare [a pattern] to the documents it finds on the Internet. Autonomy does not use keyword searches, but actually identifies the concepts involved in the text and compares them. It then assigns a relevance to the document depending on how closely it matches the patterns established by the training." (Autonomy UG at 4). "By maintaining a set of agents that correspond to a user's interests, on-line publishers and service providers can then offer a range of personalized services." (Autonomy WP at 2.) *See* Autonomy Press Release, at 1 ("Agentware 1.1, the first-ever product that delivers personalized information from the Internet by learning the preferences and needs of the user")

	Autonomy Press Release, at 2 ("They can 'live' on the user's PC (client) or on a publisher's Website (server) . . .") Autonomy Press Release, at 2 ("Agentware 1.1 requires a 486DX or microprocessor running Windows 95 or 3.11 . . ." Autonomy Technology Whitepaper, at AUT00068 ("Autonomy's technology offers a breakthrough in managing unstructured digital information, including word process and HTM-based files, email messages, and electronic news feeds.") Autonomy Technology Whitepaper, at AUT00078 ("Autonomy's software enables automatic personalization of Web content as well as flexible management of unstructured information.") Autonomy Agentware User Guide, at AUT00002 ("Autonomy Agentware is the first mass market application of Artificial Intelligence technology for the Internet") Autonomy Agentware User Guide, at AUT00004 ("Autonomy browses the web like a person, reacting to its environment, and bringing back only the relevant pages")
a) transparently monitoring user interactions with data while the user is engaged in normal use of a computer;	Autonomy Press Release, at 1 (Agentware "automatically learns user preferences to deliver personalized information on demand.") Autonomy Press Release, at 2 ("The Agent keeps learning about your user's interests and will adapt accordingly as your interests change") Autonomy Technology Whitepaper, AUT00069 ("As users' needs and interests change, the agents dynamically track those shifts without the need for manual intervention") Autonomy Technology Whitepaper, AUT00070 ("Many of Autonomy's products rely on user interest or employee expertise profiles created implicitly, simply by observing user's actions") Autonomy Technology Whitepaper, AUT00070 ("Autonomy's software applies the pattern-

	matching technology to extract key ideas from the articles a user reads online." Autonomy Technology Whitepaper, AUT00070 ("Because these services are based on a user's actual interests, they do not require him to fill out lengthy questionnaires or rate his likes and dislikes.") Autonomy Technology Whitepaper, AUT00070 ("As an individual reads additional articles online, publishes material on the corporate intranet or submits documents to the knowledge management system, the Autonomy system updates his agents by recalculating interest levels in the different ideas. . . . In this way, the Autonomy system keeps pace with an individual's changing interests. This is in contrast to an explicit preference setting, which users must remember to adjust as their interests evolve") Autonomy Agentware User Guide, at AUT00119 ("The Agents in Autonomy are intelligent because they learn all the time that you use them. Autonomy learns by retraining using samples of good documents. . . . Autonomy bases its assessment of how common phrases are on an initial knowledge of English, and what it learns from you as you use it. For example, if every document in your library is about aeroplanes, Autonomy will start to conclude that the phrase 'aeroplane' is less important than say 'safety systems' in your training text. Autonomy will adapt to the style of English that you use to train it, and the style used in documents that you decide to keep") *See also* **SCHUETZE** at 5:36-40, 11:12-14, 18:11-17, 28:65 – 29:6 **CULLISS** at 3:46-56; 5:18-21; 7:14-20. **MLADENIC**, at 3. **REFUAH** at Abstract; 3:3-11; 5:34-50; 19:20-22. **WASFI** at Abstract, 60.

	MONTEBELLO at 3.
b) updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user;	Autonomy Press Release, at 1 ("Simply name your Agent and train it by typing in a sentence or two in plain English. . . . The Agent keeps learning about your user's interests and will adapt accordingly as your interests change.") Autonomy Technology Whitepaper, at AUT00069 ("Autonomy's technology also can create highly specific user profiles by analyzing and abstracting *Concept Agents* from every article a user reads. By maintaining a set of agents that correspond to a user's interests, on-line publishers and service providers can then offer a range of personalized services.") Autonomy Technology Whitepaper, at AUT000069 ("Concept Agents are created by having the DRE (Dynamic Reasoning Engine) process a set of text. This text could be a phrase that is typed specifically to train the agent, or could consist of an existing document or set of documents") Autonomy Technology Whitepaper, at AUT00070 ("Online publishers and Intranet developers can maintain sets of agents representing an individual's interests to offer a number of personalized services.") Autonomy Technology Whitepaper, at AUT00070 ("As a user is served documents, Autonomy's software automatically creates a set of *Concept Agents*.") Autonomy Technology Whitepaper, at AUT00070 ("Many of Autonomy's products rely on user interest or employee expertise profiles created implicitly, simply by observing a user's actions.") *See also* **SCHUETZE** at 10:14-18, 10:32-39, 11:12-17, 17:47 – 18:27, 28:65 – 29:6, 34:34-37 **CULLISS** at 3:13-35; 5:36-48; 7:14-42. **MLADENIC** at 3, 8.

	REFUAH at Abstract; 5:34-50; 6:5-15; 8:31-39; 20:31-37. **WASFI** at 58, 60, 61. **MONTEBELLO** at 3-4.
c) estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files;	Autonomy Press Release, at 1 ("Autonomy's Agentware use Agents which locate information based on concepts and context, thereby selecting the most relevant information according to the individual's preferences") Autonomy Press Release, at 1 ("Simply name your Agent and train it by typing in a sentence or two in plain English. . . . The Agent can now be sent out to scour the Internet, bringing back a selection of relevant information. The Agent keeps learning about your user's interests and will adapt accordingly as your interests change.") Autonomy Technology Whitepaper, at AUT00069 ("Autonomy's technology can create highly specific user profiles by analyzing and abstracting *Concept Agents* from every article a user reads. By maintaining a set of agents that correspond to a user's interests, online publishers and service providers can then offer a range of personalized services. As users' needs and interests change, the agents dynamically track those shifts without the need for manual intervention.") Autonomy Technology Whitepaper, at AUT00070 ("typical applications maintain a set of multiple *Concept Agents* for each user.") Autonomy Technology Whitepaper, at AUT00070 ("Autonomy's software applies the pattern-matching technology to extract key ideas from the articles a user reads online. As a user is served documents, Autonomy's software automatically creates a set of *Concept Agents*. By weighting the frequency with which certain topics occur, the server can then encode a set of interests into Autonomy *Concept Agents*.") \ Autonomy Technology Whitepaper, AUT00070 ("As an individual reads additional articles online, publishes material on the corporate intranet or submits documents to the knowledge management

	system, the Autonomy system updates his agents by recalculating interest levels in the different ideas. . . . In this way, the Autonomy system keeps pace with an individual's changing interests.") Autonomy Agentware User Guide, at AUT00002 ("Autonomy learns by changing the search criteria of an Agent to reflect your choices of good documents") *See also* **SCHUETZE** at 27:44-64, 27:65-28:14 **CULLISS** at 3:57-65; 4:54 – 5:10. **MLADENIC**, at 9, 10. **REFUAH** at 2:9-35; 6:49-64; 8:30-58; *See generally* 14:21 – 15:45. **WASFI** at 58, 61, 63. **MONTEBELLO** at 3, 4.
d) analyzing a document d to identify properties of the document;	Autonomy Press Release, at 1 ("a user interested in business information could create an Agent, name it 'Traveler', and send it out to look for travel arrangement information based on their own designated personal preferences . . .") Autonomy Technology Whitepaper, at AUT00068 ("the technology can analyze a text and identify the key concepts within the document because it understands how the frequency and relationships of terms correlate with meaning. . . . Autonomy . . . extract[s] a document's digital essence and determine[s] the characteristics that give the text meaning. Once Autonomy's technology has identified and encoded the unique 'signature' of the key concepts, Concept Agents are created to seek out similar ideas in websites, news fees, email archives and other documents." Autonomy Technology Whitepaper, at AUT00069 ("Autonomy's APCM technology can analyze a piece of text and determine the main ideas through its understanding of how the frequency and

	relationships of terms correlates with meanings") Autonomy Technology Whitepaper, at AUT00069 ("the idea, or concept is represented by the pattern of terms and contextual relationships that are most commonly recognized as important in documents that address the concept") Autonomy Technology Whitepaper, at AUT00069 ("*Concept Agents* can be used to automatically sort documents into pre-defined categories. . . .") Autonomy Technology Whitepaper, at AUT00071 (describing sorting of documents into pre-defined categories) Autonomy Technology Whitepaper, at AUT00071 (describing the use of Shannon's Information Theory "which enables Agentware to determine the most important (or informative) concepts within a document.") Autonomy Agentware User Guide, at AUT00002 ("Once Autonomy has identified this 'pattern' in the training text it can compare it to the documents it finds on the Internet. Autonomy does not use keyword searches, but actually identifies the concepts involved in the text and compares them. It then assigns a relevance to the document depending on how closely it matches the pattern established by the training.") Autonomy Agentware User Guide, at AUT00117 ("In the Library Hit List, the documents are ranked in order of relevance to your query, as indicated by the length of the 'bone' shown beside them") *See also* **SCHUETZE** at 3:40-44, 5:59-63, 5:59-63, 10:20-31, 11:42 – 15:19 **CULLISS** at 2:43-46; 5:36-48; 10:47-52. **MLADENIC**, at 4, 12.

	REFUAH at 7:53 – 8:6; 20:31-37; 21:6-30. **WASFI** at 61. **MONTEBELLO** at 3, 4.
e) estimating a probability P(u|d) that an unseen document d is of interest to the user u, wherein the probability P(u|d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model; and	Autonomy Press Release, at 1 ("Autonomy's Agentware use Agents which locate information based on concepts and context, thereby selecting the most relevant information according to the individual's preferences") Autonomy Press Release, at 1 ("Simply name your Agent and train it by typing in a sentence or two in plain English. . . . The Agent can now be sent out to scour the Internet, bringing back a selection of relevant information. The Agent keeps learning about your user's interests and will adapt accordingly as your interests change.") Autonomy Press Release, at 1 ("Autonomy Agents are persistent. They can be used again and again . . .") Autonomy Press Release, at 2 ("Agentware 1.1, is based on the Dynamic Reasoning Engine (DRE) . . .") Autonomy Press Release, at 1 ("a user interested in business information could create an Agent, name it 'Traveler', and send it out to look for travel arrangement information based on their own designated personal preferences . . .") Autonomy Technology Whitepaper, at AUT00069 ("Concept agents are created by having the DRE (Dynamic Reasoning Engine) process a set of text. . . . The agent then reviews the source text and calculates the patterns of its most important concepts . . ." Autonomy Technology Whitepaper, at AUT00071 (describing the use of Bayesian analysis as "centered on calculating the probabilistic relationship between multiple variables and determining the extent to which one variable impacts another.")

	Autonomy Agentware User Guide, at AUT00119 ("Autonomy's neural network is able to identify the important words and phrases in your training text, and then apply a weighting based upon how common the phrase is according to its knowledge of language") *See also* **SCHUETZE** at 18:62 – 19:10 **CULLISS** at 2:43-51; 5:36-62. **MLADENIC** at 5, 7, 10, 12, Table 2. **REFUAH** at 3:56 – 4:4; 17:32 – 18:4. **WASFI** at 60, 61. **MONTEBELLO** at 4.
f) using the estimated probability to provide automatic, personalized information services to the user.	Autonomy Press Release, at 1 ("Agentware 1.1, the first-ever product that delivers personalized information from the Internet by learning the preferences and needs of the user") Autonomy Press Release, at 1 ("Autonomy's Agentware use Agents which locate information based on concepts and context, thereby selecting the most relevant information according to the individual's preferences") Autonomy Press Release, at 1 ("Simply name your Agent and train it by typing in a sentence or two in plain English. . . . The Agent can now be sent out to scour the Internet, bringing back a selection of relevant information. The Agent keeps learning about your user's interests and will adapt accordingly as your interests change.") Autonomy Press Release, at 1 ("Autonomy Agents are persistent. They can be used again and again . . .")