

David M. Pennock

Yahoo! Research Labs, 74 N. Pasadena Ave., 3rd Floor, Pasadena, CA 91103 USA

+1-626-229-8813

pennockd@yahoo-inc.com

<http://dpennock.com/>

Highlights

Currently a Senior Research Scientist at Yahoo! Research Labs

Prior: NEC Research, Microsoft Research, Adjunct Professor at Pennsylvania State University

50 journal and conference publications, including *PNAS*, *Science*, *IEEE Computer*, *Theoretical Computer Science*, *Electronic Commerce Research*, *Electronic Markets*, *AAAI*, *EC*,

WWW, *KDD*, *UAI*, *SIGIR*, *ICML*, *NIPS*, *INFOCOM*, *SAINT*, *ACM SIGCSE*, *VLDB*

3 patents; over 25 presentations

Press interviews include *New York Times*, *Newsweek*, *Washington Post*, *The New Yorker*, *Investors Business Daily*, *LA Business Journal*, *Dow Jones/Wall Street Journal*

Reports on my research have appeared in *Discover Magazine*, *New Scientist*, *CNN/Money*, *New York Times*, *E! Online*, *Tech TV*, Surowiecki's "The Wisdom of Crowds"

Research Interests

– Analysis of new electronic and Web markets

– Design of new markets according to both computational and economic objectives

– Design of systems for "collective intelligence": collaborative filtering, markets, groupware, ensemble machine learning

Topics: electronic commerce, artificial intelligence, recommender systems, machine learning, Internet analysis, decision theory, uncertain reasoning, decentralized coordination

Education

Ph.D. Computer Science UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, DEC 1999

Concentration in Artificial Intelligence; Graduate Certificate in Complex Systems; Attended Santa Fe Institute Complex Systems Summer School, June 1996

Dissertation: "Aggregating Probabilistic Beliefs: Market Mechanisms and Graphical Representations"

M.S. Computer Science DUKE UNIVERSITY, DURHAM, NC, AUG 1994

Masters thesis: "Learning in a Neural Network with Fractal Topology for Image Compression"

B.S. Physics, 2nd Major: Computer Science DUKE UNIVERSITY, DURHAM, NC, MAY 1993

Activities

Santa Fe Institute Complex Systems Summer School, June 1996

Duke in Cambridge Program in England, Summer 1992

Pi Kappa Alpha Fraternity, 1990–1993; Community Service Chair, 1991–1992

Honors

MIT Technology Review's TR35: Top 35 Technology Innovator Under Age 35, 2005

Fellowship, Michigan Decision Behavior Consortium, 1998

Best Student Paper Finalist, Decision Analysis Society, 1998

Graduated magna cum laude, 1993

Dean's List, 1989–1993

Golden Key National Honors Society, 1991

Phi Eta Sigma Freshman Honors Society, 1990

- Y. Chen, T. Mullen, C.-H. Chu and D.M. Pennock (2005). **Information markets vs. opinion pools: An empirical comparison.** *ACM Conference on Electronic Commerce*
- J.Feng, H. Bhargava, and D.M. Pennock (2005). **Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms.** *Inform Journal on Computing*, forthcoming
- A.I. Schein, A. Popescul, L.H. Ungar, and D.M. Pennock (2005). **CROC: A new evaluation criterion for recommender systems.** *Electronic Commerce Research*, 5(1): 51-74
- D.M. Pennock (2004). **A Dynamic pari-mutuel market for hedging, wagering, and information aggregation.** *ACM Conference on Electronic Commerce*
- O. Madani, D.M. Pennock, and G.W. Flake (2004). **Co-validation: Using disagreement on unlabeled data to validate classification algorithms.** *Neural Information Processing Systems*
- E. Servan-Schreiber, J. Wolfers, D.M. Pennock, and B. Galebach (2004). **Prediction markets: Does Money Matter?** *Electronic Markets*,14(3)
- G.W. Flake and D.M. Pennock (2004). **Self-organization, self-regulation, and self-similarity on the fractal web.** Chapter in *The Colours of Infinity*, Clear Press, UK
- J. Feigenbaum, L. Fortnow, D.M. Pennock, and R. Sami (2004). **Computation in a distributed information market.** *Theoretical Computer Science*
Conference version: *ACM Conference on Electronic Commerce, 2003*
- L. Fortnow, J. Kilian, D.M. Pennock, and M.P. Wellman (2004). **Betting boolean-style: A framework for trading in securities based on logical formulas.** *Decision Support Systems*, 39(1):87-104
Conference version: *ACM Conference on Electronic Commerce, 2003*
- S.-T. Park, D.M. Pennock, C.L. Giles, and R. Krovetz (2004). **Analysis of lexical signatures for improving information persistence on the World Wide Web.** *ACM Transactions on Information Systems*, 22(4): 540-572
Conference version: *ACM Conference on Information Retrieval, 2002*
- D. Pavlov, E. Manavoglu, D.M. Pennock and C.L. Giles (2004). **Collaborative filtering with maximum entropy.** *IEEE Intelligent Systems*, 19(6): 40-48
- A. De Bruyn, C.L. Giles, and D.M. Pennock (2004). **Offering collaborative-like recommendations when data is sparse: The case of attraction-weighted information filtering.** *Conference on Adaptive Hypermedia*
- S.-T. Park, D.M. Pennock, and C.L. Giles (2004). **Comparing static and dynamic measurements and models of the Internet's AS topology.** *Joint Conference of the IEEE Computer and Communications Societies*
- G.W. Flake, D.M. Pennock, and D.C. Fain (2003). **The self-organized web: The yin to the semantic web's yang.** *IEEE Intelligent Systems*
- S. Debnath, D.M. Pennock, S. Lawrence, and C.L. Giles (2003). **Information incorporation in online in-game sports betting markets.** *ACM Conference on Electronic Commerce*
- K. Dave, S. Lawrence and D.M. Pennock (2003). **Mining the peanut gallery: Opinion extraction and semantic classification of product reviews.** *World Wide Web Conference*
- D. Pavlov, A. Popescul, D.M. Pennock and L.H. Ungar (2003). **Mixtures of Conditional Maximum Entropy Models.** *International Conference on Machine Learning*
- S.K. Lam, D.M. Pennock, D. Cosley, and S. Lawrence (2003). **1 Billion Pages = 1 Million Dollars? Mining the Web to Play "Who Wants to be a Millionaire?".** *Conference on Uncertainty in Artificial Intelligence*
- S.-T Park, A. Khrabrov, D.M. Pennock, S. Lawrence, C.L. Giles, and L.H. Ungar (2003). **Static and Dynamic Analysis of the Internet's Susceptibility to Faults and Attacks.** *Joint Conference of the IEEE Computer and Communications Societies*
- A. Popescul, L.H. Ungar, S.Lawrence, and D.M. Pennock (2003). **Statistical relational learning for document mining.** *International Conference on Data Mining*, pp. 275-282
- D.Y. Pavlov and D.M. Pennock (2002). **A maximum entropy approach to collaborative filtering in dynamic, sparse, high dimensional domains.** *Neural Information Processing Systems*

Journal and Conference Publications (cont'd)

- E.J. Glover, D.M. Pennock, S. Lawrence, and B. Krovetz (2002). **Inferring hierarchical descriptions**. *Conference on Information and Knowledge Management*
- D.M. Pennock, G.W. Flake, S. Lawrence, E.J. Glover, C.L. Giles (2002). **Winners don't take all: Characterizing the competition for links on the web**. *Proceedings of the National Academy of Sciences*, 99 (8): 5207-5211
- D.M. Pennock, S. Debnath, E.J. Glover, C.L. Giles (2002). **Modeling information incorporation in markets with application to detecting and explaining events**. *Conference on Uncertainty in Artificial Intelligence*
- E. Glover, K. Tsioutsoulouklis, S. Lawrence, D. Pennock, G. Flake (2002). **Using web structure for classifying and describing web pages**. *World Wide Web Conference*
- S. Chakrabarti, M. Joshi, K. Punera, D.M. Pennock (2002). **The structure of broad topics on the Web**. *World Wide Web Conference*
- A.I. Schein, A. Popescul, L.H. Ungar, D.M. Pennock (2002). **Methods and metrics for cold-start recommendations**. *ACM Conference on Information Retrieval*
- D. Cosley, S. Lawrence, D.M. Pennock (2002). **REFEREE: An open framework for practical testing of recommender systems using ResearchIndex**. *Conference on Very Large Data Bases*
- D.M. Pennock, S. Lawrence, F.Å. Nielsen, C. Lee Giles (2001). **Extracting collective probabilistic forecasts from web games**. *Conference on Knowledge Discovery and Data Mining*, pp. 174–183
- A. Popescul, L.H. Ungar, D.M. Pennock, S. Lawrence (2001). **Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments**. *Conference on Uncertainty in Artificial Intelligence*, pp. 437–444
- D.M. Pennock, S. Lawrence, C.L. Giles, F.Å. Nielsen (2001). **The real power of artificial markets**. *Science*, 291(5506): 987–988
- S. Lawrence, D.M. Pennock, G.W. Flake, R. Krovetz, F.M. Coetzee, E. Glover, F.Å. Nielsen, A. Kruger, C.L. Giles (2001). **Persistence of web references in scientific research**. *Computer*, 34(2), pp. 26–31
- E.J. Glover, G.W. Flake, S. Lawrence, W.P. Birmingham, A. Kruger, C.L. Giles, D.M. Pennock (2001). **Improving category specific web search by learning query modifications**. *Symposium on Applications and the Internet*, pp. 23–31
- D.M. Pennock and M.P. Wellman (2000). **Compact securities markets for Pareto optimal reallocation of risk**. *Conference on Uncertainty in Artificial Intelligence*, pp. 481–488
- D.M. Pennock, E. Horvitz, S. Lawrence, C.L. Giles (2000). **Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach**. *Conference on Uncertainty in Artificial Intelligence*, pp. 473–480
- D.M. Pennock, P. Maynard-Reid, C.L. Giles (2000). **A normative examination of ensemble learning algorithms**. *International Conference on Machine Learning*, pp. 735–742
- D.M. Pennock, E. Horvitz, C.L. Giles (2000). **Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering**. *National Conference on Artificial Intelligence*, pp. 729–734
- D.M. Pennock and M.P. Wellman (1999). **Graphical representations of consensus belief**. *Conference on Uncertainty in Artificial Intelligence*, pp. 531–540
- D.M. Pennock (1998). **Logarithmic time parallel Bayesian inference**. *Conference on Uncertainty in Artificial Intelligence*, pp. 431–438
- D.M. Pennock and M.P. Wellman (1997). **Representing aggregate belief through the competitive equilibrium of a securities market**. *Conference on Uncertainty in Artificial Intelligence*, pp. 392–400
- D.M. Pennock and Q.F. Stout (1996). **Exploiting a theory of phase transitions in three-satisfiability problems**. *National Conference on Artificial Intelligence*, pp. 253–258
- D.M. Pennock and M.P. Wellman (1996). **Toward a market model for Bayesian inference**. *Conference on Uncertainty in Artificial Intelligence*, pp. 405–413

Journal and Conference Publications (cont'd)

- C. Connelly, A.W. Biermann, D. Pennock, P. Wu (1996). **Homestudy software: Flexible, interactive and distributed software for independent study.** *ACM SIGCSE Symposium on Computer Science Education*, 28(1), pp. 63–67
- C. Connelly, A.W. Biermann, D. Pennock, P. Wu (1996). **Homestudy software: Complementary systems for computer science courses.** *Computer Science Education*, 7, pp. 53–71
- A.W. Biermann, A.F. Fahmy, C. Guinn, D. Pennock, D. Ramm, P. Wu (1995). **A Computer animated system for demonstrating hardware and software principles.** *Journal of Computing in Small Colleges*, 10(3), p. 34.
- A.W. Biermann, D. Ramm, D. Pennock, A.F. Fahmy, P. Wu (1994). **Visualizing computation: Full color and motion demonstration of computer mechanisms.** *National Conference on College Teaching and Learning*
- A.W. Biermann, A.F. Fahmy, C. Guinn, D. Pennock, D. Ramm, P. Wu (1994). **Teaching a hierarchical model of computation with animation software in the first course.** *ACM SIGCSE Symposium on Computer Science Education*, 26(1), pp. 295–299

Other Publications

- D.M. Pennock (2004). **The Eudaemonic Pie: A review.** *AI Magazine*, 25(2): 125-128
- D.M. Pennock (2004). **Conference report: The fifth ACM conference on electronic commerce.** *SIGecom Exchanges*, 5(1): 48–56
- D.M. Pennock (2001). **NP markets, or How to get everyone else to solve your intractable problems.** *Workshop on Economic Agents, Models, and Mechanisms at the International Joint Conference on Artificial Intelligence*, pp. 89–98
- D.M. Pennock (2001). **Conference report: The second ACM conference on electronic commerce.** *SIGecom Exchanges*, 2(1): 33–38
- P. Rusmevichientong, D.M. Pennock, S. Lawrence, C.L. Giles (2001). **Methods for sampling pages uniformly from the World Wide Web,** *AAAI Fall Symposium on Using Uncertainty Within Computation*
- D.M. Pennock and M.P. Wellman (1999). **The observability of independence under monetary-based elicitation.** *Workshop on Conditional Independence Structures and Graphical Models*, pp. 56–57

Patents

- Term-based concept market (3)**, in preparation, with G. Flake, *Yahoo, Inc.*
- Continuous dynamic pari-mutuel market for trading or wagering assets of uncertain value**, in preparation, *Yahoo, Inc.*
- Methods and apparatus for predicting attribute values based on a personality diagnosis**, pending, with E. Horvitz, *Microsoft Corporation*

Selected Presentations

Markets in uncertainty: Risk, gambling, and information aggregation.

Tutorial at *National Conference on Artificial Intelligence (AAAI)*, San Jose, CA, July 2004

Tutorial at *ACM Conference on Electronic Commerce*, San Diego, CA, June 2003

A Dynamic pari-mutuel market for hedging, wagering, and information aggregation.

Microsoft Research, Redmond, WA, July, 2004

University of Michigan, Ann Arbor, MI, March, 2004

Sports betting markets: Characteristics and information aggregation. *International*

Conference on Gambling and Risk Taking, Vancouver, Canada, May 2003

Information and forecast accuracy in markets and market games.

Google, Inc., Mountain View, CA, September 2002

Overture Services, Inc., Pasadena, CA, August 2002

Modeling information incorporation in markets and market games. *Markets and Decisions*

Workshop, Arlington, VA, June 2002

Semantic explanations of market forecasts. *Controlled Market Economies Symposium*,

Cambridge, MA, May 2002

The power of play: Efficiency, information aggregation, and forecast accuracy in market games. *Institute for Operations Research and the Management Sciences National Meeting*,

Miami, FL, November 2001

Maximizing information, optimizing risk, and leveraging forecasts in securities markets.

NEC Research Symposium, Bonn, Germany, May 2001

Recommender systems. *Penn State eBusiness Research Center Academic Workshop on*

Personalization Issues in e-Business, Arlington, VA, April 2001

Winners don't take all: A model of web link accumulation. *Workshop on Data Mining and*

Learning on the Web at the 14th Conference on Neural Information Processing Systems,

Breckenridge, CO, December 2000

E-markets and uncertainty, or What Bayesians can learn by watching market prices.

Microsoft Research, Redmond, WA, June 2000

Introduction to auctions. *University of Pennsylvania*, Philadelphia, PA, April, 2000

Group Coordination: A History of Paradox and Impossibility.

NEC Research Institute, Princeton, NJ, February 2000

Microsoft Research, Redmond, WA, August, 1998

Efficient representations for aggregate belief and compact securities markets. *Institute for*

Operations Research and the Management Sciences National Meeting, Philadelphia, PA,

November 1999

Representing and forming aggregate probabilities: From graphical models to market mechanisms.

NASA Ames, Moffett Field, CA, June 1999

Washington University, St. Louis, MO, May 1999

AT&T Labs Research, Florham Park, NJ, May 1999

Combining probabilistic models: Impossibility, possibility, and decentralized mechanisms.

IBM Research, Yorktown Heights, NY, March 1999

Toward an efficient, decentralized mechanism for aggregating belief. *Stanford University*,

Palo Alto, CA, February 1999

A market framework for pooling opinions.

RAND Corporation, Santa Monica, CA, July 1998

Institute for Operations Research and the Management Sciences National Meeting,

Montréal, Canada, April 1998

Professional Experience

Senior Research Scientist

YAHOO! RESEARCH LABS, PASADENA, CA, NOV 2002 TO PRESENT

Managing research group and conducting research on information markets, auctions, recommender systems, and machine learning. Developing creative and innovative technologies to help position Yahoo! as a leader in Internet e-commerce and related scientific endeavors

Research Scientist

NEC RESEARCH INSTITUTE, PRINCETON, NJ, OCT 1999 TO OCT 2002

Conducted research on information markets, Web games, recommender systems, Web hyperlinks, social networks, consensus Bayesian networks, and ensemble learning algorithms

Adjunct Assistant Professor of Computer Science and Engineering

PENNSYLVANIA STATE UNIVERSITY, STATE COLLEGE, PA, JUN 2001 TO PRESENT

Taught CSE 597B, Computational aspects of e-commerce

Research Intern MICROSOFT RESEARCH, REDMOND, WA, JUN 1998 TO AUG 1998

Conducted research on impossibility theorems in group coordination, recommender systems, and Bayesian networks

Research Assistant UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, SEP 1995 TO MAY 1999

Conducted research on computationally efficient and decentralized mechanisms for aggregating information using markets

Teaching Assistant UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, SEP 1994 TO DEC 1994

Taught CS 380: Data Structures and Algorithms

Research Assistant DUKE UNIVERSITY, DURHAM, NC, SEP 1993 TO AUG 1994

Co-developed an educational software program titled "This is How a Computer Works", currently still in use at Duke University and elsewhere.

Professional Services

Program Co-Chair, ACM Conference on Electronic Commerce, 2006

Deputy Area Chair, International World Wide Web Conference, 2002

Organizing Committee

Workshop on Sponsored Search Auctions at the ACM Conference on Electronic Commerce, 2005; DIMACS Workshop on Information Markets, 2005; Session on Markets in Uncertainty at the Institute for Operations Research and the Management Sciences National Meeting, 2001; Workshop on Data Mining and Learning on the Web at the 14th Conference on Neural Information Processing Systems, 2000

Senior Program Committee

Conference on Uncertainty in Artificial Intelligence, 2003

Professional Services (cont'd)

Program Committee

International Joint Conference on Artificial Intelligence 2003,2005; International World Wide Web Conference, 2005; ACM Conference on Electronic Commerce, 2004; Conference on Uncertainty in Artificial Intelligence, 2000-2005; Conference on Email and Anti-Spam, 2004,2005; Conference on Very Large Databases, 2005; ACM Conference on Knowledge Discovery and Data Mining 2003; National Conference on Artificial Intelligence, 2002,2005; International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2002,2003; Web Dynamics Workshop at the International World Wide Web Conference, 2002; Workshop on Agent Mediated Electronic Commerce at the International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2002,2003; Workshop on Economic Agents, Models, and Mechanisms at the 17th International Joint Conference on Artificial Intelligence, 2001; Workshop on Internet Bots: Systems and Applications at the 12th International Conference on Database and Expert Systems Applications, 2001; International Symposium on Imprecise Probabilities and Their Applications, 2001

Other Peer Reviewing Services

Artificial Intelligence; Journal of Artificial Intelligence Research; Proceedings of the National Academy of Sciences; Journal of Machine Learning Research; ACM Transactions on Information Systems; IEEE Internet Computing; IEEE Intelligent Systems; IEEE Transactions on Knowledge and Data Engineering; Decision Analysis; Journal of Parallel and Distributed Computing; Management Science; Organizational Behavior and Human Decision Processes; ACM Conference on Electronic Commerce, 2005; National Conference on Artificial Intelligence, 2002; International Joint Conference on Autonomous Agents and Multi-Agent Systems, 2002; International Joint Conference on Artificial Intelligence, 1999, 2001; Conference on Uncertainty in Artificial Intelligence, 1998, 1999; International Conference on Distributed Computing Systems, 1998

Professional Society Memberships

Association for Computing Machinery (ACM); American Association for Artificial Intelligence (AAAI); Institute of Electrical and Electronics Engineers (IEEE)

Post-Doctoral Scientists Managed

Seung-Taek Park (now Research Associate at Yahoo!)
Tracy Mullen (now Assistant Professor at Pennsylvania State University)
Dmitry Pavlov (now Research Engineer at Yahoo!)

Students Managed or Co-Managed

Dan Cosley & Shyong Lam, U. Minnesota; Varsha Dani, U. Chicago; Kushal Davé & Rahul Sami, Yale; Sandip Debnath, Jane Feng, Seung-Taek Park, & Secil Ugurel, Pennsylvania State; Rodrigo Lopes Cançado Fonseca, U. Federal de Minas Gerais; Eric Glover, U. Michigan; A. Khrabrov & Alexandrin Popescul, U. Pennsylvania; Finn Årup Nielsen, Technical U. Denmark; Paat Rusmevichientong, Stanford; Sumit Sanghai, U. Washington

Technical Background

Computer Science

Artificial Intelligence

Electronic commerce, recommender systems, machine learning, information retrieval, decision theory, Bayesian networks, game theory, complex systems, constraint satisfaction, satisfiability, and neural networks

Theory

Parallel algorithms, cellular automata, and computational complexity

Languages

C++, C, PERL, PASCAL, PROLOG, LISP, C*, BASIC, Fortran, assembly, SuperCard, others

Physics

Chaos, non-linear dynamics, quantum mechanics, relativity, optics, holography, percolation theory, thermal (statistical) physics, electronics, classical dynamics, and electricity and magnetism

Math

Probability, non-linear time series analysis, differential equations, financial engineering, fractals, topology, linear algebra, and calculus

References

Available on request