

Ron Kohavi, PhD

ronnyk@live dot com

<http://www.kohavi.com/RonnyKResume.pdf>

<http://www.linkedin.com/in/ronnyk>

Los Altos, CA

Summary

Experienced executive that built and led multiple top-notch cross-disciplinary teams in AI/machine learning, data mining, and controlled experiments at companies including Amazon, Microsoft, SGI, and two startups (one IPO'ed). Technical, hands-on, coded and architected multiple systems, and passionate about accelerating innovation using trustworthy experimentation. Few publications, yet academically influential: over [34,000 citations](#), and **three** papers in the [top 1,000 most-cited papers in Computer Science](#). In 2016, named [5th most influential scholar in AI](#) and [26th in machine learning](#).

Work Experience

- 2014- Distinguished Engineer and General Manager, Analysis and Experimentation, Artificial Intelligence and Research, Microsoft (Bellevue, WA; Bay Area, CA). Leading a team of over 80 Data Scientists, Developers, and Program Managers to accelerate innovation through trustworthy experimentation at Microsoft. Helping change Microsoft's culture from ship-it-awards to data-driven and customer-focused with scientific rigor. Providing a Microsoft-wide experimentation platform (ExP) that runs over 12,000 experiments/year, and center-of-excellence for analysis and experimentation. Teaching classes, running hands-on labs, and organizing conferences. Surprising results worth hundreds of millions of dollars annually would not have been recognized without the platform, and many poor ideas that hurt users would have been launched without the platform. Some results shared in [Seven Rules of Thumb for Web Site Experimenters](#).
- 2010-2014 Partner-level Architect, Bing, Online Services Division, Microsoft. Enable trustworthy and scalable experimentation at Bing. See [Online Controlled Experiments at Large Scale](#).
- 2005-2010 General Manager, [Experimentation Platform](#), Microsoft. Founded, managed, and architected the Experimentation Platform team to enable running and analyzing controlled experiments at Microsoft. Grew and headed Dev/Test/PM/Analyst/Ops/Support teams from scratch to about 50 top notch people, including six principal-level team members. See [Online Experimentation at Microsoft, Sept 2009](#).

- 2003-2005 Director, Data Mining and Personalization, Amazon.com (Seattle, WA).
Manage multiple teams and grew the organization from under 50 to over 90 people. Responsibilities included multiple “two-pizza” teams, such as Amazon’s personalization (two teams), ad automation (SEO/SEM), consumer behavior / data mining, site experimentation, and automated e-mail. Introduced several features estimated to be worth several hundred million dollars in incremental revenue.
- 1998-2003 Vice President, Business Intelligence, Blue Martini Software (San Mateo, CA).
Managed the Business Intelligence Sales Demo team and the Analytic Services team. Helped drive Business Intelligence sales and marketing activities, created success stories, and helped the professional services organization in complex implementations as a center of excellence. The company IPO’ed in 2001 and was acquired in 2005. Prior to VP, Senior Director, Data Mining Applications, where I led the engineering team, including two managers. Responsible for the data collection, transfer (ETL), analysis, reporting, visualization, and campaign management in Blue Martini's products. Designed, architected, and initially coded (in Java) the data mining, reporting, and transfer (ETL) modules.
- 1995-1998 Manager, MineSet, Silicon Graphics Inc. (Mountain View, CA)
Managed the MineSet data mining and visualization engineering team of 15. Designed and coded the analytical (server-side) data mining engines of MineSet on top of [MLC++](#), which I headed at Stanford.
- 1993-1995 Started [MLC++](#), Machine Learning library in C++ for data mining, Stanford University. Designed, coded, and led a team in parallel with my Ph.D. work. Supervised four students, including one 50% research assistant dedicated for the project (paid by grants from NSF and ONR).
The library was adopted as the basis for the analytical engines in MineSet when I moved to SGI and later licensed by Blue Martini Software for \$6M.
See <http://www.sgi.com/tech/mlc/>.
- 1991 Programmer, Verification Group, IBM Research Center, Haifa, Israel.
(summer) Programmed a graphical user interface in C.
- 1985-1988 Manager (Lieutenant), Israeli Defense Forces, Israel.
Managed 12 people at a computer center in a large army base in Israel.
Prior to being the manager, I designed and coded systems.
- 1981-1984 Programmer, International Software, Tel-Aviv, Israel (startup).
Developed a database application generator, IRIS, with three other people (during high school). The program was sold commercially, including sales to the Israeli Defense Forces.

Education

- 1991-1995 [Stanford University](#), Stanford, CA
Ph.D. in Machine Learning, Computer Science.
Thesis: [Wrappers for Performance Enhancement and Oblivious Decision Graphs](#)
Thesis advisors: Jerome Friedman, Nils Nilsson, and Yoav Shoham.
- 1988-1991 [Technion](#), Haifa, Israel.
B.A. in Computer Science, Summa Cum Laude.

Honors

- 2016 Named [5th most influential scholar in AI](#) and [26th in machine learning](#)
- 2014 Distinguished Engineer at Microsoft (one of less than 50 at Microsoft).
- 2014 My papers have over [34,000 citations](#). My [h-index](#), a measure of productivity and impact of published work, is **51** according to [Google Scholar](#). Hirsch, who proposed the metric, suggested that an h-index of 10-12 is considered a useful guideline for tenure decisions at major research universities; a value of about 18 could mean a full professorship; 15-20 could mean a fellowship in the American Physical Society.
Three of my articles are in the top 1,000 most cited articles. The article [Wrappers for Feature Subset Selection](#) is in the [top 300](#) most-cited articles according to [CiteSeerX](#).
- 2009 [Online Experimentation at Microsoft](#), 2009, recognized as **top 30 Microsoft ThinkWeek paper** and an early version of it **won 3rd place** at the Third workshop on Data Mining Case Studies and Practice Prize, 2009
- 1996 IEEE *Tools With Artificial Intelligence* **Best Paper Award** for the paper [Data Mining using MLC++, a Machine Learning Library in C++](#) by Kohavi, Sommerfield, and Dougherty.
- 1992 Passed the Ph.D. Artificial Intelligence qualifying exam with distinction.
- 1989, 1990, 1991 Technion, President's award (top 5%) each year of BA degree.

Professional Activities

1. [Fourteen patents](#) granted.
2. Invited speaker at multiple conferences ([KDD 2015](#); [MIT CODE](#) 2014,2015,2016, eMetrics). See selected talks at <http://www.kohavi.com/ronnyk-talks.html>
3. General Chair, KDD 2004
4. Scientific Advisor to Trusted Opinion, 2007-2008
5. Member of Technical Advisory Board, mySimon, 1999-2000 (until they were bought by CNET)
6. MLC++: [C++ coding standards](#), [MLC++ coding standards](#), [environment](#), and [utilities](#)
7. Program committee member, Knowledge Discovery and Data Mining conference (KDD), 1997-2014
8. Program committee member, International Conference on Machine Learning, 1997-2003
9. Co-chair with Jim Gray, Industrial Track, Knowledge Discovery and Data Mining (KDD), 1999
10. Co-chair (with Carla Brodley), KDD-CUP 2000 (Aug 2000)
11. Co-chair WEBKDD'2003, [WEBKDD'2001](#), [WEBKDD'2000](#)
12. Co-editor (with Foster Provost), special issue of the International Journal Data Mining and Knowledge Discovery on [e-commerce and data mining](#). This special issue is also available as book: [Applications of Data Mining to Electronic Commerce](#)
13. Member of the editorial board, Data Mining and Knowledge Discovery journal,1997, 1998, 1999, 2000, 2001, 2002
14. Co-Editor (with Foster Provost), special issue on applications of machine learning (Volume 30, 1998), journal of Machine Learning
15. Member of the editorial board, journal of Machine Learning, 1997, 1998, 1999

Selected Publications (reverse chronological order)

1. Ron Kohavi, Alex Deng, Roger Longbotham, and Ya Xu, [Seven Rules of Thumb for Web Site Experimenters](#), KDD 2014.
2. Ron Kohavi, Alex Deng, Brian Frasca, Toby Walker, Ya Xu, Nils Pohlmann, [Online Controlled Experiments at Large Scale](#), KDD 2013.
3. Ron Kohavi, Alex Deng, Brian Frasca, Roger Longbotham, Toby Walker, Ya Xu, [Trustworthy Online Controlled Experiments: Five Puzzling Outcomes Explained](#), KDD 2012. [Powerpoint slides](#).
4. Ronny Kohavi, Thomas Crook, Roger Longbotham, Brian Frasca, Randy Henne, Juan Lavista Ferres, Tamir Melamed, [Online Experimentation at Microsoft](#), 2009. **Microsoft ThinkWeek paper recognized as top 30**. An earlier version of [Online Experimentation at Microsoft](#) appeared in the Third workshop on Data Mining Case Studies and Practice Prize, 2009. **The paper won 3rd place**.
5. Ron Kohavi, Roger Longbotham, Dan Sommerfield, and Randal M. Henne, [Controlled Experiments on the Web: Survey and Practical Guide](#), [Data Mining and Knowledge Discovery journal](#), 2009. The paper has over 300 citations according to Google Scholar.
6. Ron Kohavi, Llew Mason, Rajesh Parekh, Zijian Zheng, [Lessons and Challenges from Mining Retail E-Commerce Data](#), [Machine Learning journal](#) volume 57, p. 83-13, Special Issue on Data Mining Lessons Learned, 2004.
7. Ron Kohavi, Neal Rothleder, and Evangelos Simoudis, [Emerging Trends in Business Analytics](#), Communications of the ACM, Volume 45, Number 8, Aug 2002, pages 45-48.
8. Eric Bauer and Ron Kohavi. [An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants](#). The journal [Machine Learning](#) Vol 36, Nos. 1/2, July/August 1999, pages 105-139. The paper has over 2,300 citations according to Google Scholar.
9. Ron Kohavi and George John, [Wrappers for Feature Subset Selection](#). Artificial Intelligence 97, 1997 ([print version](#)). The paper has over 7,000 citations according to Google Scholar.
10. Ron Kohavi, Dan Sommerfield, and James Dougherty. [Data Mining using MLC++, a Machine Learning Library in C++](#). International Journal on Artificial Intelligence Tools vol. 6, No. 4, 1997. The paper received the IEEE *Tools with Artificial Intelligence* **Best Paper Award**. The paper has over 400 citations according to Google Scholar.
11. Ron Kohavi and David Wolpert. [Bias Plus Variance Decomposition for Zero-One Loss Functions](#). In Machine Learning: Proceedings of the Thirteenth International Conference, pages 275-283, July 1996. The paper has over 600 citations according to Google Scholar.
12. Ron Kohavi, [A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection](#). IJCAI 1995. The paper has over 7,000 citations according to Google Scholar.
13. James Dougherty, Ron Kohavi, and Mehran Sahami, [Supervised and Unsupervised Discretization of Continuous Features](#). Machine Learning 1995. The paper has over 2,000 citations according to Google Scholar.