

Notes

1. CCTV stands for closed-circuit television, also known as video surveillance.
2. Many of the techniques described in this subsection were employed in Schwartz et al. (2013).
3. In contrast, supervised learning techniques, employ external variables when interpreting the data. A good example of supervised learning is regression analysis, which aims to detect patterns within data that best predict a given external variable.
4. Latent Dirichlet Allocation is an example of a cluster analysis approach that is well-suited to language data. It can be used to automatically extract topics typically discussed in a sample of text documents and estimate the probability that a given document discusses a particular topic.
5. The feed from many of such public cameras can be accessed by anyone (see <https://www.cam2project.net/>).
6. Artificial neural networks are computational models inspired by animal brains. They are composed of layers of artificial neurons (analogous to biological neurons) that process information and connections (analogous to synapses).
7. Overfitting occurs when the model (e.g., regression model) describes a given sample very well.
8. Or, in fact, three variables; in most color images, each pixel is defined by the intensity of three colors: red, green, and blue.

References

- 37 mind-blowing YouTube facts, figures, and statistics—2018. (2018, August 4). *MerchDrop*. Retrieved from <https://merchdrop.com/youtube-stats/>
- Abdullah, S., Matthews, M., Frank, E., Doherty, G., Gay, G., & Choudhury, T. (2016). Automatic detection of social rhythms in bipolar disorder. *Journal of the American Medical Informatics Association*, 23, 538–543.
- Acquisti, A., & Gross, R. (2009). Predicting Social Security numbers from public data. *Proceedings of the National Academy of Sciences*, 106, 10975–10980.
- AlHanai, T. W., & Ghassemi, M. M. (2017). Predicting latent narrative mood using audio and physiologic data. *AAAI Conference on Artificial Intelligence*. Retrieved from https://groups.csail.mit.edu/sls/publications/2017/TukaAlHanai_aaai-17.pdf
- Apte, C., & Weiss, S. (1997). Data mining with decision trees and decision rules. *Data Mining*, 13, 197–210.
- Banko, M., & Brill, E. (2001). Scaling to very very large corpora for natural language disambiguation. In Association for Computational Linguistics (Ed.), *Association for Computational Linguistics 39th annual meeting and 10th conference of the European Chapter* (Vol. 1, pp. 26–33). New Brunswick, NJ: Association for Computational Linguistics.
- Barata, F., Kowatsch, T., Tinschert, P., & Filler, A. (2016). Personal MobileCoach: Tailoring behavioral interventions to the needs of individual participants. In P. Lukowicz (Ed.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1089–1094). New York, NY: ACM.
- Barchard, K. A., & Williams, J. (2008). Practical advice for conducting ethical online experiments and questionnaires for United States psychologists. *Behavior Research Methods*, 40, 1111–1128.
- Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology*, 24, 738–743.
- Bemelmans, R., Gelderblom, G. J., Jonker, P., & de Witte, L. (2012). Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *Journal of the American Medical Directors Association*, 13, 114–120.
- Ben-Zeev, D., Brenner, C. J., Begale, M., Duffecy, J., Mohr, D. C., & Mueser, K. T. (2014). Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophrenia Bulletin*, 40, 1244–1253.

- Bi, B., Shokouhi, M., Kosinski, M., & Graepel, T. (2013). Inferring the demographics of search users. In D. Schwabe (Ed.), *Proceedings of the 22nd International Conference on World Wide Web* (pp. 131–140). Geneva, Switzerland: International World Wide Web Conferences Steering Committee.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bleidorn, W., Hopwood, C. J., & Wright, A. G. (2017). Using big data to advance personality theory. *Current Opinion in Behavioral Sciences*, 18, 79–82.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489, 295–298.
- Boyd, R. L., & Pennebaker, J. W. (2015). Did Shakespeare write double falsehood? Identifying individuals by creating psychological signatures with text analysis. *Psychological Science*, 26, 570–582.
- Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinion in Behavioral Sciences*, 18, 63–68.
- Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015). Values in words: Using language to evaluate and understand personal values. In Association for the Advancement of Artificial Intelligence (Ed.), *Proceedings of the Ninth International AAAI Conference on Web and Social Media*. Palo Alto, CA: AAAI.
- Bruckman, A. (2002). Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology*, 4, 217–231.
- Buchanan, E. A., Aycock, J., Dexter, S., Dittrich, D., & Hvizdak, E. (2011). Computer science security research and human subjects: Emerging considerations for research ethics boards. *Journal of Empirical Research on Human Research Ethics*, 6, 71–83.
- Cadwalladr, C. (2016, December 4). Google, democracy and the truth about internet search. *The Guardian*. Retrieved from <https://www.theguardian.com>
- Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356, 183–186.
- Campbell, A. T., Eisenman, S. B., Lane, N. D., Miluzzo, E., Peterson, R. A., Lu, H., . . . Ahn, G. S. (2008). The rise of people-centric sensing. *IEEE Internet Computing*, 12(4), 12–21.
- Carey, A. L., Brucks, M. S., Küfner, A. C. P., Holtzman, N. S., Deters, F. G., Back, M. D., . . . Mehl, M. R. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology*, 109(3), e1–e15.
- Chamorro-Premuzic, T., Akhtar, R., Winsborough, D., & Sherman, R. A. (2017). The datafication of talent: How technology is advancing the science of human potential at work. *Current Opinion in Behavioral Sciences*, 18, 13–16.
- Chen, E. E., & Wojcik, S. P. (2016). A practical guide to big data research in psychology. *Psychological Methods*, 21, 458–474.
- Chon, J., & Cha, H. (2011). LifeMap: A smartphone-based context provider for location-based services. *IEEE Pervasive Computing*, 10(2), 58–67.
- Chow, P. L., Fua, K., Huang, Y., Bonelli, W., Xiong, H., Barnes, L. E., & Teachman, B. A. (2017). Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. *Journal of Medical Internet Research*, 19(3), e62.
- Cowie, R., Douglas-Cowie, E., Savvidou, S., McMahon, E., Sawey, M., & Schröder, M. (2000). "FEELTRACE": An instrument for recording perceived emotion in real time. In R. Cowie (Ed.), *Proceedings of the ISCA Workshop on Speech and Emotion* (pp. 19–24). Belfast, Northern Ireland: Textflow.
- D'Andrea, W., Chiu, P. H., Casas, B. R., & Deldin, P. (2012). Linguistic predictors of post-traumatic stress disorder symptoms following 11 September 2001. *Applied Cognitive Psychology*, 26, 316–323.
- Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). No country for old members: User lifecycle and linguistic change in online communities. In D. Schwabe (Ed.), *Proceedings of the 22nd International Conference on World Wide Web* (pp. 307–318). Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

- de Montjoye, Y.-A., Quoidbach, J., Robic, F., & Pentland, A. S. (2013). Predicting personality using novel mobile phone-based metrics. In A. M. Greenberg, W. G. Kennedy, & N. D. Bos (Eds.), *Social computing, behavioral-cultural modeling and prediction* (pp. 48–55). Washington, DC: Springer.
- de Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. S. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, 347, 536–539.
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78.
- Eagle, N., & Pentland, A. S. (2009). Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63, 1057–1066.
- Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G. J., Labarthe, D. R., Merchant, R. M., . . . Seligman, M. E. P. (2015). Psychological language on twitter predicts county-level heart disease mortality. *Psychological Science*, 26, 159–169.
- Frankel, M. S., & Siang, S. (1999). *Ethical and legal aspects of human subjects research on the Internet*. Washington, DC: American Association for the Advancement of Science.
- Freud, S. (2011). *On aphasia: A critical study* (E. Stengel, Trans.). Whitefish, MT: Literary Licensing. (Original work published 1891)
- Gillan, C. M., & Whelan, R. (2017). What big data can do for treatment in psychiatry. *Current Opinion in Behavioral Sciences*, 18, 34–42.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012–1014.
- Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for large-scale sentiment classification: A deep learning approach. *Proceedings of the 28th International Conference on Machine Learning*, 2011, 513–520.
- Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333, 1878–1881.
- Gosling, S. D., Sandy, C. J., John, O. P., & Potter, J. (2010). Wired but not WEIRD: The promise of the Internet in reaching more diverse samples. *Behavioral and Brain Sciences*, 33, 94–95.
- Greenberg, D. M., & Rentfrow, P. J. (2017). Music and big data: A new frontier. *Current Opinion in Behavioral Sciences*, 18, 50–56.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Hall, T., & Flynn, V. (2001). Ethical issues in software engineering research: A survey of current practice. *Empirical Software Engineering*, 6, 305–317.
- Hao, T., Xing, G., & Zhou, G. (2013). iSleep. In C. Petrioli (Ed.), *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (pp. 1–14). New York, NY: ACM.
- Harari, G. M., Müller, S. R., Aung, M. S. H., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences*, 18, 83–90.
- Hemminki, S., Nurmi, P., & Tarkoma, S. (2013). Accelerometer-based transportation mode detection on smartphones. C. Petrioli (Ed.), *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (pp. 1–14). New York, NY: ACM.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466, 29.
- Hernandez, J., Hoque, M. E., & Picard, R. W. (2012). Mood meter: Counting smiles in the wild. In A. K. Dey, H.-H. Chu, & G. R. Hayes (Eds.), *Proceedings of the ACM Conference on Ubiquitous Computing* (pp. 301–310). New York, NY: ACM.
- Herring, S. (1996). Linguistic and critical analysis of computer-mediated communication: Some ethical and scholarly considerations. *Information Society*, 12, 153–168.
- Hill, K. (2012, February 16). How Target figured out a teen girl was pregnant before her father did. *Forbes*. Retrieved from <https://www.forbes.com>

- Jacobson, R. (2013, April 24). 2.5 quintillion bytes of data created every day. How does CPG & retail manage it? *IBM Consumer Products Industry Blog*. Retrieved from <https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/>
- Jebara, B. (2014). Bio-mechanical characterization of voice for smoking detection. In European Association for Signal Processing (Ed.), *Proceedings of the 22nd European Signal Processing Conference* (pp. 2475–2479). Piscataway, NJ: IEEE.
- Jones, N. M., Wojcik, S. P., Sweeting, J., & Silver, R. C. (2016). Tweeting negative emotion: An investigation of twitter data in the aftermath of violence on college campuses. *Psychological Methods*, 21, 526–541.
- Kangas, O., Simanainen, M., & Honkanen, P. (2017). Basic income in the Finnish Context. *Intereconomics*, 52, 87–91.
- Kern, M. L., Park, G. J., Eichstaedt, J. C., Schwartz, A. H., Sap, M., Smith, L. K., & Ungar, L. H. (2016). Gaining insights from social media language: Methodologies and challenges. *Psychological Methods*, 21, 507–525.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196–217.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *Quarterly Journal of Economics*, 133, 237–293.
- Kobayashi, T., Boase, J., Suzuki, T., & Suzuki, T. (2015). Emerging from the cocoon? Revisiting the tele-cocooning hypothesis in the smartphone era. *Journal of Computer-Mediated Communication*, 20, 330–345.
- Korpela, J., Miyaji, R., Maekawa, T., Nozaki, K., & Tamagawa, H. (2015). Evaluating tooth brushing performance with smartphone sound data. In K. Mase (Ed.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 109–120). New York, NY: ACM.
- Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science*, 28, 1675–1682.
- Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D. J., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, 95, 357–380.
- Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. J. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, 70, 543–556.
- Kosinski, M., Stillwell, D. J., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110, 5802–5805.
- Kosinski, M., Wang, Y., Lakkaraju, H., & Leskovec, J. (2016). Mining big data to extract patterns and predict real-life outcomes. *Psychological Methods*, 21, 493–506.
- Kramer, A. D. L., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111, 8788–8790.
- Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. In D. D. Lee, U. von Luxburg, R. Garnett, M. Sugiyama, & I. Guyon; Neural Information Processing Systems (Eds.), *Advances in neural information processing systems* (pp. 4069–4079). Red Hook, NY: Curran.
- Lambiotte, R., & Kosinski, M. (2015). Tracking the Digital Footprints of Personality. *Proceedings of the Institute of Electrical and Electronics Engineers*, 102, 1934–1939.
- Landers, R. N., Brusso, R. C., Cavanaugh, K. J., & Collmus, A. B. (2016). A primer on theory-driven web scraping: Automatic extraction of big data from the Internet for use in psychological research. *Psychological Methods*, 21, 475–492.
- Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A survey of mobile phone sensing. *IEEE Communications Magazine*, 48(9), 140–150.
- Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety [blog]. *META Group Research Note*, 6. Retrieved from <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

- Lathia, N., Sandstrom, G. M., Mascolo, C., & Rentfrow, P. J. (2017). Happier people live more active lives: Using smartphones to link happiness and physical activity. *PLoS One*, 12, e0160589.
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, 343, 1203–1205.
- Lazer, D., Pentland, A. S., Adamic, L. A., Aral, S., Barabási, A.-L., Brewer, D., . . . Van Alstyne, M. (2009). Computational social science. *Science*, 323, 721–723.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.
- Likamwa, R., Liu, Y., Lane, N. D., & Zhong, L. (2013). MoodScope: Building a mood sensor from smartphone usage patterns. In H. H. Chu (Ed.), *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services* (pp. 389–402). New York, NY: ACM.
- Lister, M. (2018, October 17). 33 mind-boggling Instagram stats & facts for 2018 [blogpost]. Retrieved from <https://www.wordstream.com/blog/ws/2017/04/20/instagram-statistics>
- Lu, H., Pan, W., Lane, N. D., Choudhury, T., & Campbell, A. T. (2009). SoundSense: Scalable sound sensing for people-centric applications on mobile phones. In Association for Computing Machinery (Ed.), *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services* (pp. 165–178). New York, NY: ACM.
- Luhmann, M. (2017). Using Big Data to study subjective well-being. *Current Opinion in Behavioral Sciences*, 18, 28–33.
- Manago, A. M., Graham, M. B., Greenfield, P. M., & Salimkhan, G. (2008). Self-presentation and gender on MySpace. *Journal of Applied Developmental Psychology*, 29, 446–458.
- Marcus, B., Machilek, F., & Schütz, A. (2006). Personality in cyberspace: personal web sites as media for personality expressions and impressions. *Journal of Personality and Social Psychology*, 90, 1014–1031.
- Matz, S. C., Gladstone, J. J., & Stillwell, D. J. (2016). Money buys happiness when spending fits our personality. *Psychological Science*, 27, 715–725.
- Matz, S. C., & Netzer, O. (2017). Using big data as a window into consumers' psychology. *Current Opinion in Behavioral Sciences*, 18, 7–12.
- Mestyán, M., Yasserli, T., & Kertész, J. (2013). Early prediction of movie box office success based on Wikipedia activity big data. *PLoS One*, 8, e71226.
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7, 221–237.
- Molokken-Ostfold, K. (2005). Ethical concerns when increasing realism in controlled experiments with industrial participants. In Institute of Electrical and Electronics Engineers (Ed.), *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (p. 264a). Los Alamitos, CA: IEEE Computer Society.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60, 683–729.
- Murnane, E. L., Abdullah, S., Matthews, M., Kay, M., Kientz, J. A., Choudhury, T., . . . Cosley, D. (2016). Mobile manifestations of alertness: Connecting biological rhythms with patterns of smartphone app use. In F. Paternò, K. Väänänen, & K. Church (Eds.), *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 465–477). New York, NY: ACM.
- Naughton, F., Hopewell, S., Lathia, N., Schalbroeck, R., Brown, C., Mascolo, C., . . . Sutton, S. (2016). A context-sensing mobile phone app (Q Sense) for smoking cessation: A mixed-methods study. *JMIR MHealth and UHealth*, 4(3), e106.
- Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D. J., & Rentfrow, P. J. (2018). Musical preferences predict personality: Evidence from active listening and Facebook likes. *Psychological Science*, 29, 1145–1158.
- Number of mobile phone users worldwide from 2015 to 2020 (in billions). (2016, November). *Statista*. Retrieved from <https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>
- O'Neil, C. (2016). Weapons of Math Destruction. *Discover*, 37(8), 50–55.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716.

- Park, G. J., Schwartz, A. H., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., . . . Seligman, M. E. P. (2014). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*, 108, 934–952.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). New York, NY: Guilford.
- Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PLoS One*, 9, e115844.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NY: Erlbaum.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, 1296–1312.
- Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, 72, 863–871.
- Penzel, I. B., Persich, M. R., Boyd, R. L., & Robinson, M. D. (2017). Linguistic evidence for the failure mindset as a predictor of life span longevity. *Annals of Behavioral Medicine*, 51, 348–355.
- Perez, S. (2017, October 30). Skype's big redesign publicly launches to desktop users. *Tech Crunch*. Retrieved from <https://techcrunch.com/2017/10/30/skypes-big-redesign-publicly-launches-to-all-desktop-users/>
- Phan, T. Q., & Airolidi, E. M. (2015). A natural experiment of social network formation and dynamics. *Proceedings of the National Academy of Sciences*, 112, 6595–6600.
- Pressman, S. D., & Cohen, S. (2007). Use of social words in autobiographies and longevity. *Psychosomatic Medicine*, 69, 262–269.
- Quercia, D., Kosinski, M., Stillwell, D. J., & Crowcroft, J. (2011). Our Twitter profiles, our selves: Predicting personality with Twitter. In Institute of Electrical and Electronics Engineers (Ed.), *Proceedings of the International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing* (pp.180–185). Los Angeles, CA: IEEE Computer Society.
- Ratkiewicz, J., Conover, M. D., Meiss, M., Gonc, B., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media. In Association for the Advancement of Artificial Intelligence (Ed.), *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media* (pp. 297–304). Palo Alto, CA: AAAI
- Rhodes, R. E., & Smith, N. E. I. (2006). Personality correlates of physical activity: a review and meta-analysis. *British Journal of Sports Medicine*, 40, 958–965.
- Robinson, R. L., Navea, R., & Ickes, W. (2012). Predicting final course performance from students' written self-introductions: A LIWC analysis. *Journal of Language and Social Psychology*, 32, 469–479.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641.
- Schultze, U., & Mason, R. O. (2012). Studying cyborgs: Re-examining Internet studies as human subjects research. *Journal of Internet Technology*, 27, 301–312.
- Schwartz, A. H., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . . Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS One*, 8, e73791.
- Schwartz, A. H., Sap, M., Kern, M. L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., . . . Ungar, L. H. (2016). Predicting individual well-being through the language of social media. In R. Altman (Ed.), *Pacific Symposium on Biocomputing 2016*. Singapore: World Scientific.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366.
- Singer, J., & Vinson, N. (2001). Why and how research ethics matters to you. Yes, You! *Empirical Software Engineering*, 6, 287–290.
- The Skype Team. (2016, January 12). Ten years of Skype video: Yesterday, today and something new... [blogpost]. Retrieved from <https://blogs.skype.com/stories/2016/01/12/ten-years-of-skype-video-yesterday-today-and-something-new/>
- Sletto, R. F., & Meehl, P. E. (1955). Clinical versus Statistical Prediction. *American Sociological Review*, 20, 482–483.

- Smith, K. (2018, March 5). 47 incredible Facebook statistics and facts. *Brand Watch*. Retrieved from <https://www.brandwatch.com/blog/47-facebook-statistics/>
- Snijders, C., Matzat, U., & Reips, U. (2012). "Big data": Big gaps of knowledge in the field of Internet science. *International Journal of Internet Science*, 7, 1–5.
- Solberg, L. (2010). Data mining on Facebook: A free space for researchers or an IRB nightmare? *University of Illinois Journal of Law, Technology & Policy*, 2, 311–343.
- Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics*, 118, 26–40.
- Sunnebo, D. (2017, January 25). Nearly 16% of US consumers now own wearables. *Kantar Worldpanel*. Retrieved from <https://www.kantarworldpanel.com/global/News/Nearly-16-of-US-Consumers-and-9-in-EU4-Now-Own-Wearables>
- Sweeney, L. (1997). Weaving technology and policy together to maintain confidentiality. *Journal of Law, Medicine & Ethics*, 25, 98–110.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29, 24–54.
- Tay, L., Jebb, A. T., & Woo, S. E. (2017). Video capture of human behaviors: Toward a big data approach. *Current Opinion in Behavioral Sciences*, 18, 17–22.
- Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, 49, 144–159.
- The top 20 valuable Facebook statistics. (2018, October). *Zephoria Digital Marketing*. Retrieved from <https://zephoria.com/top-15-valuable-facebook-statistics/>
- Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). *The anatomy of the Facebook social graph*. Unpublished manuscript. <https://arxiv.org/pdf/1111.4503.pdf>
- US Federal Reserve System. (2016). Federal reserve payment study 2016. Retrieved from <https://www.federalreserve.gov/newsevents/press/other/2016-payments-study-20161222.pdf>
- Vazire, S., & Gosling, S. D. (2004). E-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, 87, 123–132.
- Wang, Q., Guo, B., Peng, G., Zhou, G., & Yu, Z. (2016). CrowdWatch: Pedestrian safety assistance with mobile crowd sensing. In P. Lukowicz (Ed.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 217–220). New York, NY: ACM.
- Wang, R., Aung, M. S. H., Abdullah, S., Brian, R., Campbell, A. T., Choudhury, T., . . . Ben-zeev, D. (2016). CrossCheck: Towards passive sensing and detection of mental health changes in people with schizophrenia. In P. Lukowicz (Ed.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 1–12). New York, NY: ACM.
- Wang, R., Chen, F., Chen, Z., Li, T., Harari, G. M., Tignor, S., . . . Campbell, A. T. (2014). StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In ACM Special Interest Group on Spatial Information, ACM Special Interest Group on Mobility of Systems, Users, Data and Computing, & ACM Special Interest Group on Computer-Human Interaction (Eds.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 3–14). New York, NY: ACM.
- Wang, R., Harari, G. M., Hao, P., Zhou, X., & Campbell, A. T. (2015). SmartGPA: How smartphones can assess and predict academic performance of college students. In K. Mase (Ed.), *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 295–306). New York, NY: ACM.
- Wang, Y., & Kosinski, M. (2017). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology*, 144, 246–257.
- Waskul, D. (1996). Considering the electronic participant: Some polemical observations on the ethics of on-line research. *Information Society*, 12, 129–140.
- Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of facebook research in the social sciences. *Perspectives on Psychological Science*, 7, 203–220.

- Wojcik, S. P., Hovasapian, A., Graham, J. W., Motyl, M., & Ditto, P. H. (2015). Conservatives report, but liberals display, greater happiness. *Science*, 347, 1243–1246.
- Working time. (n.d.). Wikiwand. Retrieved from https://www.wikiwand.com/en/Working_time
- Yang, C., & Srinivasan, P. (2016). Life satisfaction and the pursuit of happiness on Twitter. *PLoS One*, 11, e0150881.
- Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, 44, 363–373.
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12, 1100–1122.
- Youyou, W., Kosinski, M., & Stillwell, D. J. (2015). Computer-based personality judgements are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112, 1036–1040.
- Youyou, W., Stillwell, D. J., Schwartz, A. H., & Kosinski, M. (2017). Birds of a feather do flock together: Behavior-Based personality-assessment method reveals personality similarity among couples and friends. *Psychological Science*, 28, 276–284.