

NOTES FOR THE CENTRAL TABLES OF THE BOOK, *PREDICTIVE ANALYTICS: THE POWER TO PREDICT WHO WILL CLICK, BUY, LIE, OR DIE* — REVISED AND UPDATED EDITION

This document provides citations and comments pertaining to the book's Central Tables of 182 examples.

Rather than clicking, many URLs below that span multiple lines must be copied to your browser, and extraneous spaces manually removed.

*For more information about the book, see the book's website:
www.thepredictionbook.com.*

Central Tables

The 182 examples are organized into these 9 Central Tables:

1. Family and Personal Life
2. Marketing, Advertising, and The Web
3. Financial Risk and Insurance
4. Healthcare
5. Law Enforcement and Fraud Detection
6. Fault Detection, Safety, and Logistical Efficiency
7. Government, Politics, Nonprofit, and Education Table
8. Human Language Understanding, Thought, and Psychology
9. Workforce: Staff and Employees

For more on how these sectors apply predictive analytics, see the corresponding series of industry-focused Predictive Analytics World (PAW—www.pawcon.com) conferences:

- PAW Business
- PAW Financial
- PAW Healthcare
- PAW Manufacturing
- PAW Government
- PAW Workforce

Family and Personal Life—Central Table 1

Nokia:

- David Talbot, “A Phone That Knows Where You’re Going,” *Technology Review Online*, July 8, 2012. www.technologyreview.com/news/428441/a-phone-that-knows-where-youre-going/.
- Nokia Data Center, Nokia Mobile Data Challenge. <http://research.nokia.com/page/12000>.
- KDnuggets, “Nokia Mobile Data Challenge: Your Friends Determine Where You Are Going,” July 2012. www.kdnuggets.com/2012/07/nokia-mobile-data-challenge.html.
- Albert-László Barabási, “Human Behavior Is 93 Percent Predictable, Research Shows,” *PhysOrg.com*, February 23, 2010. <http://phys.org/news186174216.html>.

Microsoft:

- Adam Sadilek and John Krumm, “Far Out: Predicting Long-Term Human Mobility,” Microsoft Research, April 4, 2012. http://research.microsoft.com/en-us/um/people/jckrumm/Publications%202012/Sadilek-Krumm_Far-Out_AAAI-2012.pdf.

Uber:

- Ren Lu, “Making A Bayesian Model To Infer Uber Rider Destinations,” Newsroom Uber Global blog, September 2, 2014. <http://newsroom.uber.com/2014/09/inferring-uber-rider-destinations/>.

Facebook (face recognition):

- Yaniv Taigman, Ming Yang, Marc’Aurello Ranzato and Lior Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” Facebook Research post, June 24, 2014.
<https://research.facebook.com/publications/480567225376225/deepface-closing-the-gap-to-human-level-performance-in-face-verification/>.
- Tom Simonite, “Facebook Creates Software That Matches Faces Almost as Well as You Do,” *MIT Technology Review*, March 27, 2014.
www.technologyreview.com/news/525586/facebook-creates-software-that-matches-faces-almost-as-well-as-you-do/.
- Matt Hicks, “Making Photo Tagging Easier,” Notes by Facebook, updated June 30, 2011.
www.facebook.com/notes/facebook/making-photo-tagging-easier/467145887130.
- Ning Zhang, Manohar Paluri, Yaniv Taigman, Rob Fergus and Lubomir Bourdev, “Beyond Frontal Faces: Improving Person Recognition Using Multiple Cues,” *Computer Vision Foundation*, modified May 25, 2015. www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Zhang_Beyond_Frontal_Faces_2015_CVPR_paper.pdf.
- Aviva Rutkin, “Facebook can recognize you in photos even if you’re not looking,” *New Scientist Online*, June 22, 2015.
www.newscientist.com/article/dn27761-facebook-can-recognise-you-in-photos-even-if-youre-not-looking.

Facebook (news feed optimization):

- Will Oremus, “Facebook’s New Secret Sauce,” *Slate*, April 24, 2014.
www.slate.com/articles/technology/technology/2014/04/facebook_news_feed_edgerank_facebook_algorithms_facebook_machine_learning.html.
- Taylor Hatmaker, “EdgeRank is dead, and other updates about the Facebook News Feed algorithm,” *The Daily Dot*, August 1, 2014. www.dailydot.com/technology/facebook-news-feed-algorithm-edgerank/.
- Lars Backstrom, “News Feed FYI: A Window Into News Feed,” Facebook for Business, August 6, 2013.

www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed.

- Roman Shapovalov, “Machine learning in Facebook,” Computer Blindness blog, December, 2012. <http://computerblindness.blogspot.in/2012/12/machine-learning-in-facebook.html>.
- Sean Davis, “We Cracked The Code On How the Facebook News Feed Algorithm Works,” *The Federalist*, February 20, 2014. <http://thefederalist.com/2014/02/20/we-cracked-the-code-on-how-the-facebook-news-feed-algorithm-works/>.
- Josh Constine, “Why Is Facebook Page Reach Decreasing? More Competition and Limited Attention,” *TechCrunch Online*, April 3, 2014. <http://techcrunch.com/2014/04/03/the-filtered-feed-problem/>.
- Matt McGee, “EdgeRank is Dead: Facebook’s News Feed Algorithm Now Has Close To 100k Weight Factors,” *Marketing Land*, August 16, 2013. <http://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908>.
- Vindu Goel, “Facebook Profit Tripled in First Quarter,” *The New York Times*, April 23, 2014. www.nytimes.com/2014/04/24/technology/facebook-profit-tripled-in-first-quarter.html.

Airbnb (host acceptance):

- Bar Ifrach, “How Airbnb uses machine learning to detect host preferences,” *Airbnb Nerds*, April 14, 2015. <http://nerds.airbnb.com/host-preferences/>.
- Tomio Geron, “Airbnb And The Unstoppable Rise of The Share Economy,” *Forbes*, January 23, 2013. www.forbes.com/sites/tomiogeron/2013/01/23/airbnb-and-the-unstoppable-rise-of-the-share-economy/.

Facebook (friendship):

- Kaggle, “Facebook Recruiting Competition,” June 5, 2012. www.kaggle.com/c/FacebookRecruiting.

LinkedIn:

- Manu Sharma, Principal Research Scientist, LinkedIn, “Iterative, Big Data Analytics and You,” Predictive Analytics World New York City Conference, October 19, 2011, New York, NY.

www.predictiveanalyticsworld.com/newyork/2011/agenda.php#track1-1nk.

More on predicting who you may know (i.e., predicting edges in social network):

- Kaggle, “IJCNN Social Network Challenge,” November 8, 2010. www.kaggle.com/c/socialnetwork.

Match.com:

- Amarnath Thombre, Vice President of Strategy and Analytics, Match.com, “Intelligent Matching at Match.com,” Predictive Analytics World New York City Conference, October 20, 2011, New York, NY. www.predictiveanalyticsworld.com/newyork/2011/agenda.php#track2-9.

OkCupid.com:

- Christian Rudder, “Exactly What to Say in a First Message,” OkTrends from OkCupid.com, September 14, 2009. <http://blog.okcupid.com/index.php/online-dating-advice-exactly-what-to-say-in-a-first-message/>.

Target:

- Andrew Pole, Senior Manager, Media and Database Marketing, Target, “How Target Gets the Most out of Its Guest Data to Improve Marketing ROI,” Predictive Analytics World Washington, DC, Conference, October 18, 2010, Washington, DC. www.predictiveanalyticsworld.com/dc/2010/agenda.php#day1-8a.
- Charles Duhigg, “How Companies Learn Your Secrets,” *New York Times Magazine*, February 16, 2012. www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

University researchers (infidelity):

- Meredith Melnick, “Study: Predicting Whether a Partner Will Cheat Could Be Gender-Specific,” *Time Magazine Online*, Healthland, July 29, 2011. <http://healthland.time.com/2011/07/29/study-predicting-if-a-partner-will-cheat-is-gender-specific/>.
- Bonny Albo, “Can You Predict Cheating?” About.com Dating, June 11, 2012. <http://dating.about.com/b/2011/06/01/predicting-infidelity.htm>.

- Christine E. Garver-Apgar, Steven W. Gangestad, Randy Thornhill, Robert D. Miller, and Jon J. Olp, “Major Histocompatibility Complex Alleles, Sexual Responsivity, and Unfaithfulness in Romantic Couples,” *Psychological Science Research Report*, October 2006.
<http://psy2.ucsd.edu/Bmgorman/garver.pdf>.
- World Science Report, “Genes May Help Predict Infidelity, Study Reports,” World Science Online, November 30, 2006.
www.world-science.net/exclusives/061128_infidelity-genes.htm.

Clinical researchers (divorce):

- The Divorce Probability Calculator, Research page, 2010.
www.divorceprobability.com/research.php.
- Ian Ayres, *Super Crunchers: Why Thinking-by-Numbers Is the New Way to Be Smart* (Bantam Dell, 2007).
www.amazon.com/Super-Crunchers-Thinking-Numbers-Smart/dp/0553805401#reader_0553805401.
- Richard E. Heyman and Amy M. Smith Slep, “The Hazards of Predicting Divorce without Crossvalidation,” NCBI PubMed Central, *Journal of Marriage and Family*. Author manuscript; available in PMC October 25, 2006; published in final edited form as: *Journal of Marriage and Family* 63, no. 2 (May 2001): 473–479. doi:10.1111/j.1741-3737.2001.00473.x.
www.ncbi.nlm.nih.gov/pmc/articles/PMC1622921/.
- Rose McDermott, James H. Fowler, and Nicholas A. Christakis, “Breaking Up Is Hard to Do, Unless Everyone Else Is Doing It Too: Social Network Effects on Divorce in a Longitudinal Sample Followed for 32 Years,” Social Science Research Network Online (SSRN), October 18, 2009.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id51490708.
- Justin Wolfers, “Assessing Your Divorce Risk,” Freakonomics.com, December 2, 2008.
www.freakonomics.com/2008/12/02/assessing-your-divorce-risk/.
- Divorce360.com, “Marriage Calculator—Calculate Your Chances of Getting Divorced.”
www.divorce360.com/content/divorcecalculator.aspx.
- Dr. John Gottman, Gottman Relationship Institute, “Research FAQs Online.” www.gottman.com/49853/Research-FAQs.html.
Note: Drs. Astro and Danielle Teller have vetted and debunked this data analysis in their book, Sacred Cows: The Truth about Divorce and Marriage.

- James Bailey, “Driven by Data: Your Bank Can Predict Your Divorce,” *Forbes Magazine Online*, November 15, 2011. www.forbes.com/sites/teconomy/2011/11/15/driven-by-data-your-bank-can-predict-your-divorce/ (*this has been denied and is unsubstantiated but is included here as food for thought*).

Marketing, Advertising, and the Web— **Central Table 2**

PREMIER Bankcard:

- Rex Pruitt, “The Sweet Spot: PREMIER Bankcard Lifts Revenue with SAS Business Analysis,” SAS Success Online. www.sas.com/success/premierbankcard.html.
- Rex Pruitt, “Using Base SAS and SAS Enterprise Miner to Develop Customer Retention Modeling,” SAS Global Forum 2009. <http://support.sas.com/resources/papers/proceedings09/278–2009.pdf>.

First Tennessee Bank:

- Dan Marks, First Tennessee Bank, “First Tennessee Bank: Analytics Drives Higher ROI from Marketing Programs,” IBM.com, March 9, 2011. www.ibm.com/smarterplanet/us/en/leadership/firsttenbank/assets/pdf/IBM-firstTennBank.pdf.

Target:

- Andrew Pole, Target, “How Target Gets the Most out of Its Guest Data to Improve Marketing ROI,” Predictive Analytics World Washington, DC, Conference, October 19, 2010, Washington, DC. www.predictiveanalyticsworld.com/dc/2010/agenda.php#day1–8a.

Harbor Sweets:

- Kate DeBevois, “Harbor Sweets’ Billie Phillips on Driving Off-Season Sales,” Target Marketing Online, July 2008. www.targetmarketingmag.com/article/harbor-sweets-billie-phillips-driving-off-season-sales-111219/1#.

Fingerhut:

- David L. Olson, “Data Mining,” University of Nebraska, Lincoln, Korea Telecom: KM1, Slide 7.

<http://cbafiles.unl.edu/public/cbainternal/facStaffUploads/KM1D M.ppt>.

- Michael J. A. Berry and Gordon S. Linoff, *Mastering Data Mining: The Art and Science of Customer Relationship Management* (Wiley Computer Publishing, 2000).
www.amazon.com/Mastering-Data-Mining-Relationship-Management/dp/0471331236#reader_0471331236.

Vermont Country Store:

- Michael J. A. Berry and Gordon S. Linoff, *Mastering Data Mining: The Art and Science of Customer Relationship Management* (Wiley Computer Publishing, 2000).
www.amazon.com/Mastering-Data-Mining-Relationship-Management/dp/0471331236#reader_0471331236.

Harrah's Las Vegas:

- Gary Loveman, "Diamonds in the Data Mine," *Harvard Business Review: The Magazine*, May 2003.
<http://hbr.org/2003/05/diamonds-in-the-data-mine/ar/1>.

Cox Communications:

- Jim Ericson, "Perfect Pitch," *Information Management*, April 22, 2008. www.information-management.com/specialreports/2008_73/-10001184-1.html.
- Bob Wood, Cox Communications, "Cox Communications: What Happens Next? Automated Smart Demand Forecasting," Predictive Analytics World New York Conference 2011, October 19, 2011, New York, NY. www.predictiveanalyticsworld.com/newyork/2011/agenda.php#track2-2.

Mutual Fund Investment Management Firm:

- Thanks to Rexer Analytics (www.rexeranalytics.com) for this case study.

UK supermarket:

- Kaggle, "Predicting Grocery Shoppers' Spending Habits," Competition, July 29, 2011. [Kaggle webpage on this no longer available.]

Elie Tahari:

- Robert L. Mitchell, "How BI Is Helping to Predict Fashion Trends," *Computer World Online*, September 12, 2011.
www.computerworld.com/s/article/357932/The_Art_Science_of_Fashion.

- More on retail sales forecasting and prediction in fashion markets:
 - Moriah Nastasi, “Professor uses online data to predict future fashion trends,” *Penn State News*, March 30, 2015. www.predictiveanalyticsworld.com/patimes/professor-uses-online-data-to-predict-future-fashion-trends-0413151/.
 - Samaneh Beheshti-Kashi, Hamid Reza Karimi, Klaus-Dieter Thoben, Michael Lutjen and Michael Teucke, “A survey on retail sales forecasting in fashion markets,” *Systems Science & Control Engineering: An Open Access Journal*, January 20, 2015, Volume 3, Issue 1, 2015. www.tandfonline.com/doi/abs/10.1080/21642583.2014.999389#.VTSS4GTBzGd.

Lifeline Screening:

- Ozgur Dogan, Vice President, Merkle, “Segmented Modeling Applications in Health Care Industry,” Predictive Analytics World Washington, DC Conference, October 21, 2009, Washington, DC. www.predictiveanalyticsworld.com/dc/2009/agenda.php#day2-15.

PREMIER Bankcard:

- Rex Pruitt, “The Sweet Spot: PREMIER Bankcard Lifts Revenue with SAS Business Analysis,” SAS Success Online. www.sas.com/success/premierbankcard.html.
- Rex Pruitt, “Using Base SAS and SAS Enterprise Miner to Develop Customer Retention Modeling,” SAS Global Forum 2009. <http://support.sas.com/resources/papers/proceedings09/278-2009.pdf>.

FedEx:

- Rick Whiting, “Businesses Mine Data to Predict What Happens Next,” *InformationWeek*, May 29, 2006. www.informationweek.com/businesses-mine-data-to-predict-what-hap/188500520.

Optus (Australia):

- Thanks to Tim Manns. For details, see the Predictive Analytics World presentation: Tim Manns, “Know Your Customers by

Knowing Who They Know, and Who They Don't: Leveraging the Power of Social Interaction.”

www.predictiveanalyticsworld.com/TimManns_CaseStudyOptus.pdf.

Sprint:

- Junxiang Lu, PhD, Sprint Communications Company, “Predicting Customer Churn in the Telecommunications Industry—An Application of Survival Analysis Modeling Using SAS,” SAS Institute Inc., SUGI 27, Data Mining Techniques, Paper 114-27. *Proceedings of the Twenty-Seventh Annual SAS Users Group International Conference* (Cary, NC: SAS Institute Inc., 2002). www2.sas.com/proceedings/sugi27/p114-27.pdf.

Telenor (Norway):

- Suresh Vittal, “Optimizing Customer Retention Programs,” Forrester Research White Paper, 2008. www.portraitsoftware.com/resources/white-papers/optimizing-customer-retention-programs.
- Suresh Vittal, “Optimal Targeting through Uplift Modeling: Generating Higher Demand and Increasing Customer Retention While Reducing Marketing Costs,” Forrester Research White Paper, 2008. www.portraitsoftware.com/resources/white_papers/optimal-targeting-through-uplift-modeling.
- Eric Siegel, PhD, “Uplift Modeling: Predictive Analytics Can't Optimize Marketing Decisions without It,” Prediction Impact, Inc. White Paper, June 29, 2011. www.predictiveanalyticsworld.com/signup-uplift-whitepaper.php.

2degrees (New Zealand):

- 11Ants Analytics Customer Case Studies, “2degrees Case Study: 1275% Boost in Churn Identification at 2degrees with 11Ants Customer Churn Analyzer.” www.11antsanalytics.com/casestudies/2degrees_casestudy.aspx.

Lloyds TSB:

- Dr. Patrick Surry, PBS, “Gaining a Competitive Advantage through Improved Customer Experience,” Insurance-Canada.ca, sponsored by Pitney Bowes Software, April 26, 2012. www.insurance-canada.ca/social-business/webinars/presentations/PBS-Webinar-Gaining-a-Competitive-Advantage.pdf.

- Nicholas Radcliffe, “Uplift Modelling: You Should Not Only Measure but Model Incremental Response,” Predictive Analytics World London Conference, November 15, 2010, London, UK. www.predictiveanalyticsworld.com/london/2010/agenda.php#day1-16.

Reed Elsevier:

- John McConnell, “Predicting Renewal Propensities for (and Segmenting) RBI Magazine Subscribers,” Analytical People, 2009. Thanks to John McConnell, Analytical People (www.analyticalpeople.com), for this case study. For details, see the Predictive Analytics World presentation: www.predictiveanalyticsworld.com/JohnMcConnell_CaseStudyReedElsevier.pdf.

IBM:

- Tim Daciuk, IBM, and Bob Humphreys, IBM Canada, “IBM: If We Host It, Will They Come? Predictive Modeling for Event Marketing,” Predictive Analytics World Toronto Conference 2012, April 26, 2012, Toronto, Canada. www.predictiveanalyticsworld.com/toronto/2012/agenda.php#day2-17.

Hewlett-Packard:

- Thanks to Govindarajan Krishnaswamy at Hewlett-Packard for this case study.

Bella Pictures:

- Lauren McKay, “Decisions, Decisions,” *CRM Magazine Online* (CRM.com), May 2009. www.destinationcrm.com/Articles/Editorial/Magazine-Features/Decisions-Decisions-53700.aspx.

Paychex:

- Frank Fiorille, Paychex, “Paychex Staying One Step Ahead of the Competition— Development of a Predictive 401(k) Marketing and Sales Campaign,” Predictive Analytics World Washington, DC, Conference 2010, October 19, 2010, Washington, DC. www.predictiveanalyticsworld.com/dc/2010/agenda.php#day1-12.

Sun Microsystems:

- Eric Siegel, PhD, Prediction Impact, Inc., “Six Ways to Lower Costs with Predictive Analytics,” B-eye-network.com, January 5, 2010. www.b-eye-network.com/view/12269.
- Jim Porzak and Alex Kriney, Sun Microsystems, “Sun Microsystems Marrying Prediction and Segmentation to Drive Sales Leads,” Predictive Analytics World San Francisco Conference 2009, February 19, 2009, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2009/agenda.php#sun.

Amazon.com:

- Matt Marshall, “Aggregate Knowledge Raises \$5M from Kleiner, on a Roll,” Venture Beat Online, December 10, 2006. <http://venturebeat.com/2006/12/10/aggregate-knowledge-raises-5m-from-kleiner-on-a-roll/>.
- Greg Linden, Brent Smith, and Jeremy York, “Amazon.com Recommendations: Item-to-Item Collaborative Filtering,” IEEE Computer Society 1089-7801/03, January/February 2003. www.cs.umd.edu/Bsamir/498/Amazon-Recommendations.pdf.
- USPTO Patent Full-Text and Image Database, “Collaborative Recommendations Using Item-to-Item Similarity Mappings,” United States Patent 6,266,649, Linden et al., July 24, 2001. www.google.com/patents/US6266649.
- Greg Bensinger, “Amazon Wants to Ship Your Package Before You Buy It,” *Wall Street Journal Digits*, January 17, 2014. <http://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it/>.
- Lance Ulanoff, “Amazon Knows What You Want Before You Buy It,” *Predictive Analytics Times*, January 27, 2014. www.predictiveanalyticsworld.com/patimes/amazon-knows-what-you-want-before-you-buy-it.

Netflix:

- Eric Siegel, PhD, “Casual Rocket Scientists: An Interview with a Layman Leading the Netflix Prize, Martin Chabbert,” September 2009. www.predictiveanalyticsworld.com/layman-netflix-leader.php.

Tesco (UK):

- Thanks to Giley Pavey, Head of Innovation and Retail Solutions at dunnhumby, for this case study. For more, see www.tescomedia.com/ or www.dunnhumby.com/uk/about-us-what-we-do.

More cases of predictively couponing at point of sale from Kmart, Kroger, Ralph's, Safeway, Stop & Shop, Target, and Winn-Dixie:

- Doug Henschen, "Catalina Marketing Aims for the Cutting Edge of 'Big Data,'" *InformationWeek*, September 14, 2011. www.informationweek.com/global-cio/interviews/catalina-marketing-aims-for-the-cutting/231600833.
- SAS Customer Success, "Catalina Marketing Helps Predict Customer Behavior with SAS," SAS Customer Success, modified October 19, 2012. www.sas.com/success/catalina.html.

Target:

- Andrew Pole, Target, "How Target Gets the Most out of Its Guest Data to Improve Marketing ROI," Predictive Analytics World Washington, DC, Conference, October 19, 2010, Washington, DC. www.predictiveanalyticsworld.com/dc/2010/agenda.php#day1-8a.

U.S. Bank:

- Michael Grundhoefer, U.S. Bank, "US Bank Raising the Bar in Cross-Sell Marketing with Uplift Modeling," Predictive Analytics World Washington, DC, Conference 2009, October 20, 2009, Washington, DC. www.predictiveanalyticsworld.com/dc/2009/agenda.php#day1-14.

Pandora:

- Linda Tischler, "Algorhythm and Blues: How Pandora's Matching Service Cuts the Chaos of Digital Music," *FastCompany.com*, December 1, 2005. www.fastcompany.com/54817/algorhythm-and-blues.

Spotify:

- Chris Johnson, "Algorithmic Music Discovery at Spotify," Slideshare.net presentation, January 13, 2014. www.slideshare.net/MrChrisJohnson/algorithmic-music-recommendations-at-spotify.
- Sander Dieleman, PhD Student at Ghent University, "Recommending music on Spotify with deep learning," Benanne.github.io blog, August 5, 2014. <http://benanne.github.io/2014/08/05/spotify-cnns.html>.

Google:

- Google is famously tight-lipped about the degree to which machine learning (predictive modeling) is employed to improve search ranking by learning from user feedback and clicks. However, in 2011, Google's "search quality gurus," describing an update (called Panda) to Google search ranking methodology to *Wired* magazine (cf. the citation below), said they designed a method to "look for signals that recreate...that same experience [of user satisfaction]". They "came up with a classifier" by working to "find a plane [in hyperspace] which says that most things on this side of the plane are red, and most of the things on that side of the plane are the opposite of red." This is descriptive and definitional of machine learning.
- Steven Levy, "TED 2011: The 'Panda' That Hates Farms: A Q&A with Google's Top Search Engineers," Steven Levy interview with Amit Singhal and Matt Cutts of Google, March 3, 2011. www.wired.com/business/2011/03/the-panda-that-hates-farms/all/.
- Aaron Wheeler, "How Google's Panda Update Changed SEO Best Practices Forever— Whiteboard Friday," Daily SEO blog, June 23, 2011. www.seomoz.org/blog/how-google-panda-update-changed-seo-best-practices-forever-whiteboard-friday.
- Peter van der Graff, "How Search Engines Use Machine Learning for Pattern Detection," Search Engine Watch, December 1, 2011. <http://searchenginewatch.com/article/2129359/How-Search-Engines-Use-Machine-Learning-for-Pattern-Detection>.

Facebook predicts ad clicks based on context, user, device, etc.:

- Junfeng Pan, He Xinran, Ou Jin, Tianbing XU, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, Joaquin Quiñonero Candela, "Practical Lessons from Predicting Clicks on Ads at Facebook," International Workshop on Data Mining for Online Advertising, (ADKDD, August 24, 2014). <https://research.facebook.com/publications/758569837499391/practical-lessons-from-predicting-clicks-on-ads-at-facebook/>.

Education portal:

- "Case Study: How Predictive Analytics Generates \$1 Million Increased Revenue," case study provided by Prediction Impact, Inc. www.predictiveanalyticsworld.com/casestudy.php.

Google:

- D. Sculley, Robert Malkin, Sugato Basu, and Roberto J. Bayardo, “Predicting Bounce Rates in Sponsored Search Advertisements,” Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009.
www.bayardo.org/ps/kdd2009.pdf.
- Sugato Basu, Ph.D., Google. “Lessons Learned in Predictive Modeling for Ad Targeting,” Predictive Analytics World San Francisco 2011 Conference, March 14, 2011, San Francisco, CA.
www.predictiveanalyticsworld.com/sanfrancisco/2011/agenda.php#day1-7.

MTV:

- “Why Lean and Hungry Entrepreneurs Succeed,” *Futurist Update Magazine* 13, no. 5 (May 2012). www.wfs.org/futurist-update/futurist-update-2012-issues/may-2012-vol-13-no-5#1.

Google:

- Google.com, “About Gmail: So Much Time, So Little Spam.”
<https://mail.google.com/mail/help/intl/en/fightspam/spamexplained.html>.
- Alex Woodie, “Inside Sibyl, Google’s Massively Parallel Machine Learning Platform,” *Datanami*, July 17, 2014.
www.datanami.com/2014/07/17/inside-sibyl-googles-massively-parallel-machine-learning-platform/.

Researchers:

- Sitaram Asur and Bernardo A. Huberman, Social Computing Lab, HP Lab, “Predicting the Future with Social Media,” April 8, 2010. www.hpl.hp.com/research/scl/papers/socialmedia/socialmedia.pdf.
- Mark Brown, “Pop Hit Prediction Algorithm Mines 50 Years of Chart-Toppers for Data,” *Wired Magazine Online*, December 19, 2011. www.wired.com/underwire/2011/12/hit-potential-equation/.
- Malcolm Gladwell, “The Formula,” *New Yorker Magazine Online*, October 16, 2006.
www.newyorker.com/archive/2006/10/16/061016fa_fact6.

Financial Risk and Insurance—Central Table

3

Allstate:

- Kaggle, Allstate, “Predicting Liability for Injury from Car Accidents,” Competition, April 13, 2012.
www.kaggle.com/host/casestudies/allstate. Here is how the \$40 million estimate is derived. Allstate’s 2010 annual report puts bodily injury claims at roughly \$1.6 billion. Industry insiders suggest 20 percent of bodily injury claims actuarially relate to vehicle. If we assume that a 1 percent improvement in error rate impacts claims costs by 0.25 percent, even just doubling the accuracy comes to a \$40 million savings. See also www.iihs.org/research/hldi/fact_sheets/BodilyInjury_0911.pdf.

Accident Fund Insurance:

- Zubair Shams, Accident Fund Insurance Company of America, “Accident Fund Using Text Analytics to Accurately Segment Workers’ Compensation Injuries,” Text Analytics World New York City Conference, October 20, 2011, New York, NY.
www.textanalyticsworld.com/newyork/2011/agenda/full-agenda#track1-5.

Leading international commercial lines insurance provider:

- This was from a case study at a conference presentation; however, the insurance company later rescinded authorization to be named in connection with this example. Here is a reference on the general principle that predictive modeling improves upon standard actuarial methods: Guillaume Briere-Giroux, FSA, MAAA, CFA, “Predictive Modeling Applied to Variable Annuity Lapse Rates,” Predictive Modeling for Variable Annuities, Towers Watson neac 11182010.
www.claudepenland.com/2011/02/12/predictive-modeling-applied-to-variable-annuity-lapse-rates/.

Life insurance companies:

- Dr. Thomas Ashley, Dr. Guizhou Hu, and Chris Stehno, “Does Mortality Modeling Affect Life Insurance Preferred Risk Selection?” ReFocus Conference, Las Vegas, NV, March 14, 2011. www.soa.org/files/pd/2011-lv-refocus-b3.pdf.
- Guizhou Hu, MD, PhD, “Mortality Assessment Technology: A New Tool for Life Insurance Underwriting,” May 23, 2002.

www.knowyournumber.com/pdf/MAT_Validation_NHANES1.pdf.

- Yuhong (Jason) Xue, FSA, MAAA, “Predictive Modeling as Applied to Life Insurance,” Predictive Modelling SOA Annual Meeting 2008, October 21, 2008.
www.claudepenland.com/2011/02/12/predictive-modeling-as-applied-to-life-insurance/.

Chase:

- See the detailed case study in Chapter 4.
- Ali Moazami and Shaolin Li, “Mortgage Business Transformation Program Using CART-Based Joint Risk Modeling (and Including a Practical Discussion of TreeNet),” Salford Systems Data Mining 2005, New York, May 10, 2005.
www.salford-systems.com/doc/MoazamiLi.pdf.

Citigroup:

- Stephan Kudyba, *Managing Data Mining: Advice from Experts* (Cybertech Publishing, 2004).

Canadian Tire:

- Charles Duhigg, “What Does Your Credit-Card Company Know about You?” *New York Times Online*, May 12, 2009.
www.nytimes.com/2009/05/17/magazine/17credit-t.html.

PREMIER Bankcard:

- Rex Pruitt, “The Sweet Spot: PREMIER Bankcard Lifts Revenue with SAS Business Analysis,” SAS Success Online.
www.sas.com/success/premierbankcard.html.
- Rex Pruitt, “Using Base SAS and SAS Enterprise Miner to Develop Customer Retention Modeling,” SAS Global Forum 2009.
<http://support.sas.com/resources/papers/proceedings09/278-2009.pdf>.

Mimoni (Mexico):

- Carlos Soltero, Head of Analytics and Software Development, Mimoni, “Using Predictive Models to Provide More Favorable Short-Term Credit Solutions for the Unbanked,” Predictive Analytics World Chicago Conference, June 9, 2015, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2015/agenda.php#day1_1055c. Full presentation slide deck:
www.predictiveanalyticsworld.com/chicago/2015/presentations/pdf/1055_Track3_Soltero.pdf.

Brasil Telecom (now Oi, which means “hi”):

- Carlos André R. Pinheiro, Alexandre G. Evsukoff, and Nelson F. F. Ebecken, “Revenue Recovering with Insolvency Prevention on a Brazilian Telecom Operator,” SIGKDD Explorations, June 1, 2006. www.sigkdd.org/explorations/issues/8-1-2006-06/9-Pinheiro.pdf.

DTE Energy:

- Ozgur Tuzcu, Risk Manager, DTE Energy, as told to Colin Beasty, “Predicting Debt,” *CRM Magazine*, December 2007. www.destinationcrm.com/Articles/Columns-Departments/Secret-of-My-Success/Predicting-Debt-46927.aspx.

Financial institutions:

- Don Davey, “Collect More for Less: Strategic Predictive Analytics Is Key to Growing Collections and Reducing Costs,” First Data White Paper, April 2009. www.firstdata.com/downloads/thought-leadership/fd_collectmoreforless_whitepaper.pdf.

John Elder:

- See Chapter 1 for this case study.
- For a short autobiographical essay by John Elder, see: Mohamed Medhat Gaber, *Journeys to Data Mining: Experiences from 15 Renowned Researchers* (Springer, 2012), 61–76.

London Stock Exchange:

- “Black Box Traders Are on the March,” *The Telegraph*, August 27, 2006. www.telegraph.co.uk/finance/2946240/Black-box-traders-are-on-the-march.html#disqus_thread.
- Kendall Kim, *Electronic and Algorithmic Trading Technology* (Elsevier, 2007) www.elsevier.com/wps/find/bookdescription.cws_home/711644/description#description.

Various firms:

- Dave Andre, PhD, Cerebellum Capital, “Black Box Trading: Analytics in the Land of the Vampire Squid,” Predictive Analytics World San Francisco Conference, March 14, 2011, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2011/agenda.php#day1-8a.

- Spencer Greenberg, Rebellion Research, “Humans, Rules & Machine Learning: Three Prediction Paradigms,” Prediction Analytics World New York City Conference, October 20, 2011, New York, NY.
www.predictiveanalyticsworld.com/newyork/2011/agenda.php#track3-9.
- Randy Saaf, AlphaGenius, “Using Twitter & the Social Internet to Obtain Above Market Returns,” Text Analytics World San Francisco Conference, March 7, 2012, San Francisco, CA.
www.textanalyticsworld.com/sanfrancisco/2012/agenda/full-agenda#day1430-450-2.

Hopper:

- Dr. Patrick Surry, “How the Bunny Saves You Money,” *Hopper, Research & Data*, January 28, 2015.
www.hopper.com/research/bunny-saves-money/.
- Dr. Patrick Surry, Chief Data Scientist, Hopper, “Buy or Wait? How the Bunny Predicts When to Buy Your Plane Ticket,” Predictive Analytics World Chicago Conference, June 10, 2015, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2015/agenda.php#day2_1115a. Full presentation slide deck:
www.predictiveanalyticsworld.com/chicago/2015/presentations/pdf/1115_Track1_Surry.pdf.

Airbnb (pricing):

- Ellen Huet, “How Airbnb Uses Big Data and Machine Learning to Guide Hosts To The Perfect Price,” *Forbes*, June 5, 2015.
www.forbes.com/sites/ellenhuet/2015/06/05/how-airbnb-uses-big-data-and-machine-learning-to-guide-hosts-to-the-perfect-price/.
- Hector Yee and Bar Ifrach, “Have you ever wondered how Airbnb’s price tips for hosts works?” *Airbnb Nerds*, June 4, 2015.
<http://nerds.airbnb.com/aerosolve/>.

Epagogix:

- Tom Whipple, “Slaves to the Algorithm,” *Intelligent Life Magazine*, May/June 2013.
<http://moreintelligentlife.com/content/features/anonymous/slaves-algorithm?page=full>.
- Malcolm Gladwell, “The Formula: What if you built a machine to predict hit movies?” *The New Yorker*, October 16, 2006 Issue.
www.newyorker.com/magazine/2006/10/16/the-formula.

Healthcare—Central Table 4

Overview of predictive analytics for healthcare—leading textbook:

- Linda Miner, Pat Bolding, Joseph Hilbe, Mitchell Goldstein, Thomas Hill, Robert Nisbet, Nephi Walton, and Gary Miner, *Practical Predictive Analytics and Decisioning Systems for Medicine: Informatics Accuracy and Cost-Effectiveness for Healthcare Administration and Delivery Including Medical Research* (Academics Press, 2014).

A top-five U.S. health insurance company:

- This was disclosed to the author as an anonymous discussion.
- See Chapter 2 for a few more details.

Riskprediction.org.uk:

- Jason J. Smith and Paris P. Tekkis, “Risk Prediction in Surgery,” updated April 2010. <http://riskprediction.org.uk/>. This risk-calculation website includes many references— see <http://riskprediction.org.uk/refs.php>.

University of Iowa Hospitals & Clinics - thanks to John Cromwell for this example:

- John Cromwell, MD, Associate Professor, University of Iowa Hospitals & Clinics, “Real-Time Modeling of Surgical Site Infections,” Predictive Analytics World for Healthcare Boston Conference, October 7, 2014, Boston, MA. www.predictiveanalyticsworld.com/health/2014/agenda.php#day2-115.

For more recent examples of surgery mortality prediction, see:

- C. Wu, F. T. Camacho, A. S. Wechsler, S. Lahey, A. T. Culliford, D. Jordan, J. P. Gold, R. S. Higgins, C. R. Smith, and E. L. Hannan, “Risk Score for Predicting Long-Term Mortality after Coronary Artery Bypass Graft Surgery,” U.S. National Library of Medicine, National Institutes of Health, May 22, 2012. www.ncbi.nlm.nih.gov/pubmed/22547673.
- N. Fukuda, J. Wada, M. Niki, Y. Sugiyama, and H. Mushiake, “Factors Predicting Mortality in Emergency Abdominal Surgery in the Elderly,” U.S. National Library of Medicine, National

Institutes of Health, May 11, 2012. www.ncbi.nlm.nih.gov/pubmed/22578159.

- Juan J. Fibla, corresponding author, Division of General Thoracic Surgery, Mayo Clinic; Alessandro Brunelli, Stephen D. Cassivi, and Claude Deschamps, “Aggregate Risk Score for Predicting Mortality after Surgical Biopsy for Interstitial Lung Disease,” *Interactive CardioVascular and Thoracic Surgery*, May 17, 2012. <http://icvts.oxfordjournals.org/content/early/2012/05/17/icvts.ivs174.abstract>.

Google Flu Trends:

- “Google Beats the CDC: Web Tool Predicts Flu-Related Ed Surge,” Advisory Board Company, The Daily Briefing, January 13, 2012. www.advisory.com/Daily-Briefing/2012/01/13/Google-flu.

Stanford University:

- Sebastian Anthony, “Computer More Accurate Than Human Doctor at Breast Cancer Diagnosis,” ExtremeTech Online, November 10, 2011. www.extremetech.com/extreme/104407-computer-more-accurate-than-human-doctor-at-breast-cancer-diagnosis.
- Andrew H. Beck, Ankur R. Sangoi, Samuel Leung, Robert J. Marinelli, Torsten O. Nielsen, Marc J. van de Vijver, Robert B. West, Matt van de Rijn, and Daphne Koller, “Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features Associated with Survival,” *Science Magazine Online; Science Translational Medicine* 3, issue 108 (November 9, 2011): 108ra113. doi:10.1126/scitranslmed.3002564. <http://stm.sciencemag.org/content/3/108/108ra113>.
- Andrew Myers, “Stanford Team Trains Computer to Evaluate Breast Cancer,” Stanford School of Medicine, November 9, 2011. <http://med.stanford.edu/ism/2011/november/computer.html>.

Sisters of Mercy Health Systems:

- Jeni Fan, “Framework for Detection of Clinical States & Disease Onset Using Electronic Health Record (EHR) Data,” Predictive Analytics World San Francisco Conference, March 6, 2012, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2012/agenda.php#day2-30a.

Researchers:

- Kaggle, “Predict HIV Progression,” Competition, April 27, 2010. <https://www.kaggle.com/c/hivprogression/details/Background>.

Pfizer:

- Max Kuhn, Pfizer, “Right Medicine, Right Patient,” Predictive Analytics World San Francisco Conference, March 6, 2012, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2012/agenda.php#day2-hc4.
- Max Kuhn and Kjell Johnson, *Applied Predictive Modeling* (Springer-Verlag, 2013).

Brigham Young University and University of Utah:

- Heidi Toth, “Research Shows Progress in Predicting Premature Births,” *Daily Herald*, April 20, 2011. www.heraldextra.com/news/local/central/provo/article_c13c1e9a-ef75-585d-a22f-8a09a3cbb500.html.
- M. Sean Esplin, MD, Karen Merrell, PhD, Robert Goldenberg, MD, Yinglei Lai, PhD, Jay D. Iams, MD, Brian Mercer, MD, Catherine Y. Spong, MD, Menachem Miodovnik, MD, Hygriv N. Simhan, MD, Peter van Dorsten, MD, and Mitchell Dombrowski, MD, Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal- Fetal Medicine Units Network, “Proteomic Identification of Serum Peptides Predicting Subsequent Spontaneous Preterm Birth,” *American Journal of Obstetrics & Gynecology* 204, issue 5 (May 2011): 391.e1–391.e8. www.ajog.org/article/S0002-9378%2810%2901167-1/.

Pfizer:

- Pfizer, Inc. Case Study, “Pfizer Enlists CART to Score Male Erectile Dysfunction Diagnostic Test,” Salford Systems. www.salford-systems.com/en/resources/case-studies/118-pfizer-inc.

Heritage Provider Network:

- Health Prize, “Improve Healthcare, Win \$3,000,000,” Competition, April 4, 2011. www.heritagehealthprize.com/c/hhp.

University of Pittsburgh Medical Center:

- Scott Zasadil, UPMC Health Plan, “A Predictive Model for Hospital Readmissions,” Predictive Analytics World Washington, DC, Conference, October 19, 2010, Washington, DC.

www.predictiveanalyticsworld.com/dc/2010/agenda.php#day1-17a.

- Jennifer Zaino, “Baptist Health sees big payoff using predictive analytics,” *Healthcare Finance*, February 11, 2015. www.healthcarefinancenews.com/news/baptist-health-sees-big-payoff-using-predictive-analytics.

For more information regarding analytics and hospital admissions:

- Peter Horner and Atanu Basu, “Analytics & the Future of Healthcare,” *Analytics Magazine Online*, January/February 2012. www.analytics-magazine.org/januaryfebruary-2012/503-analytics-a-the-future-of-healthcare.

FICO:

- Todd Steffes, “Predictive Analytics: Saving Lives and Lowering Medical Bills,” *Analytics Magazine*, Analytics Informs, January/February 2012. www.analytics-magazine.org/januaryfebruary-2012/505-predictive-analytics-saving-lives-and-lowering-medical-bills.

GlaxoSmithKline (UK):

- Vladimir Anisimov, GlaxoSmithKline, “Predictive Analytic Patient Recruitment and Drug Supply Modelling in Clinical Trials,” Predictive Analytics World London Conference, November 30, 2011, London, UK. www.predictiveanalyticsworld.com/london/2011/agenda.php#day1-16.
- Vladimir V. Anisimov, “Statistical Modelling of Clinical Trials (Recruitment and Randomization),” *Communications in Statistics—Theory and Methods* 40, issue 19–20 (2011): 3684–3699. www.tandfonline.com/toc/lsta20/40/19-20.

MultiCare Health System (four hospitals in Washington):

- Karen Minich-Pourshadi for HealthLeaders Media, “Hospital Data Mining Hits Paydirt,” HealthLeaders Media Online, November 29, 2010. www.healthleadersmedia.com/page-1/FIN-259479/Hospital-Data-Mining-Hits-Paydirt.

Medical centers and healthcare providers:

- Customer Potential Management Marketing Group, “Predictive Market Segmentation in Healthcare: Increasing the Effectiveness of Disease Prevention and Early Intervention,” CMP White Paper, CiteSeer^x_B Online, January 14, 2002.

[http://citeseerx.ist.psu.edu/
viewdoc/summary?doi=10.1.1.134.756](http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.134.756).

Blue Cross Blue Shield of Tennessee:

- Rick Whiting, “Businesses Mine Data to Predict What Happens Next,” *InformationWeek*, May 29, 2006.
www.informationweek.com/businesses-mine-data-to-predict-what-hap/ 188500520.

Law Enforcement and Fraud Detection— Central Table 5

Internal Revenue Service:

- Case study described by John Elder and Cheryl Howard in Mohamed Medhat Gaber’s *Journeys to Data Mining: Experiences from 15 Renowned Researchers* (Springer, 2012), 61.

New York State:

- Gazette staff writers, “Tax and Finance Dept. expected to stop \$500 million in fraudulent claims this year,” *The Legislative Gazette Online*, April 14, 2015.
www.legislativegazette.com/Articles-Top-Stories-c-2015-04-14-91421.113122-Tax-and-Finance-Dept-expected-to-stop-500-million-in-fraudulent-claims-this-year.html.

U.S. Department of Defense’s Defense Finance and Accounting Service:

- Thanks to Dean Abbott, Abbott Analytics (www.abbottanalytics.com), for this case study: “Case Study: Defense Finance and Accounting Service; Project Title: Invoice Fraud Detection.” www.abbottanalytics.com/data-mining-case-study-1.php.
- Dean Abbott, Haleh Vafaie, PhD, Mark Hutchins, and David Riney, “Improper Payment Detection in Department of Defense Financial Transactions,” Data Mining Symposium, Washington, DC, March 28–29, 2000.
www.abbottanalytics.com/assets/pdf/Abbott-Analytics-Improper-Payment-Detection-2000.pdf.

U.S. Postal Service:

- Antonia de Medinaceli, “Fraud Detection: Fraught with Frightful Modeling Hurdles,” Predictive Analytics World Chicago Conference, June 25, 2012, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day1-1040b.
- Shawn Hessinger, “Data Mining for Fraud at the US Postal Service,” All Analytics Online, October 21, 2011.
www.allanalytics.com/author.asp?section_id51412&doc_id5234817.

U.S. Postal Service:

- Antonia de Medinaceli, Director of Business Analytics and Fraud Detection, Elder Research, and Bryan Jones, Director, Data Mining Group, USPS OIG, Office of Investigations, “Fighting the Good Fraud Fight: USPS OIG Contract Fraud Detection.” Predictive Analytics World for Government, Washington, DC Conference, September 13, 2011.
www.predictiveanalyticsworld.com/gov/2011/agenda.php#day2-1020-1105.

New York City Medicaid administrators:

- Saratu Ghartey, Chief Program Accountability Officer, NYC Human Resources Administration Investigation, Revenue and Enforcement Administration, “How Luxury Cars and Predictive Analytics Impact & Protect the Integrity of Your Public Assistance Programs,” Predictive Analytics World for Government Conference, September 15, 2014, Washington, DC.
www.predictiveanalyticsworld.com/gov/2014/agenda.php#day1-430b.

The Centers for Medicare and Medicaid Services:

- Federal News Radio Staff, “CMS head Tavenner announces resignation,” *Federal News Radio*, 1500 AM, January 16, 2015. www.federalnewsradio.com/458/3781167/CMS-head-Tavenner-announces-resignation.

\$40+ billion U.S. insurance company:

- Case study at a conference presentation; the insurance company later rescinded authorization to be named in connection with this example.

Aviva Insurance (UK):

- Thanks to Chekuri Swarochish of Capillary Technologies (www.capillary.sg) for this case study example (which was conducted with WNS for knowledge process outsourcing).

Hewlett-Packard:

- Joe Fallon, Investigations Manager, Hewlett-Packard, “Hewlett-Packard Warranty Fraud: Using Data Mining to Identify Fraud,” presentation at Data Mining to Identify Fraud, an Association of Certified Fraud Examiners Boston Chapter Seminar, May 2010. http://rexeanalytics.com/ACFE-decks/ACFE_Presentation_Dist.pdf.

Citizens Bank:

- Jay Zhou, “Building In-Database Predictive Scoring Model: Check Fraud Detection Case Study,” Predictive Analytics World Washington, DC, Conference, October 20, 2009, Washington, DC. www.predictiveanalyticsworld.com/dc/2009/agenda.php#day1-19.

U.S. Social Security Administration also conducts fraud detection (not shown in table):

- Stephanie Kanowitz, “Social Security to step up fraud detection with predictive analytics,” *GNC*, April 22, 2014. www.predictiveanalyticsworld.com/patimes/social-security-step-fraud-detection-predictive-analytics/.

Maryland:

- Nadya Labi, “Misfortune Teller,” *The Atlantic*, January/February 2012. www.theatlantic.com/magazine/archive/2012/01/misfortune-teller/8846/.

Baltimore:

- Jennifer Brown and Karen E. Crummy, “Technology, quick-reaction programs guiding parole reform in other states,” *The Denver Post*, September 23, 2013.

www.denverpost.com/parole/ci_24152942/technology-quick-reaction-programs-guiding-reform-other-states.

Israel Institute of Technology:

- Kira Radinsky and Eric Horvitz, “Mining the Web to Predict Future Events,” Israel Institute of Technology, ACM 978-1-4503-1869-3-/13/02, April 27, 2013.
<http://tx.technion.ac.il/~kirar/files/Radinsky-webtorealworld.pdf>.

Chicago, Los Angeles, Miami, Memphis (TN), Richmond (VA), Santa Cruz (CA), and Vineland (NJ):

- Melissa Fellet, “Cops on the Trail of Crimes That Haven’t Happened,” *New Scientist*, October 12, 2011.
www.newscientist.com/article/mg21128333.400-cops-on-the-trail-of-crimes-that-havent-happened.html.
- Erica Goode, “Sending the Police Before There’s a Crime,” *New York Times on the Web*, August 15, 2011.
www.nytimes.com/2011/08/16/us/16police.html.
- Stephen Hollifield, Information Services Manager, Richmond Police Department, “Curbing Crime with Predictive Analytics,” Predictive Analytics World for Government Washington, DC, Conference, September 12, 2011, Washington, DC.
www.predictiveanalyticsworld.com/gov/2011/agenda.php#day1-1115-1145.
- Carrie Kahn, “At LAPD, Predicting Crimes Before They Happen,” National Public Radio Online, August 14, 2011.
www.npr.org/2011/11/26/142758000/at-lapd-predicting-crimes-before-they-happen.
- Sergeant Christopher Fulcher, Vineland, NJ, Police Department, “To Predict and Serve: Predictive Intelligence Analysis.” Part I, July 5, 2011, <http://policeledintelligence.com/2011/07/05/to-predict-and-serve-predictive-intelligence-analysis-part-i/>. Part II, July 6, 2011, <http://policeledintelligence.com/2011/07/06/to-predict-and-serve-predictive-intelligence-analysis-part-ii/>.
- Emily Badger, “How to Catch a Criminal with Data,” Atlantic Cities Online, March 14, 2012.
www.theatlanticcities.com/technology/2012/03/how-catch-criminal-data/1477/.
- Samuel Greengard, “Policing the Future,” *Communications of the ACM* 55, no. 3 (March 2012), 19–21.
<http://cacm.acm.org/magazines/2012/3/146249-policing-the-future/abstract#comments>.

- Nadya Labi, “Misfortune Teller,” *The Atlantic*, January/February 2012.
www.theatlantic.com/magazine/archive/2012/01/misfortune-teller/8846/.

U.S. Armed Forces:

- Neal Ungerleider, “A Computer Program That Predicts Terrorist Attacks,” Co.EXIST. www.fastcoexist.com/1680540/a-computer-program-that-predicts-terrorist-attacks.
- A. Mannes, J. Shakarian, A. Sliva, and V. S. Subrahmanian, “A Computationally-Enabled Analysis of Lashkar-e-Taiba Attacks in Jammu & Kashmir,” Laboratory for Computational Cultural Dynamics, July 1, 2011. https://lccd-content.umiacs.umd.edu/main/papers/let_eisic_camera.pdf.
- Ryan Jacobs, “Afghan War Games: Computer Scientists Accurately Predict Attacks,” *Mother Jones Online*, July 30, 2012. www.motherjones.com/mojo/2012/07/afghan-war-games-researchers-predict-conflicts.
- *The Economist* Staff, “What Makes Heroic Strife: Computer Models That Can Predict the Outbreak and Spread of Civil Conflict Are Being Developed,” *The Economist*, April 21, 2012. www.economist.com/node/21553006.
- Paulo Shakarian, Margo K. Nagel, Brittany E. Schuetzle, and V. S. Subrahmanian, “Abductive Inference for Combat: Using SCARE-S2 to Find High-Value Targets in Afghanistan,” Association for the Advancement of Artificial Intelligence, last modified August 4, 2011. www.aaai.org/ocs/index.php/IAAI/IAAI-11/paper/view/2740.
- Andrew Zammit-Mangion, Michael Dewar, Visakan Kadirkamanathan, and Guido Sanguinetti, “Point Process Modeling of the Afghan War Diary,” *Proceedings of the National Academy of Sciences of the United States of America*, edited by Stephen E. Fienberg, Carnegie Mellon University, Pittsburgh, PA, and approved June 8, 2012 (received for review February 25, 2012). www.pnas.org/content/early/2012/07/11/1203177109.abstract.

New York City:

- Alex, Howard, “Predictive data analytics is saving lives and taxpayer dollars in New York City,” O’Reilly, June 26, 2012. <http://strata.oreilly.com/2012/06/predictive-data-analytics-big-data-nyc.html>.

Oregon and Pennsylvania:

- Public Safety Checklist Website in Oregon, “The Public Safety Checklist for Oregon,” Criminal Justice Commission, last updated August 11, 2012. <https://risktool.ocjc.state.or.us/psc/>.
- Nadya Labi, “Misfortune Teller,” *The Atlantic*, January/February 2012.
www.theatlantic.com/magazine/archive/2012/01/misfortune-teller/8846/.

The Florida Department of Juvenile Justice:

- Mark Greenwald, “Improving Juvenile Justice for the State of Florida,” A Smarter Planet Blog, April 14, 2010.
<http://asmarterplanet.com/blog/2010/04/improving-juvenile-justice-for-the-state-of-florida.html>.
- Ira M. Schwartz, Peter York, Mark Greenwald, Ana Ramos-Hernandez, and Lisa Feeley, "Using Predictive Analytics and Machine Learning to Improve Juvenile Justice Risk Assessments," *American Society of Criminology*. 2016. The Florida Case Study. Volume 1, No. 1.

Chicago Police Department:

- Megan A. Alderden and Timothy A. Lavery, “Predicting Homicide Clearances in Chicago: Investigating Disparities in Predictors across Different Types of Homicide,” *Sage Journals Homicide Studies*, May 5, 2007.
<http://hsx.sagepub.com/content/11/2/115.abstract>.

Amazon.com:

- Amazon Information Security Data Science Competition.
<http://sites.google.com/site/amazonaccessdatacompetition>.

Researchers (hacker and virus detection):

- Chih-Fong Tsai, Yu-Feng Hsu, Chia-Ying Lin, and Wei-Yang Lin, “Intrusion Detection by Machine Learning: A Review.” *ScienceDirect Online*, May 29, 2009.
www.sciencedirect.com/science/article/pii/S0957417409004801.

Reverse engineering spam filters:

- Blaine Nelson, Marco Barreno, Fuching Jack Chi, Anthony D. Joseph, Benjamin I. P. Rubinstein, Udam Saini, Charles Sutton, J. D. Tygar, and Kai Xia, University of California, Berkeley, “Exploiting Machine Learning to Subvert Your Spam Filter,” April 8, 2008.

http://static.usenix.org/event/leet08/tech/full_papers/nelson/nelson_html/.

Computers detect computer-generated text (for search engine spamming):

- A. Lavoie and M. Krishnamoorthy, “Algorithmic Detection of Computer Generated Text,” 2010.
<http://arxiv.org/abs/1008.0706v1>.

Fault Detection for Safety and Efficiency— Central Table 6

Large commercial satellite company:

- Thanks to Matthew Strampe at Elder Research, Inc. (www.datamininglab.com) for this case study. “Reliability Analytics: Predicting Equipment Failure,” Predictive Analytics World San Francisco Conference, February 7, 2010, San Francisco, CA.
www.predictiveanalyticsworld.com/sanfrancisco/2010/agenda.php#day2-12.

Argonne National Laboratory:

- S. Mohanty, S. Majumdar, and K. Natesan, “A Review of Stress Corrosion Cracking/ Fatigue Modeling for Light Water Reactor Cooling System Components,” Argonne National Laboratory, June 2012. https://inlportal.inl.gov/portal/server.pt/document/106726/environmental_fatigue_pdf.
- Also see a related case study described in John Elder’s chapter in Mohamed Medhat Gaber, *Journeys to Data Mining: Experiences from 15 Renowned Researchers* (Springer, 2012), 61–76.

Con Edison:

- Philip Gross, Albert Boulauger, Marta Arias, David Waltz, Philip M. Long, Charles Lawson, Roger Anders (Columbia University), Matthew Koenig, Mark Mastrocinque, William Fairechio, John A. Johnson, Serena Lee, Frank Doherty, and Arthur Kressner (Consolidated Edison Company of New York), “Predicting Electricity Distribution Feeder Failures Using Machine Learning Susceptibility Analysis,” March 31, 2006.
www.phillong.info/publications/GBAetal06_susc.pdf. This work has been partly supported by a research contract from Consolidated Edison.

Con Edison (manholes):

- Rachel Ehrenberg, “Predicting the Next Deadly Manhole Explosion,” *Wired Online Magazine*, July 7, 2010.
www.wired.com/2010/07/manhole-explosions/.

BNSF Railway:

- C. Tyler Dick, Christopher P. L. Barkan, Edward R. Chapman, and Mark P. Stehly, “Multivariate Statistical Model for Predicting Occurrence and Location of Broken Rails,” Transportation Research Board of the National Academies, January 26, 2007.
<http://trb.metapress.com/content/v2j6022171r41478/>. See also:
http://ict.uiuc.edu/railroad/cee/pdf/Dick_et_al_2003.pdf.

TTX:

- Thanks to Mahesh Kumar at Tiger Analytics for this case study, “Predicting Wheel Failure Rate for Railcars.”

Fortune 500 global technology company:

- Thanks to Dean Abbott, Abbot Analytics (<http://abbottanalytics.com/index.php>) for information about this case study. “Inductive Business-Rule Discovery in Text Mining.” Text Analytics World San Francisco Conference, March 7, 2012, San Francisco, CA.
www.textanalyticsworld.com/sanfrancisco/2012/agenda/full-agenda#day11040-11-2.

Leading payments processor:

- Thanks to Robert Grossman, Open Data Group (<http://opendatagroup.com>), for this case study. “Scaling Health and Status Models to Large, Complex Systems,” Predictive Analytics World San Francisco Conference, February 16, 2010, San Francisco, CA.
www.predictiveanalyticsworld.com/sanfrancisco/2010/agenda.php#day1-17.

Universities in Iran:

- Vahid K. Alilou and Mahammad Teshnehlab, “Prediction of 28-Day Compressive Strength of Concrete on the Third Day Using Artificial Neural Networks,” *International Journal of Engineering (IJE)* 3, issue 6, February 2, 2010.

www.cscjournals.org/csc/manuscript/Journals/IJE/volume3/Issue6/IJE-126.pdf.

Poll: 18 percent of IT departments:

- Esther Shein, “Predictive Analytics Key to Unlocking Big Data’s Value, Reducing Cloud Complexity,” *Network Computing Online*, March 8, 2012. www.networkcomputing.com/storage-networking-management/predictive-analytics-key-to-unlocking-bi/232602217.

Washing machine manufacturer:

- G. Dounias, G. Tselentis, and V. S. Moustakis, “Machine Learning Based Feature Extraction for Quality Control in a Production Line,” *Integrated Computer Aided Engineering* 8, no. 4 (2001), 325–336.
<http://iospress.metapress.com/content/b2546qy4aqb8akty/>.

National Iranian South Oil Company:

- A. Mirzaei-Paiaman and S. Salavati, “The Application of Artificial Neural Networks for the Prediction of Oil Production Flow Rate,” *Taylor & Francis Online* 34, issue 19 (2012), 1834–1843.
www.tandfonline.com/doi/abs/10.1080/15567036.2010.492386.

Shell:

- Thanks to Esther Bongenaar, Lead HR Analytics at Shell International.

RightShip:

- Rebecca Merrett, “Predictive analytics ensures ships stay in shipshape for cargo company,” *CIO*, April 1, 2015.
www.cio.com.au/article/571734/predictive-analytics-ensures-ships-stay-shipshape-cargo-company/.
- Bryan Guenther, QI Program Manager, RightShip, “The Impact of Predictive Analytics on Maritime Safety and Efficiency,” *Predictive Analytics World San Francisco Conference*, March 31, 2015, San Francisco, CA.
www.predictiveanalyticsworld.com/sanfrancisco/2015/agenda.php#day1-1145a. Full presentation slide deck:
www.predictiveanalyticsworld.com/sanfrancisco/2015/presentations/pdf/1145_Track1_Guenther.pdf,

UPS:

- Jack Levis, Senior Director, Process Management, UPS, interview with Eric Siegel, Founder of Predictive Analytics World, “Wise Practitioner – Predictive Analytics World Interview Series, Jack Levis of UPS,” *Predictive Analytics Times*, April 27, 2015.
www.predictiveanalyticsworld.com/patimes/wise-practitioner-predictive-analytics-interview-series-jack-levis-of-ups-04272015/.
- Jack Levis, Senior Director, Process Management, UPS, “UPS Analytics – The Road to Optimization,” Predictive Analytics World Chicago Conference, June 9, 2015, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2015/agenda.php#keynote_910.
- John Dix, “How UPS uses analytics to drive down costs (and no, it doesn’t call it big data),” *NetworkWorld*, December 1, 2014. www.networkworld.com/article/2850874/big-data-business-intelligence/how-ups-uses-analytics-to-drive-down-costs-and-no-it-doesn-t-call-it-big-data.html.

Canadian Automobile Association:

- Thanks to Richard Boire, BoireFillerGroup (www.boirefillergroup.com), for this case study. “The Diminishing Marginal Returns of Variable Creation in Predictive Analytics Solutions,” Predictive Analytics World London Conference, November 15, 2010, London, UK.
www.predictiveanalyticsworld.com/london/2010/agenda.php#day1-16a. For more, see:
www.campana.com/axis/products/SSERS0308.pdf.

Analytics leaders:

- G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and B. Nisbet, *Practical Text Mining and Statistical Analysis for Non-Structured Data Text Applications* (Academic Press, 2012), Part II, Tutorial B, p. 181, by Jennifer Thompson and Thomas Hill.

Continental Airlines:

- Tim Grovac, PASSUR Aerospace, “Predicting the Health of Air Traffic Systems,” Predictive Analytics World San Francisco Conference, February 17, 2010, San Francisco, CA.

www.predictiveanalyticsworld.com/sanfrancisco/2010/agenda.php#day2-8.

A large U.S. carrier:

- Kumar Satyam, Manager, Advisory at PricewaterhouseCoopers, LLP, interview with Bala Deshpande, Conference Co-Chair, Predictive Analytics World for Manufacturing 2015, “Wise Practitioner – Manufacturing Predictive Analytics Interview Series: Kumar Satyam of PricewaterhouseCoopers, LLP,” *Predictive Analytics Times*, May 13, 2015.
www.predictiveanalyticsworld.com/patimes/wise-practitioner-manufacturing-predictive-analytics-interview-series-kumar-satyam-of-pricewaterhousecoopers-llc05132015/.
- Kumar Satyam, Manager, Advisory, PricewaterhouseCoopers, LLP and Alex Mannella, Principal, PricewaterhouseCoopers, LLP, “Utilizing On Board Technologies to Improve Maintenance Practices in Airlines,” Predictive Analytics World for Manufacturing Chicago Conference, June 10, 2015, Chicago, IL.
www.predictiveanalyticsworld.com/mfg/2015/agenda.php#day2_1120.

New South Wales, Australia:

- Kaggle, “RTA Freeway Travel Time Prediction,” Competition, November 23, 2010. www.kaggle.com/c/RTA.

Nokia Siemens Networks:

- Shirish Nagaraj and Kashyap Kamdar, Nokia Siemens Networks, “Understanding Mobile User Outages: Predictive Analytics in Wireless Broadband Networks,” Predictive Analytics World Chicago Conference, June 25, 2012, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day2-1045b.

Government, Politics, Nonprofit, and Education—Central Table 7

Obama for America 2012 Campaign:

- See the sidebar at the end of Chapter 7.

Charlotte Rescue Mission:

- JMP Statistical Discovery Customer Brief, “A Smarter Approach to Fundraising,” August 9, 2011.
www.jmp.com/software/success/pdf/102799_charlotte_rescue.pdf

The Nature Conservancy:

- John Blackwell and Tracy DeCanio, The Nature Conservancy, “Successfully Implementing Predictive Analytics in Direct Marketing,” NESUG 2009. www.nesug.org/Proceedings/nesug09/sa/sa09.pdf

JustGiving:

- Mike Bugembe, JustGiving, “Analytics at JustGiving—Using Data to Unlock and Increase Individual Generosity,” Predictive Analytics World London Conference, November 30, 2011, London, UK.
www.predictiveanalyticsworld.com/london/2011/agenda.php#day1-8b

University of Utah’s School of Business:

- Rick Whiting, “Businesses Mine Data to Predict What Happens Next,” *InformationWeek*, May 29, 2006.
www.informationweek.com/businesses-mine-data-to-predict-what-hap/188500520

City of Boston and other cities - restaurant health code violations:

- DrivenData competition: “Keeping it Fresh: Predict Restaurant Inspections,” DrivenData, July 7, 2015.
www.drivendata.org/competitions/5/
- Peter Bull, “DrivenData Competition: Keeping Boston Fresh,” KDNuggets Blog Post, May, 2015.
www.kdnuggets.com/2015/05/drivendata-competition-keeping-boston-fresh.html
- Jun Seok Kang, Polina Kuznetsova, Michael Luca and Yejin Choi, “Where *Not* to Eat? Improving Public Policy by Predicting Hygiene Inspections Using Online Reviews,” Harvard Business School, September 20, 2013.
<http://people.hbs.edu/mluca/hygiene.pdf>
- Tom Schenk, Jr., Chief Data Officer, Department of Innovation and Technology, City of Chicago, “Transforming Food Inspections Using Predictive Analytics,” Predictive

Analytics World Chicago Conference, June 10, 2015,
Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2015/agenda.php#day2_1000a.

- Mohana Ravindranath, “In Chicago, food inspectors are guided by big data,” *The Washington Post*, September 28, 2014. www.washingtonpost.com/business/on-it/in-chicago-food-inspectors-are-guided-by-big-data/2014/09/27/96be8c68-44e0-11e4-b47c-f5889e061e5f_story.html.

City of Chicago:

- Eric Potash, Subhabrata Majumdar, Eric Rozier, Joe Brew, Andrew Reece, Emile Jorgensen, Rayid Ghani, Alexander Loewi, Joe Walsh and Raed Mansour, “Predictive Modeling for Public Health: Preventing Childhood Lead Poisoning,” University of Chicago, June 19, 2015. <http://dssg.uchicago.edu/2000/02/03/org-cdph.html>.
- Chicago Department of Public Health, Healthy Chicago Project Description, “Predictive Analytics to Prevent Lead Poisoning in Children,” *Data Science For Social Good*, 2014 Project. <http://learndatamining.com/portfolio-items/predictive-analytics-to-prevent-lead-poisoning-in-children>.
- Michael Hawthorne, “Could Chicago prevent childhood lead poisoning before it happens?” *Chicago Tribune*, July 16, 2015. www.chicagotribune.com/news/watchdog/ct-lead-poisoning-solutions-20150707-story.html#page=1.

City of New York:

- Alexander B. Howard, “New York City moves to apply predictive data analytics to preventing fires,” E Pluribus Unum blog, May 20, 2013. <http://e-pluribusunum.org/2013/05/20/new-york-city-predictive-data-analytics-fire/>.
- Elizabeth Dwoskin, “How New York’s Fire Department Uses Data Mining,” *Wall Street Journal Digits*, January 24, 2014. <http://blogs.wsj.com/digits/2014/01/24/how-new-yorks-fire-department-uses-data-mining/>.

University of Melbourne:

- Kaggle, University of Melbourne, “Predict Grant Applications,” Competition, December 13, 2010. www.kaggle.com/c/unimelb.

Energex (Australia):

- Jared McKee, Energex, “Spatially Simulating Electricity Demand Growth,” Predictive Analytics World London Conference, November 30, 2011, London, UK. www.predictiveanalyticsworld.com/london/2011/agenda.php#day1-9a.

U.S. Postal Service Office of Inspector General - thanks to Elder Research, Inc. for this case study example:

- Elder Research Institute Case Study, “Elder Research Prioritizes Building Lease Negotiations with Highest Potential Savings (USPS),” Elder Research Institute, January 8, 2014. http://elderresearch.com/images/case-studies/ERI_Lease_Renewal_Risk_Model.pdf.

Kiva:

- Danielle Muoio, “Students Compete in First Annual Duke DataFest,” *Duke Chronicle*, April 16, 2012. www.dukechronicle.com/article/students-compete-first-annual-duke-datafest. For final results, see <http://stat.duke.edu/datafest/results>.
- G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and B. Nisbet, *Practical Text Mining and Statistical Analysis for Non-Structured Data Text Applications* (Academic Press, 2012), Part II, Tutorial K, p. 417, by Richard Foley of SAS.
- Tap directly into Kiva’s loan database with the Kiva API: <http://build.kiva.org>.

U.S. Social Security Administration:

- Thanks to John Elder, PhD, Elder Research, Inc. (www.datamininglab.com), for this case study. John Elder, PhD, “Text Mining to Fast-Track Deserving Disability Applicants,” Elder Research, Inc., August 7, 2010. http://videlectures.net/site/normal_dl/tag573772/kdd2010_elder_tmft_01.pdf.
- John Elder, PhD, “Text Mining: Lessons Learned,” Text Analytics World San Francisco Conference, March 7, 2012, San Francisco, CA.

www.textanalyticsworld.com/sanfrancisco/2012/agenda/full-agenda#day1520-605.

Infinity Insurance:

- James Taylor, JT on EDM, “Putting Predictive Analytics to Work at Infinity Insurance,” DecisionManagementSolutions.com, September 15, 2009. <http://jtonedm.com/2009/09/15/putting-predictive-analytics-to-work-at-infinity-insurance/>. Thanks to James Taylor, Decision Management Solutions (www.decisionmanagementsolutions.com), for this case study.
- James Taylor, “Putting Predictive Analytics to Work,” Predictive Analytics World Washington, DC, Conference, October 19, 2009, Washington, DC. www.predictiveanalyticsworld.com/dc/2009/agenda.php#day1-5.

British Broadcasting Corporation:

- John McConnell, “Identifying and Helping the Most Vulnerable,” Predictive Analytics World San Francisco Conference, February 16, 2010, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2010/agenda.php#day1-16.

The American Public University System, Arizona State University, Iowa State University, Netherlands’ Eindhoven University, Oklahoma State University, and University of Alabama:

- Tim Daciuk, “Taking the ‘Risk’ out of ‘At-Risk’—Identifying Students Before They Drop Out,” Predictive Analytics World Washington, DC, Conference, October 20, 2010, Washington, DC. www.predictiveanalyticsworld.com/dc/2010/agenda.php#day2-7-2.
- Gerben W. Dekker, Mykola Pechenizkiy, and Jan M. Vleeshouwers, “Predicting Students Drop Out: A Case Study,” Educational Data Mining 2009. www.educationaldatamining.org/EDM2009/uploads/proceedings/dekker.pdf.
- Dursun Delen, “Predicting Student Attrition with Data Mining Methods,” *Journal of College Student Retention: Research, Theory and Practice* 13, no. 1 (2011–2012). <http://baywood.metapress.com/app/home/contribution.asp?referrer5par>

ent&backto5issue,2,7;journal,5,51;linkingpublicationresults,1:300319,1.

- Chong Ho Yu, Samuel DiGangi, Angel Jannasch-Pennell, and Charles Kaprolet, University of Arizona, “A Data Mining Approach for Identifying Predictors of Student Retention from Sophomore to Junior Year,” *Journal of Data Science* 8 (2010): 307–325. www.jds-online.com/file_download/242/JDS-574.pdf.
- Hanover Research, “How Data Mining Helped 11 Universities Improve Student Retention Strategies,” Higher Education blog, January 18, 2012. www.hanoverresearch.com/2012/01/how-11-universities-will-improve-student-retention/.
- SAS Customer Success, “Alabama Posts Gains with Recruiting, Retention,” SAS Customer Story, July 27, 2007. www.sas.com/success/uofalabama.html.

Hewlett Foundation:

- Kaggle, The Hewlett Foundation: Automated Essay Scoring, “Develop and Automated Scoring Algorithm for Student-Written Essays,” Competition, February 10, 2012. www.kaggle.com/c/asap-aes.
- Mark D. Shermis and Ben Hammer, “Contrasting State-of-the-Art Automated Scoring of Essays: Analysis,” National Council on Measurement in Education (NCME), March 29, 2012. http://dl.dropbox.com/u/44416236/NCME%202012%20Paper3_29_12.pdf.
- Randal Stross, “The Algorithm Didn’t Like My Essay,” *New York Times*, June 9, 2012. www.nytimes.com/2012/06/10/business/essay-grading-software-as-teachers-aide-digital-domain.html.
- Kaggle, The Hewlett Foundation: Short Answer Scoring, “Develop a Scoring Algorithm for Student Responses,” Competition, June 25, 2012. www.kaggle.com/c/asap-sas.

University of Phoenix:

- Rebecca Barber and Mike Sharkey, Apollo Group, “Course Correction: Using Analytics to Predict Course Success,” *Learning Analytics and Knowledge*, May 2012, 259–262. <http://dl.acm.org/citation.cfm?id=52330664&dl5ACM&coll5DL>.
- **Georgia State University also predicts academic struggle:** Sophie Quinton, “Are Colleges Invading Their Students’ Privacy?” *The Atlantic*, April 6, 2015. www.theatlantic.com/education/archive/2015/04/is-big-brothers-eye-on-campus/389643/#disqus_thread.

Rio Salado Community College:

- Marc Parry, “Big Data on Campus,” *New York Times*, July 28, 2012. www.nytimes.com/2012/07/22/education/edlife/colleges-awakening-to-the-opportunities-of-data-mining.html.

▪ **Further examples in education:**

Predicting at-risk students in U.S. school districts:

- Himabindu Lakkaraju, David Miller, Everaldo Aguiar, Nasir Bhanpuri, Kecia L. Addison, Carl Shan and Rayid Ghani, “A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes,” University of Chicago, June 23, 2015. http://dssg.uchicago.edu/papers/montgomery_education.pdf.

"Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle":

- Eduventures, Inc., “Predictive Analytics in Higher Education,” January 2013. www.eduventures.com/wp-content/uploads/2013/02/Eduventures_Predictive_Analytics_White_Paper1.pdf.

University of Michigan Department of Physics also predicts final semester grades:

- Tim McKay, Kate Miller, Jared Tritz, “What to do with actionable intelligence: E²Coach as an intervention engine,” University of Michigan Department of Physics, 2012. http://sitemaker.umich.edu/ecoach/files/lak_2012_submission.pdf.

***Jeopardy!* winner:**

- See Chapter 6 for more details.
- Roger Craig, “Data Science Meets the Quiz Show *Jeopardy!*,” Predictive Analytics World Chicago Conference, June 26, 2012, Chicago, IL. www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day2-11.
- NPR Staff, “How One Man Played ‘Moneyball’ with ‘Jeopardy!’,” National Public Radio Online, November 20, 2011. www.npr.org/2011/11/20/142569472/how-one-man-played-moneyball-with-jeopardy.

Facebook, Elsevier, IBM, Pittsburgh Science of Learning Center:

- ACM KDD Cup 2010 Annual Data Mining “Student Performance Evaluation” Challenge.
www.sigkdd.org/kddcup/index.php?section52010&method5info.

Grockit:

- Kaggle, “Improve the State of the Art in Student Evaluation by Predicting Whether a Student Will Answer the Next Test Question Correctly,” Competition, November 18, 2011.
www.kaggle.com/c/WhatDoYouKnow.
- Grockit, “Machine Learning for Student Achievement,” Competition, May 9, 2012.
<http://grockit.com/blog/main/2012/05/09/kaggle-results/>.

For more overviews of education applications of data mining:

- International Educational Data Mining Society.
www.educationaldatamining.org/.
- Ryan S. J. D. Baker and Kalina Yacef, “The State of Educational Data Mining in 2009: A Review and Future Visions,” *International Educational Data Mining Society* 1, issue 1 (October 28, 2009).
www.educationaldatamining.org/JEDM/images/articles/vol1/issue1/JEDMVol1Issue1_BakerYacef.pdf.
- C. Romero and S. Ventura, “Educational Data Mining: A Review of the State of the Art,” *IEEE Transactions* 40, issue 6 (November 2010). <http://ieeexplore.ieee.org/xpl/login.jsp?reload5true&tp5&arnumber55524021>. For more information on the author: (www.uco.es/Bin1romoc/).

Human Language Understanding, Thought, and Psychology—Central Table 8

IBM:

- See Chapter 6, which focuses entirely on this story.
- “The AI Behind Watson—The Technical Article,” *AI Magazine*, Fall 2010. www.aaai.org/Magazine/Watson/watson.php.

University at Buffalo:

- N. Bhaskaran, I. Nwogu, M. G. Frank, and V. Govindaraju, “Lie to Me: Deceit Detection via Online Behavioral Learning,” Automatic Face & Gesture Recognition and Workshops, 2011 IEEE International Conference, March 21–25, 2011.
http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?reload5true&arnumber55771407.

Researchers (lie detection):

- G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and B. Nisbet, *Practical Text Mining and Statistical Analysis for Non-Structured Data Text Applications* (Academic Press, 2012), Part II, Tutorial N, p. 509.
- Christie M. Fuller, David P. Biro, and Rick L. Wilson, “Decision Support for Determining Veracity via Linguistic-Based Cues,” *Decision Support Systems* 46, issue 3 (February 2009), 695–703.
www.sciencedirect.com/science/article/pii/S0167923608001991.

Imperium:

- Kaggle, “Predict Whether a Comment Posted during a Public Discussion Is Considered Insulting to One of the Participants,” Competition, August 7, 2012. www.kaggle.com/c/detecting-insults-in-social-commentary.

British Broadcasting Company:

- Mark Tabladillo, “Data Mining for Social Moderation,” Predictive Analytics World Chicago Conference, June 25, 2012, Chicago, IL. www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day1-315b.

The Hebrew University:

- Oren Tsur, Dmitry Davidov, and Ari Rappoport, “ICWSM—A Great Catchy Name: Semi-Supervised Recognition of Sarcastic Sentences in Online Product Reviews,” Association for the Advancement of Artificial Intelligence (www.aaai.org), 2010. <http://staff.science.uva.nl/Botsur/papers/sarcasmAmazonICWSM10.pdf>.

PayPal:

- Han Sheong Lai, PayPal, “Identifying Customers Who Expressed Intend-to-Churn or Defect from Large Number of Surveyed Verbatim,” Text Analytics World New York Conference, October 18, 2011, New York, NY.
www.textanalyticsworld.com/newyork/2011/agenda/full-agenda#day1-gold.

Citibank:

- Ramendra Sahoo, Citibank, and Hui Xiong, “Analyzing and Scoring Customer Problem Descriptions through Multi-Focal Learning,” Text Analytics World New York Conference, October

20, 2011, New York, NY.
www.textanalyticsworld.com/newyork/2011/agenda/full-agenda#track1-tba4.

Ford Motor Company:

- Kaggle, “Stay Alert! The Ford Challenge,” Competition, January 29, 2011. www.kaggle.com/c/stayalert.

Averitt:

- Qualcomm.com, “C.R. England and Averitt Express Expand Adoption of FleetRisk Advisors’ Predictive Analytics and Remediation Services,” press release, April 11, 2012.
www.qualcomm.com/media/releases/2012/04/11/c-r-england-and-averitt-express-expand-adoption-fleetrisk-advisors-predict.

Air Force:

- Qiang Ji, Zhiwei Zhu, and Peilin Lan, “Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue,” *IEEE Transactions on Vehicular Technology* 53, no. 4 (July 2004).
www.ecse.rpi.edu/~scs/papers/qji/Papers/IEEE_vt.pdf.

Online Privacy Foundation:

- Kaggle, “Identify People Who Have a High Degree of Psychopathy Based on Twitter Usage,” Competition, May 14, 2012. www.kaggle.com/c/twitter-psychopathy-prediction.
- Kashmir Hill, “Using Twitter to Identify Psychopaths,” *Forbes Magazine Online*, July 20, 2012.
www.forbes.com/sites/kashmirhill/2012/07/20/using-twitter-to-help-expose-psychopaths/.
- Lauren Dugan, “Psychopaths Might Give Themselves Away in Their Tweets,” MediaBistro Online, October 13, 2011.
www.mediabistro.com/alltwitter/psychopaths-might-give-themselves-away-in-their-tweets_b14805.

Analytics leaders and a psychiatry professor:

- G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and B. Nisbet, *Practical Text Mining and Statistical Analysis for Non-Structured Data Text Applications* (Academic Press, 2012), Part II, Tutorial I, p. 395.

Yahoo! Labs:

- Physics arXiv, “The Algorithm That Sees Beauty in Photographic Portraits: Scientists at Yahoo Labs have developed a machine learning algorithm that distinguishes

beautiful portraits from not-so,” The Physics arXiv blog, February 4, 2015. <https://medium.com/the-physics-arxiv-blog/the-algorithm-that-sees-beauty-in-photographic-portraits-435ab8064646>.

- Miriam Redi, Nikhil Rasiwasia, Gaurav Aggarwal and Alejandro Jaimes, “The Beauty of Capturing Faces: Rating the Quality of Digital Portraits,” Cornell University Library, arXiv:1501.07304[cs.CV], January 28, 2015. <http://arxiv.org/abs/1501.07304>.

University of California, Berkeley:

- These researchers write, “This [research] is a critical step toward the creation of brain reading devices that can reconstruct dynamic perceptual experiences.” The researchers speculate on applying this technology to reveal what you see in dreams. Practical applications include assisting those unable to communicate verbally (e.g., stroke victims and coma patients). Watch this YouTube clip: www.youtube.com/watch?v=56FsH7RK1S2E.
- Shinji Nishimoto, An T. Vu, Thomas Naselaris, Yuval Benjamini, Bin Yu, and Jack L. Gallant, “Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies,” *Current Biology*, 2011, doi:10.1016/j.cub.2011.08.031. www.sciencedirect.com/science/article/pii/S0960982211009377.
- Laura Shin, “Video: Mind-Reading Tech Reconstructs Moving Images Others See,” *SmartPlanet Online*, September 22, 2011. www.smartplanet.com/blog/science-scope/video-mind-reading-tech-reconstructs-moving-images-others-see/10507.
- Yasmin Anwar, “Scientists Use Brain Imaging to Reveal the Movies in Our Mind,” *UCBerkeley News Center*, September 22, 2011. <http://newscenter.berkeley.edu/2011/09/22/brain-movies/>.

Researchers (thought detection):

- Francisco Pereira, Greg Detre, and Matthew Botvinick, “Generating Text from Functional Brain Images,” *Frontiers Journals Online*, August 23, 2011. www.frontiersin.org/human_neuroscience/10.3389/fnhum.2011.00072/abstract.
- In 2011, IBM predicted that mind-reading technology would be mainstream within five years. Steve Harrim, “The IBM 5 in 5: Our Forecast of Five Innovations That Will Alter the Tech Landscape within Five Years,” December 19, 2011. <http://asmarterplanet.com/blog/2011/12/the-next-5-in-5-our->

forecast-of-five-innovations-that-will-alter- the-landscape-
within-five-years.html.

Radical games:

- Classic 20Q games. <http://20Q.net>.
- Karen Schrock, “Twenty Questions, Ten Million Synapses,” *Scienceline*, July 28, 2006. <http://scienceline.org/2006/07/tech-schrock-20q/>.

Staff and Employees—Human Resources— Central Table 9

Google also practices HR analytics (not listed in the Table):

- Saul Hansell, “Google Answer to Filling Jobs Is an Algorithm,” *The New York Times*, January 3, 2007. www.nytimes.com/2007/01/03/technology/03google.html?_r=2&oref=slogin&.
- Scott Morrison, “Google Searches for Staffing Answers,” *The Wall Street Journal*, May 19, 2009. www.wsj.com/news/articles/SB124269038041932531.
- Chris DeRose, “How Google Uses Data to Build a Better Worker,” *The Atlantic*, October 7, 2013. www.theatlantic.com/business/archive/2013/10/how-google-uses-data-to-build-a-better-worker/280347/.
- Dr. John Sullivan, “How Google Is Using People Analytics To Completely Reinvent HR,” *Ere Media, Talent Management and HR*, February 26, 2013. www.eremedia.com/tlnt/how-google-is-using-people-analytics-to-completely-reinvent-hr/.

Hewlett-Packard:

- See Chapter 2 for this case study.
- Anindya Dey, Hewlett-Packard, and Gitali Halder, Hewlett-Packard, “Attrition Driver Analysis,” Predictive Analytics World London Conference, November 30, 2011, London, UK. www.predictiveanalyticsworld.com/london/2011/agenda.php#day1-5a.
- Jyotirmay Nag, Hewlett-Packard, “An Innovative Approach to Analyze Employee Satisfaction Response in Light of Customer Satisfaction Response,” Predictive Analytics World Chicago Conference, June 26, 2012, Chicago, IL.

www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day2-235a.

Wikipedia:

- Kaggle, “This Competition challenges data-mining experts to build a predictive model that predicts the number of edits that an editor will make five months from the end date of the training dataset,” Competition, June 28, 2011.
www.kaggle.com/c/wikichallenge.
- Karthik Sethuraman, “Crowdsourcing Predictive Analytics: Why 25,000 Heads Are Better Than One,” Predictive Analytics World Chicago Conference, June 25, 2012, Chicago, IL.
www.predictiveanalyticsworld.com/chicago/2012/agenda.php#day1-250a.
- Wendy Zukerman, “Kaggle Contest Aims to Boost Wikipedia Editors,” *New Scientist Online*, July 4, 2011.
www.newscientist.com/blogs/onepercent/2011/07/kaggle-competition-aims-to-giv.html.
- Dell Zhang, “Wikipedia Edit Number Prediction Based on Temporal Dynamics Only,” Cornell University Library, arXiv:1110.5051v1[cs.LG], October 23, 2011.
<http://arxiv.org/abs/1110.5051v1>.
- Xiangju Qin, Derek Greene, Pádraig Cunningham, “A Latent Space Analysis of Editor Lifecycles in Wikipedia,” Cornell University Library, arXiv:1407.7736 [cs:SI], January 29, 2014.
<http://arxiv.org/abs/1407.7736>.

A midsize U.S. retail bank:

- Thanks for this example to Talent Analytics, Corp., who executed on this predictive analytics project for the bank.
- Greta Roberts, “Case Study: Reducing the Costs of Employee Churn with Predictive Analytics,” O’Reilly Media Inc. (2014).
www.talentanalytics.com/wp-content/uploads/2014/12/Talent-Analytics_EmployeeChurn-OReilly-Case-Study.pdf.

Wells Fargo:

- Katie Kuehner-Herbert, “Predictive Analytics for Hiring,” *BAI Banking Strategies*, September 6, 2015.
www.predictiveanalyticsworld.com/patimes/predictive-analytics-for-hiring-0521153/.

University researchers:

- Donald H. Kluemper, Peter A. Rosen, and Kevin W. Mossholder, “Social Networking Websites, Personality Ratings, and the

Organizational Context: More Than Meets the Eye?” *Journal of Applied Social Psychology* 42, issue 5 (May 2012): 1143–1172. Wiley Online Library, 2012.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1559-1816.2011.00881.x/full>.

- Todd Wasserman, “Facebook Profiles Accurately Predict Job Performance [Study],” Mashable Business Online, February 21, 2012. <http://mashable.com/2012/02/21/facebook-profiles-job-performance/>.

U.S. Naval Special Warfare Command:

- Dean Abbott, President, Abbott Analytics, “Hiring and Selecting Special Forces Personnel Using Predictive Analytics,” Predictive Analytics World for Workforce San Francisco Conference, March 31, 2015, San Francisco, CA. www.predictiveanalyticsworld.com/workforce/2015/agenda.php#day1-1030b. Full presentation slide deck: www.predictiveanalyticsworld.com/workforce/2015/presentations/pdf/1030_Track2_Abbott.pdf.

LinkedIn:

- Manu Sharma, LinkedIn, “Data Science at LinkedIn: Iterative, Big Data Analytics and You,” Predictive Analytics World New York Conference, October 19, 2011, New York, NY. www.predictiveanalyticsworld.com/newyork/2011/agenda.php#track1-lnk.
- Scott Nicholson, LinkedIn, “Beyond Big Data: Better Living through Data Science,” Predictive Analytics World Boston Conference, October 1, 2012, Boston, MA. www.predictiveanalyticsworld.com/boston/2012/agenda.php#keynote-900.
- Scott Nicholson, LinkedIn, “Econometric Applications & Extracting Economic Insights from the LinkedIn Dataset,” Predictive Analytics World San Francisco Conference, March 5, 2012, San Francisco, CA. www.predictiveanalyticsworld.com/sanfrancisco/2012/agenda.php#day1-20a.

Careerbuilder:

- Kaggle, “Predict Which Jobs Users Will Apply To,” Competition, August 3, 2012. www.kaggle.com/c/job-recommendation.

- Careerbuilder.com, “Make Your Job Search Easier: Get Job Recommendations from Careerbuilder.”
www.careerbuilder.com/JobSeeker/Resumes/JobRecommendationsNoDID.aspx.

For more on HR analytics, as applied internally by IBM:

- Stephen Baker with Bremen Leak, “Math Will Rock Your World,” *Bloomberg BusinessWeek Magazine Online*, January 22, 2006. www.businessweek.com/stories/2006-01-22/math-will-rock-your-world.

Other analytically-based HR management applications (not included in the Table) and related privacy concerns:

- Michael Fitzpatrick, “Minority office report: Warning over software now being used by bosses that predicts if you're going to steal from the firm, have a nervous breakdown and even have an office affair,” *Daily Mail Online*, February 17, 2015. www.dailymail.co.uk/news/article-2956766/Minority-office-report-Warning-software-used-bosses-predicts-going-steal-firm-nervous-breakdown-office-affair.html.