

NOTES

1. HYBRIDIZATION: COMBINING ALGORITHMS, AUTOMATION, AND PEOPLE IN NEWSWORK

1. Bastian Obermayer and Frederik Obermaier, *The Panama Papers: Breaking the Story of How the Rich and Powerful Hide Their Money* (London, UK: Oneworld Publications, 2017).
2. Mar Cabra, “How We Built the Data Team behind the Panama Papers,” *Source*, November 2017, <https://source.opennews.org/articles/how-we-built-data-team-behind-panama-papers/>.
3. Mar Cabra and Erin Kissane, “Wrangling 2.6TB of Data: The People and the Technology behind the Panama Papers,” April 26, 2016, <https://panamapapers.icij.org/blog/20160425-data-tech-team-ICIJ.html>; the actual tool is open sourced and can be found at <https://github.com/ICIJ/extract>, accessed October 21, 2018.
4. James Manyika et al., “Harnessing Automation for a Future That Works,” McKinsey Global Institute, 2017, <https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works>.
5. Mary (Missy) Cummings, “Man Versus Machine or Man + Machine?” *IEEE Intelligent Systems* 29, no. 5 (2014).
6. Peter Denning, “Is Computer Science Science?” *Communications of the ACM* 48, no. 4 (2005).
7. Raja Parasuraman, Thomas Sheridan, and Christopher Wickens, “A Model for Types and Levels of Human Interaction with Automation,” *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30, no. 3 (2000).
8. Mark Hansen et al., *Artificial Intelligence: Practice and Implications for Journalism* (New York: Tow Center for Digital Journalism, 2017).
9. Erik Brynjolfsson and Andrew McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (New York: W. W. Norton, 2014).
10. H. P. Luhn, “The Automatic Creation of Literature Abstracts,” *IBM Journal of Research and Development* 2, no. 2 (1958).
11. Susan Straus, “Testing a Typology of Tasks: An Empirical Validation of McGrath’s (1984) Group Task Circumplex,” *Small Group Research* 30, no. 2 (1999).
12. Sarah Cohen, James T. Hamilton, and Fred Turner, “Computational Journalism,” *Communications of the ACM* 54, no. 10 (2011).
13. Geoffrey Bowker et al., *Sorting Things Out: Classification and Its Consequences* (Cambridge, MA: MIT Press, 2000).
14. Reuben Binns et al., “Like Trainer, Like Bot? Inheritance of Bias in Algorithmic Content

- Moderation,” presented at the International Conference on Social Informatics (SocInfo), Oxford, UK, 2017.
15. Michael Schudson, *The Sociology of News*, 2nd. ed. (New York: W.W. Norton, 2011).
 16. Pierre Bourdieu, “The Political Field, The Social Science Field and the Journalistic Field,” in *Bourdieu and The Journalistic Field*, ed. Rodney Benson and Erik Neveu (Cambridge, UK: Polity, 2005).
 17. Matt Carlson, “The Many Boundaries of Journalism,” in *Boundaries of Journalism Professionalism, Practices and Participation*, ed. Seth C. Lewis and Matt Carlson (London: Routledge, 2015).
 18. Mark Deuze, “What Is Journalism?: Professional Identity and Ideology of Journalists Reconsidered,” *Journalism* 6, no. 4 (2005).
 19. David Ryfe, *Journalism and the Public (Key Concepts in Journalism)* (Cambridge, UK: Polity, 2017); Jane Singer, “Out of Bounds: Professional Norms as Boundary Markers,” in *Boundaries of Journalism Professionalism, Practices and Participation*, ed. Seth C. Lewis and Matt Carlson (London: Routledge, 2015).
 20. Bob Kovach and Tom Rosenstiel, *The Elements of Journalism*, 3rd ed. (New York: Three Rivers Press, 2014); Ryfe, *Journalism and the Public*.
 21. Society of Professional Journalists (SPJ) Code of Ethics, September 2014, <https://www.spj.org/ethicscode.asp>.
 22. For an example of this model applied to the context of newswire services, see: Barbara Czarniawska, *Cyberfactories: How News Agencies Produce News* (Cheltenham, UK: Edward Elgar, 2011).
 23. Robert S. Taylor, *Value-Added Processes in Information Systems* (Norwood, NJ: Ablex Publishing Corporation, 1986).
 24. Nicholas Diakopoulos, *Cultivating the Landscape of Innovation in Computational Journalism* (New York: Tow-Knight Center for Entrepreneurial Journalism, 2012).
 25. Sašo Slaček Brlek, Jurij Smrke, and Igor Vobič, “Engineering Technologies for Journalism in the Digital Age,” *Digital Journalism* 5, no. 8 (2017).
 26. Mary Lynn Young and Alfred Hermida, “From Mr. and Mrs. Outlier to Central Tendencies,” *Digital Journalism* 3, no. 3 (2014); James T. Hamilton and Fred Turner, “Accountability through Algorithm: Developing the Field of Computational Journalism,” Report from the Center for Advanced Study in the Behavioral Sciences Summer Workshop, 2009, <https://web.stanford.edu/~fturner/Hamilton%20Turner%20Acc%20by%20Alg%20Final.pdf>; Nicholas Diakopoulos, “A Functional Roadmap for Innovation in Computational Journalism,” January 2010, <http://www.nickdiakopoulos.com/2011/04/22/a-functional-roadmap-for-innovation-in-computational-journalism/>; James T. Hamilton, *Democracy’s Detectives* (Cambridge, MA: Harvard University Press, 2016); C. W. Anderson and David Caswell, “Computational Journalism,” in *The International Encyclopedia of Journalism Studies*, ed. Tim Vos and Folker Hanusch (London: Sage, 2018). For a description of the emergence and various facets of the field over its first dozen years see Neil Thurman, “Computational Journalism,” in *The Handbook of Journalism Studies*, Second Edition, ed. Karin Wahl-Jorgensen and Thomas Hanitzsch (New York: Routledge, 2019).
 27. Matt Carlson, “Automating Judgment? Algorithmic Judgment, News Knowledge, and Journalistic Professionalism,” *New Media & Society* 8, no. 4 (2017); Wendy Espeland and Mitchell Stevens, “A Sociology of Quantification,” *European Journal of Sociology* 49, no. 3 (2009).
 28. Sergio Splendore, “Quantitatively Oriented Forms of Journalism and Their Epistemology,” *Sociology Compass* 10, no. 5 (2015); Nikki Usher, *Interactive Journalism: Hackers, Data, and Code* (Urbana: University of Illinois Press, 2016); Mark Coddington, “Clarifying Journalism’s Quantitative Turn,” *Digital Journalism* 3, no. 3 (2014), does an excellent job parsing out some of the nuances of these terms.
 29. John Pavlik, “The Impact of Technology on Journalism,” *Journalism Studies* 1, no. 2 (2000); Henrik Örnebring, “Technology and Journalism-as-Labour: Historical Perspectives,” *Journalism* 11, no. 1 (2010).
 30. Matthew Powers, “‘In Forms That Are Familiar and Yet-to-Be Invented’: American Journalism and the Discourse of Technologically Specific Work,” *Journal of Communication Inquiry* 36, no. 1 (2011).
 31. Building computer systems becomes a form of “engineering activism”: Helen Nissenbaum, “How Computer Systems Embody Values,” *Computer* 34, no. 3 (2001).

32. Coddington, "Clarifying Journalism's Quantitative Turn."
33. For a treatment of ethics with respect to automated journalism, see Konstantin Dörr and Katharina Hollnbuchner, "Ethical Challenges of Algorithmic Journalism," *Digital Journalism* 5, no. 4 (2017).
34. Ryfe, *Journalism and the Public*.
35. Leo Leppänen et al., "Finding and Expressing News from Structured Data," presented at AcademicMindtrek, Helsinki, Finland, 2017.
36. Batya Friedman et al., "Value Sensitive Design and Information Systems," in *Human-Computer Interaction and Management Information Systems*, ed. Ping Zhang, and Denis Galletta (Oxford, UK: Routledge, 2006).
37. In terms of ethical implications Kirsten Martin argues that "developers make a moral choice as to the delegation of who-does-what between algorithms and individuals" because design determines where in the system careful human consideration of a decision is employed, and where potentially inferior automated decisions may inadequately substitute for humans in uneven ways. For more detail, see Kirsten Martin, "Ethical Implications and Accountability of Algorithms," *Journal of Business Ethics*, published ahead of print (2018).
38. Frank Levy and Richard J. Murnane, *The New Division of Labor* (Princeton, NJ: Princeton University Press, 2004).
39. Nathalie Riche et al., eds., *Data-Driven Storytelling* (Boca Raton, FL: CRC Press, 2018); Karin Assmann and Nicholas Diakopoulos, "Negotiating Change: Audience Engagement Editors as Newsroom Intermediaries," *#ISOJ Journal* 7, no. 1 (2017).
40. For a clear comparison of the relative capabilities of humans and machines, see section 3.3.6 of Ben Shneiderman et al., *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 6th ed. (Boston: Pearson, 2017). For a comparison that is explicitly applied to automated content creation, see Arjen van Dalen, "The Algorithms behind the Headlines," *Digital Journalism* 6, nos. 5–6 (2012).
41. Noam Lemelshtrich Latar, *Robot Journalism: Can Human Journalism Survive?* (Singapore: World Scientific Publishing Company, 2018).
42. Erik Brynjolfsson and Tom Mitchell, "What Can Machine Learning Do? Workforce Implications," *Science* 358, no. 6370 (2017).
43. Jeanette Wing, "Computational Thinking," *Communications of ACM* 49, no. 3 (2006).
44. Peter Denning, "Remaining Trouble Spots with Computational Thinking," *Communications of the ACM* 60, no. 6 (2017).
45. Nicholas Diakopoulos, "Computational Journalism and the Emergence of News Platforms," in *The Routledge Companion to Digital Journalism Studies*, ed. Bob Franklin and Scott Eldridge II (London: Routledge, 2016).
46. Gregor Aisch, "Seven Features You'll Want in Your Next Charting Tool," *Vis4.net* (blog), 2015, <http://vis4.net/blog/posts/seven-features-youll-wantin-your-next-charting-tool/>.
47. Carl-Gustav Linden, "Algorithms for Journalism: The Future of News Work," *Journal of Media Innovations* 4, no. 1 (2017).
48. Mark Ackerman, "The Intellectual Challenge CSCW—The Gap Between Social Requirements and Technical Feasibility," *Human-Computer Interaction* 15, no. 2 (2000).
49. Thomas B. Sheridan, "Function Allocation: Algorithm, Alchemy or Apostasy?" *International Journal of Human-Computer Studies*, 52, no. 2 (2000); Richard Patton and Peter Patton, "What Can Be Automated? What Cannot Be Automated?" in *Springer Handbook of Automation*, ed. Shimon Y. Nof (Heidelberg: Springer, 2009).
50. Robert Picard, "The Economics of Journalism and News Provision," in *Journalism (Handbooks of Communication Science)*, ed. Tim Vos (Berlin, Boston: Mouton De Gruyter, 2018).
51. Seth C. Lewis and Oscar Westlund, "Actors, Actants, Audiences, and Activities in Cross-Media News Work," *Digital Journalism* 3, no. 1 (2014).
52. Matthew Weber and Allie Kosterich, "Coding the News," *Digital Journalism* 6, no. 3 (2018).
53. Brynjolfsson and McAfee, *The Second Machine Age*.

54. Michael Bernstein et al., “Soylent: A Word Processor with a Crowd Inside,” *Communications of the ACM* 58, no. 8 (2015); Aniket Kittur et al., “CrowdForge: Crowdsourcing Complex Work,” *Proceedings of the Symposium on User Interface Software and Technology (UIST)* (New York: ACM, 2011); Elena Agapie, Jaime Teevan, and Andrés Monroy-Hernández, “Crowdsourcing in the Field: A Case Study Using Local Crowds for Event Reporting,” presented at the third AAAI Conference on Human Computation and Crowdsourcing, San Diego, CA, 2015.
55. Tanja Aitamurto, “Crowdsourcing in Open Journalism,” in *The Routledge Companion to Digital Journalism Studies*, ed. Bob Franklin and Scott Eldridge II (London: Routledge, 2016); Tanja Aitamurto, “Crowdsourcing as a Knowledge-Search Method in Digital Journalism,” *Digital Journalism* 4, no. 2 (2015); Elliot Higgins, “Crowdsourcing Conflict and Beyond,” in *Data Journalism Past, Present, Future*, ed. John Mair, Richard Lance Keeble, and Megan Lucero (Bury St Edmunds, UK: Abramis, 2017).
56. Aniket Kittur et al., “The Future of Crowd Work,” *Proceedings of the Conference on Computer Supported Cooperative Work and Social Computing (CSCW)* (New York: ACM, 2013).
57. Gianluca Demartini, “Hybrid Human–Machine Information Systems: Challenges and Opportunities,” *Computer Networks* 90, no. 29 (2015).
58. Daniela Retelny, Michael Bernstein, and Melissa Valentine, “No Workflow Can Ever Be Enough: How Crowdsourcing Workflows Constrain Complex Work,” *Proceedings on the ACM on Human Computer Interaction* 1 (2017).
59. Ackerman, “The Intellectual Challenge of CSCW.”
60. Laura Hazard Owen, “How 7 News Organizations Are Using Slack to Work Better and Differently,” Nieman Lab, July 30, 2015, <http://www.niemanlab.org/2015/07/how-7-news-organizations-are-using-slack-to-work-better-and-differently/>.
61. Agapie, Teevan, and Monroy-Hernández, “Crowdsourcing in the Field.”
62. Ibid.
63. Steven Rich, “What I Learned by Submitting the Same FOIA 1,033 Times (and Counting),” presented at the National Institute for Computer Assisted Reporting (NICAR) Conference, Jacksonville, FL, 2017, <http://slides.com/stevenrich/allthefoias#/>.
64. Kittur et al., “The Future of Crowd Work”; Daniel Weld et al., “Artificial Intelligence and Collective Intelligence,” presented at the Association for the Advancement of Artificial Intelligence (AAAI) Workshop on Human Computation, San Francisco, CA, 2011.
65. John Micklethwait, “The Future of News,” Bloomberg, May 3, 2018, <https://www.bloomberg.com/news/articles/2018-05-03/john-micklethwait-the-future-of-news>.

2. JOURNALISTIC DATA MINING

1. Carrie Teegardin et al., “Doctors & Sex Abuse,” *Atlanta Journal Constitution*, July 6, 2016, http://doctors.ajc.com/doctors_sex_abuse.
2. “How the Doctors & Sex Abuse Project Came About,” *Atlanta Journal Constitution*, 2016, http://doctors.ajc.com/about_this_investigation/.
3. Jeff Ernsthansen, “Doctors & Sex Abuse,” presented at the National Institute for Computer Assisted Reporting (NICAR) Conference, Jacksonville, FL, 2017, https://docs.google.com/presentation/d/1keGeDk_wpBPQgUOOhbRarPPFbyCculTObGLeAhOMmEM/edit
4. Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, “The KDD Process for Extracting Useful Knowledge from Volumes of Data,” *Communications of the ACM* 39, no. 11 (1996); Mehmed Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, 2nd ed. (Hoboken, NJ: Wiley, 2011).
5. Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth, “From Data Mining to Knowledge Discovery in Databases,” *AI Magazine* 17, no. 3 (1996).
6. Mary Branscombe, “Artificial Intelligence’s Next Big Step: Reinforcement Learning,” *New Stack*, January 2017, <https://thenewstack.io/reinforcement-learning-ready-real-world/>.
7. Zorroa, <https://zorroa.com/core-features/>, accessed October 28, 2018, and Snapstream, <https://www>

[.snapstream.com/](http://snapstream.com/), accessed October 28, 2018 are both companies that offer such archive tagging services to media companies.

8. Martha Stone, *Big Data for Media* (Oxford: Reuters Institute for the Study of Journalism, 2014), <https://reutersinstitute.politics.ox.ac.uk/our-research/big-data-media>.

9. Matt Lindsay and Patrick Tornabene, “Newsday Data Analytics Boost Subscriber Retention,” *Ideas Blog* (blog), September 2016, <http://www.inma.org/blogs/ideas/post.cfm/newsday-data-analytics-boost-subscriber-retention>.

10. Bob Kovach and Tom Rosenstiel, *The Elements of Journalism*, 3rd ed. (New York: Three Rivers Press, 2014).

11. Pamela Shoemaker and Tim Vos, *Gatekeeping Theory* (New York: Routledge, 2009); Pamela Shoemaker and Stephen Reese, *Mediating the Message in the 21st Century: A Media Sociology Perspective* (New York: Routledge, 2013). For additional background on the economic forces that shape news content, see Robert Picard, “The Economics of Journalism and News Provision,” in *Journalism (Handbooks of Communication Science)*, ed. Tim Vos (Berlin, Boston: Mouton De Gruyter, 2018).

12. Tony Harcup and Deirdre O’Neill, “What Is News?: News Values Revisited (Again),” *Journalism Studies* 18, no. 12 (2017).

13. Franziska Badenschier and Holger Wormer, “Issue Selection in Science Journalism: Towards a Special Theory of News Values for Science News?,” in *The Sciences’ Media Connection—Public Communication and Its Repercussions. Sociology of the Sciences Yearbook*, ed. Simone Rödder, Martina Franzen, and Peter Weingart (Dordrecht: Springer Netherlands, 2011).

14. Heidi Blake and John Templon, “The Tennis Racket,” *BuzzFeed News*, January 2016, <https://www.buzzfeed.com/heidiblake/the-tennis-racket>. For a critique of the methods used, see Nicholas Diakopoulos, “Enabling Accountability of Algorithmic Media: Transparency as a Constructive and Critical Lens,” in *Towards Glass-Box Data Mining for Big and Small Data*, ed. Tania Cerquitelli, Daniele Quercia, and Frank Pasquale (Cham, Switzerland: Springer, 2017). Another example of statistical anomaly detection (Monte Carlo simulation) to find a story is in Susan Pulliam and Rob Barry, “Executives’ Good luck in Trading Own Stock,” *Wall Street Journal*, November 27, 2012, <https://www.wsj.com/articles/SB10000872396390444100404577641463717344178>.

15. Matt Shearer, Basile Simon, and Clément Geiger, “Datastringer: Easy Dataset Monitoring for Journalists,” presented at the Computation + Journalism Symposium, New York, 2014.

16. Måns Magnusson, Jens Finnäs, and Leonard Wallentin, “Finding the News Lead in the Data Haystack: Automated Local Data Journalism Using Crime Data,” presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016. To try out the Newsworthy service, see <http://www.newsworthy.se/en/>, accessed October 28, 2018; for examples of output, see <https://www.europeandatajournalism.eu/eng/Tools-for-journalists/Stats-Monitor>, accessed October 28, 2018.

17. Titus Plattner, Didier Orel, and Olivier Steiner, “Flexible Data Scraping, Multi-Language Indexing, Entity Extraction and Taxonomies: Tadam, a Swiss Tool to Deal with Huge Amounts of Unstructured Data,” presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016.

18. Derek Willis, “Seeing the Real Campaign with ProPublica’s Election DataBot,” ProPublica, September 2016, <https://www.propublica.org/article/seeing-the-real-campaign-with-propublica-election-databot>.

19. Robert Faturechi and Derek Willis, “‘On Like Donkey Kong’: How a Dubious Super PAC Boosted a Questionable Penny Stock,” ProPublica, May 2016, <https://www.propublica.org/article/how-a-dubious-super-pac-boosted-a-questionable-penny-stock>.

20. William Perrin, “Local News Engine: Can the Machine Help Spot Diamonds in the Dust?” in *Data Journalism Past, Present, Future*, ed. John Mair, Richard Lance Keeble, and Megan Lucero (Bury St Edmunds, UK: Abramis, 2017).

21. Peter Aldhous, “BuzzFeed News Trained a Computer to Search for Hidden Spy Planes. This Is What We Found,” *BuzzFeed News*, August 2017, <https://www.buzzfeed.com/peteraldhous/hidden-spy-planes>.

22. Vlad Herasymenko et al. “Leprosy of the Land,” Texty.org.ua, March 2018, <http://texty.org.ua/d>

/2018/amber_eng/; for the methods for the project, see Amber Methodology, <https://github.com/texty/amber-methodology>.

23. Ryan J. Foley, Larry Fenn, and Nick Penzenstadler, “Chronicle of Agony: Gun Accidents Kill at Least 1 Kid Every Other Day,” *USA Today*, October 14, 2016, <https://www.usatoday.com/story/news/2016/10/14/ap-usa-today-gun-accidents-children/91906700/>.

24. William Lyon, “Combining the BuzzFeed Trumpworld Graph with Government Contracting Data in Neo4j,” *William Lyon* (blog), January 2017, <http://www.lyonwj.com/2017/01/30/trumpworld-us-contracting-data-neo4j/>.

25. Jonathan Stray, “Network Analysis in Journalism: Practices and Possibilities,” presented at the Knowledge Discovery and Data Mining (KDD) Workshop on Data Science + Journalism, Halifax, CA, 2017.

26. “The Russian Laundromat Exposed,” Organized Crime and Corruption Reporting Project (OCCRP), March 2017, <https://www.occrp.org/en/laundromat/the-russian-laundromat-exposed>.

27. The corpint library is open sourced and available here, <https://github.com/alephdata/corpint>, accessed August 7, 2018.

28. For an example of how Dedupe is interactively trained, see <https://dedupe.io/#demo>, accessed October 28, 2018.

29. Bedfellows, accessed August 7, 2018, <https://github.com/TheUpshot/Bedfellows/blob/master/introduction.md>.

30. Derek Willis, “Campaign Donations Reflect the Sharp Split in Congress among Republicans,” *ProPublica*, January 2016, <https://www.propublica.org/article/campaign-donations-reflect-the-sharp-split-in-congress-among-republicans>.

31. Joan Biskupic, Janet Roberts, and John Shiffman, “The Echo Chamber,” *Reuters*, December 8, 2014, <http://www.reuters.com/investigates/special-report/scotus/>.

32. Peter Tolmie et al., “Supporting the Use of User-Generated Content in Journalistic Practice,” *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM, 2017).

33. Matthew Brehmer et al., “Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool for Investigative Journalists,” *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 20, no. 12 (2014).

34. Overview News Stories, accessed August 7, 2018, <https://github.com/overview/overview-server/wiki/News-stories>.

35. Cristian Felix et al., “RevEx: Visual Investigative Journalism with a Million Healthcare Reviews,” presented at the Computation + Journalism Symposium, New York, 2015; Charles Ornstein, “On Yelp, Doctors Get Reviewed like Restaurants—And It Rankles,” *National Public Radio*, August 6, 2015, <https://www.npr.org/sections/health-shots/2015/08/06/429624187/on-yelp-doctors-get-reviewed-like-restaurants-and-it-rankles>.

36. Seid Muhie Yimam et al., “new / s / leak—Information Extraction and Visualization for Investigative Data Journalists,” presented at the Association for Computational Linguistics—Systems Demonstrations, Berlin, Germany, 2016.

37. Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo, “Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors,” *Proceedings of the World Wide Web Conference* (New York: ACM, 2010).

38. Neil Thurman, “Social Media, Surveillance, and News Work,” *Digital Journalism* 6, no. 1 (2018).

39. Xiaomo Liu et al., “Reuters Tracer: Toward Automated News Production Using Large Scale Social Media Data,” *Proceedings of the IEEE Big Data* (2017); Xiaomo Liu et al., “Reuters Tracer—A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter,” *Proceedings of the Conference on Information and Knowledge Management* (New York: ACM, 2016).

40. Tom Reilly (Cloudera) and Khalid Al-Kofahi (Thomson Reuters), “Becoming Smarter about Credible News,” presented at Strata + Hadoop World, San Francisco, CA, March 2017, <https://www.youtube.com/watch?v=1rwYOiWZmzc>.

41. "Timeliness Analysis of 31 Breaking News," accessed August 7, 2018, <https://docs.google.com/document/d/1iECb9y9aKRe0z58Fz0SdqhTOMEQU2pLZ4yQ6PUvgAw/edit#head>
42. Xiaomo Liu, et al., "Real-time Rumor Debunking on Twitter," *Proceedings of the International Conference on Information and Knowledge Management* (New York: ACM, 2015).
43. Raz Schwartz, Mor Naaman, and Ranny Theodoro, "Editorial Algorithms: Using Social Media to Discover and Report Local News," *Proceedings of the International Conference on Web and Social Media* (Palo Alto: The AAAI Press, 2015).
44. An example of video monitoring is Face-O-Matic, which can alert users when the faces of the president or congressional leaders appear on major TV news channels: Nancy Watzman, "Internet Archive TV News Lab: Introducing Face-O-Matic, Experimental Slack Alert System Tracking Trump & Congressional Leaders on TV News," *Internet Archive* (blog), July 19, 2017, <https://blog.archive.org/2017/07/19/introducing-face-o-matic>.
45. Kalev Leetaru and Philip A. Schrodt, "GDELT: Global Data on Events, Location, and Tone," presented at the International Studies Association meeting, San Francisco, CA, April 2013.
46. Konstantina Papanikolaou et al., "'Just the Facts' with PALOMAR: Detecting Protest Events in Media Outlets and Twitter," presented at the Workshop on Social Media in the Newsroom, Cologne, Germany, 2016; Elizabeth Boschee, Premkumar Natarajan, and Ralph Weischedel, "Automatic Extraction of Events from Open Source Text for Predictive Forecasting," in *Handbook of Computational Approaches to Counterterrorism*, ed. V. S. Subrahmanian (New York: Springer New York, 2012).
47. "Documenting Hate News Index," ProPublica, August 2017, <https://projects.propublica.org/hate-news-index/>; Simon Rogers, "A New Machine Learning App for Reporting on Hate in America," *Google* (blog), August 2017, <https://www.blog.google/topics/journalism-news/new-machine-learning-app-reporting-hate-america/>.
48. Katherine Keith et al., "Identifying Civilians Killed by Police with Distantly Supervised Entity-Event Extraction," *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Association for Computational Linguistics, 2017).
49. See, for example, "2017 NFL Predictions," FiveThirtyEight, accessed August 7, 2018, <https://projects.fivethirtyeight.com/2017-nfl-predictions/>; "2016 U.S. Open Predictions," FiveThirtyEight, accessed August 7, 2018, <https://projects.fivethirtyeight.com/2016-us-open-tennis-predictions/>; "CARMELO NBA Player Projections," FiveThirtyEight, accessed August 7, 2018, <https://projects.fivethirtyeight.com/carmelo/>.
50. Jeff Ernsthansen, "Predict-a-Bill: How It Works," *Atlanta Journal Constitution* (blog), February 2014, <http://ajcnewsapps.tumblr.com/post/77008124090/predict-a-bill-how-it-works>; Jeff Ernsthansen, "Predict-a-Bill Goes Real-Time (More or Less)," *Atlanta Journal Constitution* (blog), February 2015, <http://ajcnewsapps.tumblr.com/post/112039443036/predict-a-bill-goes-real-time-more-or-less>.
51. For a solid critique of this and related problems, see Nate Silver, "The Media Has a Probability Problem," FiveThirtyEight, September 2017, <https://fivethirtyeight.com/features/the-media-has-a-probability-problem/>.
52. For a treatment of how probabilistic horse race coverage may impact voter perceptions and turnout, see Sean Westwood, Solomon Messing, and Yphtach Lelkes, "Projecting Confidence: How the Probabilistic Horse Race Confuses and Demobilizes the Public," *Social Science Research Network (SSRN)*, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3117054.
53. Nate Silver, "A User's Guide to FiveThirtyEight's 2016 General Election Forecast," FiveThirtyEight, June 2016, <https://fivethirtyeight.com/features/a-users-guide-to-fivethirtyeights-2016-general-election-forecast/>.
54. Nate Silver, "The Real Story of 2016," FiveThirtyEight, January 2017, <http://fivethirtyeight.com/features/the-real-story-of-2016/>.
55. Tolmie et al., "Supporting the Use of User Generated Content in Journalistic Practice."
56. Steve Schifferes et al., "Identifying and Verifying News through Social Media," *Digital Journalism* 2, no. 3 (2014).

57. Ben Posten, Joel Rubin, and Anthony Pesce, "LAPD Underreported Serious Assaults, Skewing Crime Stats for 8 Years," *LA Times*, October 15, 2015, <http://www.latimes.com/local/cityhall/la-me-crime-stats-20151015-story.html>.
58. Freia Nahser, "Three Examples of Machine Learning in the Newsroom," *Global Editors Network* (blog), March 2018, <https://medium.com/global-editors-network/three-examples-of-machine-learning-in-the-newsroom-1b47d1f7515a>.
59. Alfred Hermida, Seth C. Lewis, and Rodrigo Zamith, "Sourcing the Arab Spring: A Case Study of Andy Carvin's Sources on Twitter during the Tunisian and Egyptian Revolutions," *Journal of Computer-Mediated Communication*, 19, no. 3 (2014).
60. For details on the development of one score that considers Twitter activity of a user, see Richard Fletcher, Steve Schifferes, and Neil Thurman, "Building the 'Truthmeter': Training Algorithms to Help Journalists Assess the Credibility of Social Media Sources," *Convergence: The International Journal of Research into New Media Technologies*, published ahead of print (2017).
61. Nicholas Diakopoulos, Munmun De Choudhury, and Mor Naaman, "Finding and Assessing Social Media Information Sources in the Context of Journalism," *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM, 2012).
62. Rui Fang et al., "Witness Identification in Twitter," presented at the Fourth International Workshop on Natural Language Processing for Social Media, Austin, TX, 2016.
63. Claire Wardle, "Fake News. It's Complicated," *First Draft News* (blog), February 2017, <https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79>.
64. Arkaitz Zubiaga et al., "Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads," *PloS one* 11, no. 3 (2016).
65. Klas Backholm et al., "Crises, Rumours and Reposts: Journalists' Social Media Content Gathering and Verification Practices in Breaking News Situations," *Media and Communication* 5, no. 2 (2017).
66. Carlos Castillo, Marcelo Mendoza, Barbara Poblete, "Information Credibility on Twitter," *Proceedings of the World Wide Web Conference* (New York: ACM, 2011).
67. Christina Boididou et al., "Learning to Detect Misleading Content on Twitter," *Proceedings of the International Conference on Multimedia Retrieval* (New York: ACM, 2017).
68. The efforts are part of an EU-funded project called "InVid." See "InVID Verification Plugin Open Beta Release," accessed August 7, 2018, <http://www.invid-project.eu/invid-verification-plugin-open-beta-release/>.
69. See the MediaEval Benchmarking Initiative for Multimedia evaluation <http://www.multimediaeval.org/>, accessed October 28, 2018; for a summary of RumourEval results from 2017, see Leon Derczynski et al., "SemEval-2017 Task 8: RumourEval: Determining Rumour Veracity and Support for Rumours," presented at the Workshop on Semantic Evaluation, Vancouver, Canada, 2017, <https://arxiv.org/abs/1704.05972v1>.
70. Sarah Harrison Smith, *The Fact Checker's Bible: A Guide to Getting It Right* (New York: Anchor, 2004); Lucinda Southern, "Inside Spiegel's 70-Person Fact-Checking Team," *Digiday*, August 2017, <https://digiday.com/media/inside-spiegels-70-person-fact-checking-team/>.
71. Elizabeth Jensen, "Behind the Scenes: NPR Fact Checks First Debate in Near Real Time," National Public Radio, October 3, 2016, <http://www.npr.org/sections/ombudsman/2016/09/27/495654679/behind-the-scenes-npr-fact-checks-first-debate-in-near-real-time>; Shan Wang, "How NPR Factchecked the First Presidential Debate in Realtime, on Top of a Live Transcript," Nieman Lab, September 2016, <http://www.niemanlab.org/2016/09/how-npr-factchecked-the-first-presidential-debate-in-realtime-on-top-of-a-live-transcript/>.
72. Brooks Jackson and Kathleen Hall Jamieson, *unSpun: Finding Facts in a World of Disinformation* (New York: Random House, 2007).
73. Mevar Babakar and Will Moy, "The State of Automated Factchecking," *FullFact*, August 2016, https://fullfact.org/media/uploads/full_fact-the_state_of_automated_factchecking_aug_2016.pdf.
74. Nicholas Diakopoulos, "Scaling Fact-Checking," *Tow-Knight Center for Entrepreneurial Journalism*

(blog), June 2012, <http://towknight.org/research/thinking/scaling-fact-checking/>. See also a summary of automated fact checking initiatives which reinforces the need for a hybrid approach in Lucas Graves, *Understanding the Promise and Limits of Automated Fact-Checking* (Oxford, UK: Reuters Institute for the Study of Journalism, 2018).

75. Naeemul Hassan et al., “Toward Automated Fact-Checking: Detecting Check-Worthy Factual Claims by ClaimBuster,” *Proceedings of the Conference on Knowledge Discovery and Datamining* (New York: ACM, 2017).

76. ClaimBuster is available online and you can enter your own text to score it: <http://idir-server2.uta.edu/claimbuster/demo>, accessed October 28, 2018.

77. Daniel Funke, “This Washington Post Fact Check Was Chosen by a Bot,” Poynter, January 2018, <https://www.poynter.org/news/washington-post-fact-check-was-chosen-bot>.

78. Lev Konstantinovskiy et al. “Towards Automated Factchecking: Developing an Annotation Schema and Benchmark for Consistent Automated Claim Detection,” presented at The First Workshop on Fact Extraction and Verification (FEVER), Brussels, Belgium, 2018.

79. For demos, see Will Moy and Mevan Babakar, “Automated Factchecking in the Era of Fake News,” presented at PyData, London, 2017, <https://www.youtube.com/watch?v=VPtfezx8WSg>. See also Daniel Funke, “In a Step toward Automation, Full Fact Has Built a Live Fact-Checking Prototype,” Poynter, November, 2017, <https://www.poynter.org/news/step-toward-automation-full-fact-has-built-live-fact-checking-prototype>.

80. James Thorne and Andreas Vlachos, “An Extensible Framework for Verification of Numerical Claims,” presented at the European Chapter of the Association for Computational Linguistics, Valencia, Spain, 2017.

81. Nathalie Jomini Stroud, Joshua Scacco, Alexander Curry, “The Presence and Use of Interactive Features on News Websites,” *Digital Journalism* 4, no. 3 (2016).

82. Nicholas Diakopoulos and Mor Naaman, “Towards Quality Discourse in Online News Comments,” *Proceedings of the Conference on Computer Supported Cooperative Work* (New York: ACM, 2011).

83. Ashley A. Anderson et al., “The ‘Nasty Effect’: Online Incivility and Risk Perceptions of Emerging Technologies,” *Journal of Computer-Mediated Communication* 19, no. 3 (2014).

84. AutoModerator, Reddit, accessed August 7, 2018, <https://www.reddit.com/wiki/automoderator>.

85. For more details on the performance of the system, see Shailesh Prakash, “Journalism & Technology: Big Data, Personalization & Automation,” keynote presented at the Computation + Journalism Symposium, Evanston, IL, 2017, <https://youtu.be/PqMvxo89AQ4?list=PL3kjMDQ990L556im9F1qozenDo2rMpcCH&t=2737>.

86. Jillian York, “Google’s Anti-Bullying AI Mistakes Civility for Decency,” Motherboard, August 2017, https://motherboard.vice.com/en_us/article/qvvv3p/googles-anti-bullying-ai-mistakes-civility-for-decency.

87. Nicholas Diakopoulos, “The Editor’s Eye: Curation and Comment Relevance on the New York Times,” *Proceedings of the Conference on Computer Supported Cooperative Work* (New York: ACM, 2015).

88. Nicholas Diakopoulos, “Picking the NYT Picks: Editorial Criteria and Automation in the Curation of Online News Comments,” *ISOJ Journal* 6, no. 1 (2016).

89. Deok Gun Park et al., “Supporting Comment Moderators in Identifying High Quality Online News Comments,” *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM, 2016).

90. James T. Hamilton, *Democracy’s Detectives* (Cambridge, MA: Harvard University Press, 2016).

91. For statistics on the average time spent on an investigative report, see Hamilton, *Democracy’s Detectives*, 159.

92. Ariana Giorgi, “An Analysis of Methods for Information Retrieval,” MA thesis, Columbia Journalism School, 2015, https://github.com/arianagiorgi/masters-proj/blob/master/Giorgi_MP2015.pdf.

93. Meredith Broussard, “Artificial Intelligence for Investigative Reporting,” *Digital Journalism* 3, no. 6

(2015).

94. Shoemaker and Vos, *Gatekeeping Theory*.

95. François Heinderyckx and Tim Vos, “Reformed Gatekeeping,” *Communication and Media* 11, no. 38 (2016).

96. Shoemaker and Vos, *Gatekeeping Theory*.

97. Nicholas Diakopoulos, “Diversity in the Robot Reporter Newsroom,” *Tow Center for Digital Journalism* (blog), July 2014, <https://towcenter.org/diversity-in-the-robot-reporter-newsroom/>.

98. See for instance the ClaimRank system, which is trained on and can be adapted based on data from nine different fact-checking organizations: Israa Jaradat et al., “ClaimRank: Detecting Check-Worthy Claims in Arabic and English,” presented at the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) Demonstrations, New Orleans, LA, 2018.

99. Eirik Stavelin, “Computational Journalism. When Journalism Meets Programming,” PhD diss., University of Bergen, 2014, http://stavelin.com/uib/ComputationalJournalism_EirikStavelin.pdf.

100. Neil Thurman et al., “Giving Computers a Nose for News,” *Digital Journalism* 4, no. 7 (2016).

101. Jens Finnäs and Måns Magnusson, “Marple: ‘We Tried to Automate Story Finding in Data—This Is What We Learnt,’ ” *Data Driven Journalism*, August 2017, http://datadrivenjournalism.net/featured_projects/marple_we_tried_to_automate_story_finding_in_data_this_is_what_we_learn#.WZLWkcDXK9Y.twitter.

102. Schwartz, Naaman, and Theodoro, “Editorial Algorithms: Using Social Media to Discover and Report Local News”; and Thurman, “Social Media, Surveillance, and News Work.”

103. Hamilton, *Democracy’s Detectives*.

104. For a survey, while not applied in the domain of journalism, see Richard J. Bolton and David J. Hand, “Statistical Fraud Detection: A Review,” *Statistical Science* 17, no. 3 (2002).

105. For a review of the evolution in journalism, see Stephen Ward, “Epistemologies of Journalism,” in *Journalism (Handbooks of Communication Science)*, ed. Tim Vos (Berlin, Boston: Mouton De Gruyter, 2018).

106. Seth Lewis and Oscar Westlund, “Big Data and Journalism,” *Digital Journalism* 3, no. 3 (2015); Sylvain Parasié, “Data-Driven Revelation?” *Digital Journalism* 3, no. 3 (2015).

107. James Ettema and Theodore Glasser, *Custodians of Conscience: Investigative Journalism and Public Virtue* (New York: Columbia University Press, 1998).

108. For an example of a public-facing decision tree in journalism, see Stefano Ceccon, “The Election Explained by the Times Data Team,” *Digital Times* (blog), May 18, 2015, <https://medium.com/digital-times/the-election-explained-by-the-times-data-team-482d3ab95f5c>.

109. Jennifer Stark and Nicholas Diakopoulos, “Towards Editorial Transparency in Computational Journalism,” presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016.

110. For a more thorough treatment of “the dial” and its relationship to the understanding of certainty and uncertainty in journalism, see C. W. Anderson, *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford: Oxford University Press, 2018).

3. AUTOMATED CONTENT PRODUCTION

1. Neil Thurman, Konstantin Dörr, and Jessica Kunert, “When Reporters Get Hands-On with Robo-Writing: Professionals Consider Automated Journalism’s Capabilities and Consequences,” *Digital Journalism* 5, no. 10 (2017).

2. Dongwan Kim, “An Algorithmic Approach to Personalized and Interactive News Generation,” PhD diss., Seoul National University, 2017, <http://s-space.snu.ac.kr/bitstream/10371/120411/1/000000141346.pdf>.

3. Ehud Reiter and Robert Dale, *Building Natural Language Generation Systems* (Cambridge, UK: Cambridge University Press, 2000); Ehud Reiter, “Natural Language Generation,” in *The Handbook of Computational Linguistics and Natural Language Processing*, ed. Alexander Clark, Chris Fox, and Shalom

Lappin (Oxford, UK: Wiley-Blackwell, 2010); Amanda Stent and Srinivas Bangalore, eds., *Natural Language Generation in Interactive Systems* (New York: Cambridge University Press, 2014).

4. Eli Goldberg, Norbert Driedger, and Richard Kittredge, "Using Natural-Language Processing to Produce Weather Forecasts," *IEEE Expert* 9, no. 2 (1994).

5. Leo Leppänen et al., "Finding and Expressing News From Structured Data," presented at AcademicMindtrek, Helsinki, Finland, 2017.

6. Bastian Haarmann and Lukas Sikorski, "Natural Language News Generation from Big Data," *International Journal of Computer, Electrical, Automation, Control and Information Engineering* 9, no. 6 (2015).

7. Nicholas D. Allen et al., "StatsMonkey: A Data-Driven Sports Narrative Writer," presented at the Association for the Advancement of Artificial Intelligence (AAAI) Fall Symposium, Arlington, VA, 2010.

8. Lawrence A. Birnbaum et al., System and Method for Using Data to Automatically Generate a Narrative Story, US Patent, US8688434B1, filed May 13, 2010, and issued April 1, 2014, <https://www.google.com/patents/US8688434>.

9. Espen Waldal, "Building a Robot Journalist," *B&B Stories* (blog), November 18, 2016, <https://medium.com/bakken-b%C3%A6ck/building-a-robot-journalist-171554a68fa8>.

10. Albert Gatt and Ehud Reiter, "SimpleNLG: A Realisation Engine for Practical Applications," presented at the European Workshop on Natural Language Generation, Athens, Greece, 2009; Al Johri, Eui-Hong Han, and Dhruvil Mehta, "Domain Specific Newsbots," presented at the Computation + Journalism Symposium, Palo, Alto, CA, 2016.

11. Regina Barzilay and Mirella Lapata, "Collective Content Selection for Concept-to-Text Generation," *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing* (Stroudsburg, PA: Association for Computational Linguistics, 2005).

12. Joy Mahapatra, Sudip Naskar, and Sivaji Bandyopadhyay, "Statistical Natural Language Generation from Tabular Non-Textual Data," *Proceedings of the International Natural Language Generation Conference* (2016).

13. Ravi Kondadadi, Blake Howald, and Frank Schilder, "A Statistical NLG Framework for Aggregated Planning and Realization," *Proceedings of the Association for Computational Linguistics* (Stroudsburg, PA: Association for Computational Linguistics, 2013).

14. Freddy Chong Tat Chua and Sitaram Asur, "Automatic Summarization of Events from Social Media," *Proceedings of the Conference on Weblogs and Social Media* (Palo Alto: The AAAI Press, 2013); Philip J. McParlane, Andrew J. McMinn, Joemon M. Jose, "Picture the Scene ...:Visually Summarising Social Media Events," *Proceedings of the International Conference on Information and Knowledge Management* (New York: ACM, 2014); Jeffrey Nichols, Jalal Mahmud, and Clemens Drews, "Summarizing Sporting Events Using Twitter," *Proceedings of the Conference on Intelligent User Interfaces* (New York: ACM, 2012).

15. Elena Lloret and Manuel Palomar, "Towards Automatic Tweet Generation: A Comparative Study from the Text Summarization Perspective in the Journalism Genre," *Expert Systems with Applications* 40, no. 16 (2013); Kristian Woodsend and Mirella Lapata, "Automatic Generation of Story Highlights," *Proceedings of the Annual Meeting of the Association for Computational Linguistics* (Stroudsburg, PA: Association for Computational Linguistics, 2010); Bonnie Dorr, David Zajic, and Richard Schwartz, "Hedge Trimmer: A Parse-and-Trim Approach to Headline Generation," (presented at the Human Language Technologies - North American Chapter of the Association for Computational Linguistics (HLT-NAACL) Workshop on Text Summarization, Edmonton, Canada, 2003; Yonatan Oren, "Flipboard's Approach to Automatic Summarization," *Flipboard* (blog), October 2014, <https://about.flipboard.com/engineering/flipboards-approach-to-automatic-summarization/>.

16. Ani Nenkova and Kathleen McKeown, "A Survey of Text Summarization Techniques," in *Mining Text Data*, ed. Charu C. Aggarwal and ChengXiang Zhai (Boston: Springer, 2012).

17. Stefano Bocconi, Frank Nack, and Lynda Hardman, "Automatic Generation of Matter-of-Opinion Video Documentaries," *Web Semantics: Science, Services and Agents on the World Wide Web* 6, no. 2

(2008); Diogo Delgado, Joao Magalhaes, and Nuno Correia, “Assisted News Reading with Automated Illustration,” *Proceedings of the International Conference on Multimedia* (New York: ACM, 2010); Jack Jansen et al., “Just-in-Time Personalized Video Presentations,” *Proceedings of the Symposium on Document Engineering* (New York: ACM, 2012); Nate Nichols and Kristian Hammond, “Machine-Generated Multimedia Content,” presented at the Conference on Advances in Computer-Human Interactions, Washington, DC, 2009.

18. Nathalie Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale, eds. *Data-Driven Storytelling* (Boca Raton, FL: CRC Press, 2018).

19. Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar, “Contextifier: Automatic Generation of Annotated Stock Visualizations,” *Proceedings of the Conference on Human Factors in Computing Systems [CHI]* (New York: ACM, 2013).

20. Stephen Feiner and Kathleen McKeown, “Automating the Generation of Coordinated Multimedia Explanations,” *Computer* 24 no. 10 (1991); Nancy Green et al., “Autobrief—An Experimental System for the Automatic Generation of Briefings in Integrated Text and Information Graphics,” *International Journal of Human-Computer Studies* 61, no. 1 (2004); Rola Alhalaseh et al., “Towards Data-Driven Generation of Visualizations for Automatically Generated News Articles” presented at Academic Mindtrek, Tampere, Finland, 2018.

21. Vibhu Mittal et al., “Describing Complex Charts in Natural Language: A Caption Generation System,” *Computational Linguistics* 24, no. 3 (1998).

22. Hullman, Diakopoulos, and Adar, “Contextifier”; Tong Gao et al., “NewsViews: An Automated Pipeline for Creating Custom Geovisualizations for News,” *Proceedings of the Conference on Human Factors in Computing* (New York: ACM, 2014).

23. Patrick Stotz and Danial Montazeri, “Taktiktafeln: You’ve Never Seen Soccer Like This Before!” (“SPON-Taktiktafeln—So Haben Sie Fußball Noch Nie Gesehen!”) *Spiegel Online*, January 27, 2017, <http://www.spiegel.de/sport/fussball/bundesliga-taktiktafeln-paesse-und-formationen-der-bundesligateams-a-1131647.html>. Reuters has experimented with automatically including visualizations produced by Graphiq, see: Lora Kolodny, “Reuters Is the Latest Large News Agency to Embrace Content Automation,” *Tech Crunch*, September 13, 2016, <https://techcrunch.com/2016/09/13/reuters-is-the-latest-large-news-agency-to-embrace-content-automation/>.

24. For an example of voice generation see Vocal Avatar, Lyrebird, accessed August 9, 2018, <https://lyrebird.ai/demo>.

25. Mattias Niessner, “Face2Face: Real-Time Face Capture and Reenactment of RGB Videos,” YouTube, March 2016, <https://www.youtube.com/watch?v=ohmajJTcpNk>.

26. Kevin Roose, “Here Come the Fake Videos, Too,” *New York Times*, March 4, 2018, <https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html>.

27. See video linked here: Karen Hao, “Researchers Have Figured Out How to Fake News Video with AI,” *Quartz*, July 2017, <https://qz.com/1031624/researchers-have-figured-out-how-to-fake-news-video-with-ai/>; Supasorn Suwajanakorn, Steven Seitz, and Ira Kemelmacher-Shlizerman, “Synthesizing Obama: Learning Lip Sync From Audio,” *ACM Transactions On Graphics* 36, no. 4 (2017).

28. Tero Karras et al., “Progressive Growing of GANs for Improved Quality, Stability, and Variation,” presented at the International Conference on Learning Representations, Vancouver, Canada, 2018.

29. Yuanshun Yao, “Automated Crowdturfing Attacks and Defenses in Online Review Systems,” *Proceedings of Computer and Communications Security* (New York: ACM, 2017).

30. James Vincent, “Artificial Intelligence Is Going to Make It Easier than Ever to Fake Images and Video,” *Verge*, December 20, 2016, <https://www.theverge.com/2016/12/20/14022958/ai-image-manipulation-creation-fakes-audio-video>.

31. Valentina Conotter et al., “Physiologically-Based Detection of Computer Generated Faces in Video,” *Proceedings of the IEEE International Conference on Image Processing* (2014); Andreas Rössle et al., “FaceForensics: A Large-Scale Video Dataset for Forgery Detection in Human Faces,” ArXiv pre-print, <https://arxiv.org/abs/1803.09179>. See also Hany Farid, “How to Detect Faked Photos,” *American Scientist*

105, no. 2 (March–April, 2017).

32. Goldberg, Driedger, and Kittredge, “Using Natural-Language Processing to Produce Weather Forecasts.”

33. “Robots at ‘Le Monde’ During the Departmental Elections? Yes and No” (“Des robots au ‘Monde’ pendant les élections départementales? Oui ... et non”) *Back Office* (blog), March 23, 2015, <http://makingof.blog.lemonde.fr/2015/03/23/des-robots-au-monde-pendant-les-elections-departementales-oui-et-non/>; Johanna Vehkoo, “News Robots Finally Arrived in Finland,” “Uutisrobotit Tulivat Vihdoin Suomeen,” *Journalisti*, March 2017, <https://www.journalisti.fi/artikkelit/2017/4/uutisrobotit-tulivat-vihdoin-suomeen/>; Leo Leppänen et al., “Data-Driven News Generation for Automated Journalism,” *Proceedings of the International Conference on Natural Language Generation* (Stroudsburg, PA: Association for Computational Linguistics, 2017).

34. Andreas Graefe, *Computational Campaign Coverage* (New York: Tow Center for Digital Journalism, 2017), https://www.cjr.org/tow_center_reports/computational-campaign-coverage.php/.

35. Alexander Fanta, *Putting Europe’s Robots on the Map: Automated Journalism in News Agencies* (Oxford: Reuters Institute for the Study of Journalism, 2017), <https://reutersinstitute.politics.ox.ac.uk/our-research/putting-europes-robots-map-automated-journalism-news-agencies>.

36. Allen et al., “StatsMonkey”; see <http://statsmonkey.infolab.northwestern.edu/baseball/> for sample output. For information on Yle’s open source hockey story writer, see <https://github.com/Yleisradio/avoivoitto>, accessed October 31, 2018; Rich McCormick, “AP’s ‘Robot Journalists’ Are Writing about Minor League Baseball Now,” *Verge*, July 4, 2016, <https://www.theverge.com/2016/7/4/12092768/ap-robot-journalists-automated-insights-minor-league-baseball>.

37. “The Washington Post to Use Artificial Intelligence to Cover Nearly 500 Races on Election Day,” *Washington Post*, October 19, 2016, <https://www.washingtonpost.com/pr/wp/2016/10/19/the-washington-post-uses-artificial-intelligence-to-cover-nearly-500-races-on-election-day/>; “The Washington Post Leverages Automated Storytelling to Cover High School Football,” *Washington Post*, September 1, 2017, <https://www.washingtonpost.com/pr/wp/2017/09/01/the-washington-post-leverages-heliograf-to-cover-high-school-football/>.

38. Scott Klein, “How to Edit 52,000 Stories at Once,” *ProPublica*, January 2013, <https://www.propublica.org/nerds/item/how-to-edit-52000-stories-at-once>; Mary Lynn Young and Alfred Hermida, “From Mr. and Mrs. Outlier to Central Tendencies,” *Digital Journalism* 3, no. 3 (2014); Will Oremus, “The First News Report on the L.A. Earthquake Was Written by a Robot,” *Slate*, March 17, 2014, http://www.slate.com/blogs/future_tense/2014/03/17/quakebot_los_angeles_times_robot_journalist_writes_article_on_la_earthquake.html.

39. Li L’Estrade, “MittMedia Homeowners Bot Boosts Digital Subscriptions with Automated Articles,” *INMA* (blog), June 2018, <https://www.inma.org/blogs/ideas/post.cfm/mittmedia-homeowners-bot-boosts-digital-subscriptions-with-automated-articles>.

40. Shuguang Wang, Eui-Hong (Sam) Han, and Alexander Rush, “Headliner: An Integrated Headline Suggestion System,” presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016.

41. Rob Hotakainen, Lindsay Wise, Frank Matt, and Samantha Ehlinger, “Irradiated,” *McClatchy*, 2016, <http://media.mcclatchydc.com/static/features/irradiated/>.

42. Andreas Graefe, *Guide to Automated Journalism* (New York: Tow Center for Digital Journalism, 2016), https://www.cjr.org/tow_center_reports/guide_to_automated_journalism.php/.

43. Nikki Usher, “Breaking News Production Processes in US Metropolitan Newspapers: Immediacy and Journalistic Authority,” *Journalism* 19, no. 1 (2017); and Justin Lewis and Stephen Cushion, “The Thirst to Be First,” *Journalism Practice* 3, no. 3 (2009).

44. Gaye Tuchman, “Making News by Doing Work: Routinizing the Unexpected,” *American Journal of Sociology*, 79, no. 1 (1973).

45. Jonathan Bright, “The Social News Gap: How News Reading and News Sharing Diverge,” *Journal of Communication* 66, no. 3 (2016).

46. Matt Carlson, “Automated Journalism: A Posthuman Future for Digital News?” in *The Routledge*

Companion to Digital Journalism Studies, ed. Bob Franklin and Scott Eldridge II (London: Routledge, 2016).

47. Elizabeth Blankespoor, Ed deHaan, and Christina Zhu, "Capital Market Effects of Media Synthesis and Dissemination: Evidence from Robo-Journalism," *Review of Accounting Studies* 21, no. 2 (2016).

48. Oremus, "The First News Report on the L.A. Earthquake."

49. Graefe, *Computational Campaign Coverage*.

50. Ben Ashwell, "How Automated Financial News Is Changing Quarterly Earnings Coverage," *IR Magazine*, June 2018, <https://www.irmagazine.com/reporting/how-automated-financial-news-changing-quarterly-earnings-coverage>.

51. Daniel Billsus and Michael Pazzani, "Adaptive News Access," in *The Adaptive Web: Lecture Notes in Computer Science* (Heidelberg: Springer, 2007); Neil Thurman, "Making 'The Daily Me': Technology, Economics, and Habit in the Mainstream Assimilation of Personalized News," *Journalism* 12, no. 4 (2011).

52. Krishna Bharat, Tomonari Kamba, and Michael Albers, "Personalized, Interactive News on the Web," *Multimedia Systems* 6, no. 5 (1998).

53. Eytan Adar et al., "PersaLog: Personalization of News Article Content," *Proceedings of Human Factors in Computing Systems* (New York: ACM, 2017).

54. Billsus and Pazzani, "Adaptive News Access."

55. Kim, "An Algorithmic Approach to Personalized and Interactive News Generation."

56. Somayajulu Sripada et al., "A Case Study: NLG Meeting Weather Industry Demand for Quality and Quantity of Textual Weather Forecasts," *Proceedings of the International Conference on Natural Language Generation* (New York: ACM, 2014).

57. "Google Funds Automated News Project," *BBC News*, July 6, 2017, <http://www.bbc.com/news/technology-40517420>. See also the RADAR website: <http://radarai.com/>, accessed October 31, 2018.

58. David Sharman, "Four Journalists Join PA Robot Reporting Unit as Scheme Expands," *HoldTheFrontPage*, June 18, 2018, <https://www.holdthefrontpage.co.uk/2018/news/four-journalists-join-pa-robot-reporting-unit-as-scheme-expands/>.

59. Shan Wang, "This Hyperlocal News Site in San Francisco Is Reinventing Itself with an Automated Local News Wire," Nieman Lab, February 2018, <http://www.niemanlab.org/2018/02/this-hyperlocal-news-site-in-san-francisco-is-reinventing-itself-with-an-automated-local-news-wire/>. For an example article, see: "Explore These 5 New Montrose Businesses," February 2018, <http://abc13.com/food/explore-these-5-new-montrose-businesses/3039818/>.

60. Gregor Aisch et al., "Where the Poor Live Longer: How Your Area Compares," *New York Times*, April 11, 2016, <https://www.nytimes.com/interactive/2016/04/11/upshot/where-the-poor-live-longer-how-your-area-compares.html>; Gregor Aisch et al., "The Best and Worst Places to Grow Up: How Your Area Compares," *New York Times*, May 4, 2015, <https://www.nytimes.com/interactive/2015/05/03/upshot/the-best-and-worst-places-to-grow-up-how-your-area-compares.html>; Kevin Quealy and Margot Sanger-Katz, "The Experts Were Wrong about the Best Places for Better and Cheaper Health Care," *New York Times*, December 15, 2016, <https://www.nytimes.com/interactive/2015/12/15/upshot/the-best-places-for-better-cheaper-health-care-arent-what-experts-thought.html>; Ted Mellnik et al., "America's Great Housing Divide: Are You a Winner or Loser?" *Washington Post*, April 28, 2016, <https://www.washingtonpost.com/graphics/business/wonk/housing/overview/>; Sarah Kliff, Soo Oh, and Sarah Frostenson, "Today's Teens ... Less Than You Did," *Vox*, June 2016, <https://www.vox.com/a/teens>.

61. Lucia Moses, "The Washington Post's Robot Reporter Has Published 850 Articles in the Past Year," *Digiday*, September 2017, <https://digiday.com/media/washington-posts-robot-reporter-published-500-articles-last-year/>.

62. L'Estrade, "MittMedia Homeowners."

63. Thurman, Dörr, and Kunert, "When Reporters Get Hands-On with Robo-Writing."

64. Katherine Fink and C. W. Anderson, "Data Journalism in the United States," *Journalism Studies* 16, no. 4 (2014).

65. Fanta, *Putting Europe's Robots on the Map*.

66. Fergus Pitt, *Sensors and Journalism* (New York: Tow Center for Digital Journalism, 2014), <http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Sensors-and-Journalism.pdf>; Sarah Cohen, James T. Hamilton, and Fred Turner, “Computational Journalism,” *Communications of the ACM* 54, no. 10 (2011).
67. For instance, many of the quality heuristics available in “The Quartz Guide to Bad Data” required human attention; see <https://github.com/Quartz/bad-data-guide>, accessed October 31, 2018. For a framework for assessing data quality for automated journalism developed, see Laurence Dierickx, “News Bot for the Newsroom: How Building Data Quality Indicators Can Support Journalistic Projects Relying on Real-Time Open Data,” presented at the Global Investigative Journalism Conference, Johannesburg, South Africa, 2017, <http://ijec.org/2018/02/02/research-news-bot-for-the-newsroom-how-building-data-quality-indicators-can-support-journalistic-projects-relying-on-real-time-open-data/>.
68. David Caswell and Konstantin Dörr, “Automated Journalism 2.0: Event-Driven Narratives,” *Journalism Practice* 12, no. 4 (2018); David Caswell, “Computable News Ecosystems: Roles for Humans and Machines,” presented at the Second Workshop on Computing News Storylines, Austin, TX, 2016.
69. Matt Carlson, “The Robotic Reporter,” *Digital Journalism* 3, no. 3 (2015).
70. Tarleton Gillespie, “The Relevance of Algorithms,” in *Media Technologies: Essays on Communication, Materiality, and Society*, ed. Tarleton Gillespie, Pablo Boczkowski, and Kirstin Foot (Cambridge, MA: MIT Press, 2014). See also the case study of structured journalism presented in C. W. Anderson, *Apostles of Certainty: Data Journalism and the Politics of Doubt* (Oxford: Oxford University Press, 2018).
71. Carl-Gustav Linden, “Algorithms for Journalism: The Future of News Work,” *Journal of Media Innovations* 4, no. 1 (2017); Konstantin Dörr and Katharina Hollnbuchner, “Ethical Challenges of Algorithmic Journalism,” *Digital Journalism* 5, no. 4 (2017).
72. Aylin Caliskan, Joanna Bryson, and Arvind Narayanan, “Semantics Derived Automatically from Language Corpora Contain Human-Like Biases,” *Science*, 356, no. 6334 (2017).
73. Fink and Anderson, “Data Journalism in the United States”; Dörr and Hollnbuchner, “Ethical Challenges of Algorithmic Journalism.”
74. Cohen, Hamilton, and Turner, “Computational Journalism.”
75. Zvi Reich, “The Impact of Technology on News Reporting: A Longitudinal Perspective,” *Journalism & Mass Communication Quarterly* 90, no. 3 (2013).
76. Noam Lemelshtrich Latar, “The Robot Journalist in the Age of Social Physics: The End of Human Journalism?” in *The New World of Transitioned Media. The Economics of Information, Communication, and Entertainment*, ed. Gali Einav (Cham, Switzerland: Springer International Publishing, 2015).
77. Eric Horvitz, “Reflections on Challenges and Promises of Mixed-Initiative Interaction,” *AI Magazine*, Summer 2007.
78. Jeffrey Nichols and Jeon-Hyung Kang, “Asking Questions of Targeted Strangers on Social Networks,” *Proceedings of the Conference on Computer Supported Cooperative Work* (New York: ACM, 2012).
79. Dan Berkowitz, “Non-Routine News and Newswork: Exploring a What-a-Story,” *Journal of Communication* 42, no. 1 (1992).
80. Raja Parasuraman, Thomas Sheridan, and Christopher Wickens, “A Model for Types and Levels of Human Interaction with Automation,” *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30, no. 3 (2000).
81. Fanta, *Putting Europe’s Robots on the Map*.
82. Leppänen et al., “Data-Driven News Generation for Automated Journalism.”
83. Rong-Gong Lin II, “Revenge of Y2K? A Software Bug Might Have Caused False Alert for Big (and Very Old) Earthquake,” *LA Times*, June 22, 2017, <http://www.latimes.com/local/lanow/la-me-earthquakes-earthquake-68-quake-strikes-near-isla-vista-calif-jyhwh-tmlstory.html>.
84. Graefe, *Guide to Automated Journalism*.
85. Alexander Fanta, “Robots Can Save Local Journalism. But Will They Make It More Biased?” *Data Driven Journalism*, March 2018, http://datadrivenjournalism.net/news_and_analysis/robots_can_save_local

[_journalism._but_will_they_make_it_more_biased.](#)

86. Carlson, "Automated Journalism."

87. Nicholas Diakopoulos, "Algorithmic Defamation: The Case of the Shameless Autocomplete," *Tow Center* (blog), August 2013, <https://towcenter.org/algorithmic-defamation-the-case-of-the-shameless-autocomplete/>. For a definition of "defamation," see Digital Media Law Project, <http://www.dmlp.org/legal-guide/defamation>, accessed October 31, 2018.

88. For a more international perspective on the laws governing automatic speech, see Meg Leta Jones, "Silencing Bad Bots: Global, Legal and Political Questions for Mean Machine Communication," *Communication Law and Policy* 23, no. 2 (2018).

89. Meg Leta Ambrose and Ben Ambrose, "When Robots Lie: A Comparison of Auto-Defamation Law," presented at the IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), Evanston, IL, 2014.

90. Seth C. Lewis, Amy Kristin Sanders, and Casey Carmody, "Libel by Algorithm? Automated Journalism and the Threat of Legal Liability," *Journalism & Mass Communication Quarterly* 80, no. 1 (2018).

91. Dan Ring, "Natural Language Generation Software Turns Data into Plain English," *Tech Target*, September 2016, <http://searchfinancialapplications.techtarget.com/feature/Natural-language-generation-software-turns-data-into-plain-English>.

92. Andreas Graefe et al., "Readers' Perception of Computer-Generated News: Credibility, Expertise, and Readability," *Journalism* 7, no. 6 (2017); Christer Clerwall, "Enter the Robot Journalist," *Journalism Practice* 8, no. 5 (2014); Jaemin Jung et al., "Intrusion of Robots into Journalism: The Public's and Journalists' Perceptions of News Written by Robots and Human Journalists," *Computers in Human Behavior* 71 (2017); Hilla van der Kaa and Emiel Kraemer, "Journalist versus News Consumer: The Perceived Credibility of Machine Written News," presented at the Computation + Journalism Symposium, New York, 2014.

93. Clerwall, "Enter the Robot Journalist."

94. Graefe et al., "Readers' Perception of Computer-Generated News."

95. Anja Wölker and Thomas Powell, "Algorithms in the Newsroom? News Readers' Perceived Credibility and Selection of Automated Journalism," *Journalism* 70, no. 14 (2018).

96. Magnus Melin et al., "No Landslide for the Human Journalist - An Empirical Study of Computer-Generated Election News in Finland," *IEEE Access*, published ahead of print, (2018).

97. Romain Paulus, Caiming Xiong, and Richard Socher, "A Deep Reinforced Model for Abstractive Summarization," arXiv.org, 2017, <https://arxiv.org/abs/1705.04304>.

98. Fanta, *Putting Europe's Robots on the Map*.

99. "Artificial Intelligence, Automation, and the Economy," Executive Office of the President of the United States, 2016.

100. Konstantin Dörr, "Mapping the Field of Algorithmic Journalism," *Digital Journalism* 4, no. 6 (2016).

101. Ken Schwencke, "How to Break News while You Sleep," *Source*, March 24, 2014, <https://source.opennews.org/en-US/articles/how-break-news-while-you-sleep/>.

102. Bocconi, Nack, Hardman, "Automatic Generation of Matter-of-Opinion Video Documentaries."

103. Graefe, *Computational Campaign Coverage*.

104. Ben Shneiderman, et al., *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 6th ed. (Boston: Pearson, 2017); John Lee and Katrina See, "Trust in Automation: Designing for Appropriate Reliance," *Human Factors* 46 no. 1 (2016); Horvitz, "Reflections on Challenges and Promises of Mixed-Initiative Interaction"; Daniel Repperger and Chandler Phillips, "The Human Role in Automation," in *Springer Handbook of Automation*, ed. Shimon Y. Nof (Heidelberg: Springer, 2009).

105. "Artificial Intelligence, Automation, and the Economy."

106. Klein, "How to Edit 52,000 Stories at Once."

107. Sripada et al., "A Case Study: NLG Meeting Weather Industry Demand."

108. Franco Moretti, *Distant Reading* (London: Verso, 2013).
109. Adar et al., “PersaLog.”
110. Henrik Örnebring, “Technology and Journalism-as-Labour: Historical Perspectives,” *Journalism* 11, no. 1 (2010); Young and Hermida, “From Mr. and Mrs. Outlier to Central Tendencies.”
111. Hanna Tuulonen, “A Possibility, a Threat, a Denial? How News Robots Affect Journalists’ Work Practices and Professional Identity,” MA thesis, University of Gothenburg, 2017.
112. Caswell, “Computable News Ecosystems: Roles for Humans and Machines.”
113. Peter Bro, Kenneth Reinecke Hansen, and Ralf Andersson, “Improving Productivity in the Newsroom?” *Journalism Practice* 10, no. 8 (2016); Chang-de Liu, “De-Skilling Effects on Journalists: ICTs and the Labour Process of Taiwanese Newspaper Reporters,” *Canadian Journal of Communication* 31, no. 3 (2006).
114. Caswell and Dörr, “Automated Journalism 2.0.”
115. Robert O. Wyatt and David Badger, “A New Typology for Journalism and Mass Communication Writing,” *Journalism Educator*, Spring 1993.
116. Bocconi, Nack, and Hardman, “Automatic Generation of Matter-of-Opinion Video Documentaries”; Michael Mateas, Paul Vanouse, and Steffi Domike, “Generation of Ideologically-Biased Historical Documentaries,” *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence* (2000).
117. Stotz and Montazeri, “Taktiktafeln: You’ve Never Seen Soccer Like This Before!” (“SPON-Taktiktafeln—So Haben Sie Fußball Noch Nie Gesehen!”).

4. NEWSBOTS: AGENTS OF INFORMATION

1. Nicholas Diakopoulos, “Picking the NYT Picks: Editorial Criteria and Automation in the Curation of Online News Comments,” *#ISOJ Journal* 6, no. 1 (2015).
2. Nicholas Diakopoulos, “Enabling Accountability of Algorithmic Media: Transparency as a Constructive and Critical Lens,” in *Transparent Data Mining for Big and Small Data*, ed. Tania Cerquitelli, Daniele Quercia, and Frank Pasquale (Cham, Switzerland: Springer International Publishing, 2017).
3. AnecbotalNYT (@AnecbotalNYT), Twitter, January 22, 2017, <https://twitter.com/AnecbotalNYT/status/823351538198446080>.
4. Tim Hwang, Ian Pearce, and Max Nanis, “Socialbots: Voices from the Fronts,” *Interactions*, March / April, 2012; Timothy Graham and Robert Ackland, “Do Socialbots Dream of Popping the Filter Bubble?,” in *Socialbots and Their Friends: The Role of Socialbots in Promoting Deliberative Democracy in Social Media*, ed. Robert W. Gehl and Maria Bakardjieva (New York: Routledge, 2017).
5. Nicholas Confessore et al., “The Follower Factory,” *New York Times*, January 27, 2018, <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>.
6. Stefan Wojcik et al., *Bots in the Twittersphere* (Washington, DC: Pew Research Center, 2018), <http://www.pewinternet.org/2018/04/09/bots-in-the-twittersphere/>.
7. Onur Varol et al., “Online Human-Bot Interactions: Detection, Estimation, and Characterization,” *Proceedings of the International Conference on Web and Social Media* (Palo Alto: The AAAI Press, 2017).
8. Norah Abokhodair, Daisy Yoo, and David W. McDonald, “Dissecting a Social Botnet,” *Proceedings of the Conference on Computer Supported Cooperative Work and Social Computing* (New York: Association for Computing Machinery, 2015); Hwang, Pearce, and Nanis, “Socialbots”; Claudia Wagner et al., “When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks,” presented at the second Workshop on Making Sense of Microposts, Lyon, France, 2012; Robert W. Gehl and Maria Bakardjieva, eds., *Socialbots and their Friends* (New York: Routledge, 2017). For a thorough review of definitions of and distinctions between crawlers, chatbots, spambots, social bots, sock puppets, and cyborgs, see Robert Gorwa and Douglas Guilbeault, “Understanding Bots for Policy and Research: Challenges, Methods, and Solutions,” presented at the International Communication Association (ICA) Conference, Prague, Czechia, 2018.
9. Christian Grimme et al., “Social Bots: Human-Like by Means of Human Control?” *Big Data* 5, no. 4

(2017).

10. Robert Dale, "The Return of the Chatbots," *Natural Language Engineering* 22, no. 5 (2016).
11. "Bad Bot Report," Distil Networks, 2017, <https://resources.distilnetworks.com/white-paper-reports/2017-bad-bot-report>.
12. Emilio Ferrara, et al., "The Rise of Social Bots," *Communications of the ACM* 59, no. 7 (2016).
13. Ewa Luger and Abigail Sellen, "Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents," *Proceedings of the Conference on Human Factors in Computing Systems [CHI]* (New York: Association for Computing Machinery, 2016).
14. Kedar Dhamdhere, "Analyza—Exploring Data with Conversation," *Proceedings of International Conference on Intelligent User Interfaces* (New York: Association for Computing Machinery, 2017).
15. Amanda Stent and Srivinas Bangalore, "Introduction," in *Natural Language Generation in Interactive Systems*, eds. Amanda Stent and Srivinas Bangalore (Cambridge, UK: Cambridge University Press, 2014); Dale, "The Return of the Chatbots"; Justine Cassell, "Embodied Conversational Interface Agents," *Communications of the ACM* 43, no. 4 (2000); Robert Moore et al., "Conversational UX Design," presented at the Conversational UX Design CHI Workshop, Denver, CO, 2017.
16. Al Johri, Eui-Hong (Sam) Han, and Dhruvil Mehta, "Domain Specific Newsbots," presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016.
17. Marisa Vasconcelos, Heloisa Candello, and Claudio Pinhanez, "Bottester: Testing Conversational Systems with Simulated Users," presented at the Conversational UX Design CHI Workshop, Denver, CO, 2017.
18. For a light treatment, see Alex Hern, "Please, Facebook, Don't Make Me Speak to Your Awful Chatbots," April 29, 2016, *Guardian*, <https://www.theguardian.com/technology/2016/apr/29/please-facebook-dont-make-me-speak-to-your-awful-chatbots>. For a more rigorous comparison between GUIs and CUIs, see Michael Lewis, "Designing for Human-Agent Interaction," *AI Magazine* 19, no. 2 (1998); and Ben Shneiderman, "Direct Manipulation versus Agents: Paths to Predictable, Controllable, and Comprehensible Interfaces," in *Software Agents*, ed. Jeffrey M. Bradshaw (Menlo Park, CA: AAAI Press, 1997).
19. "Meet Our Experimental Guardian Sous-Chef Facebook Messenger Bot," *Guardian*, June 9, 2016, <https://www.theguardian.com/technology/2016/jun/09/meet-our-experimental-guardian-sous-chef-facebook-messenger-bot>; "Want to Join the Vogue Club?" *Vogue*, February 2017, <http://www.vogue.co.uk/article/vogue-daily-update-messaging-service>.
20. For an overview of the emerging area of human-machine communication which considers the "creation of meaning among humans and machines," see Andrea L. Guzman, "What Is Human-Machine Communication, Anyway?" in *Human-Machine Communication: Rethinking Communication, Technology, and Ourselves*, ed. Andrea L. Guzman (New York: Peter Lang, 2018).
21. For a full description of the sample and our methodology for analysis, see Tanya Lokot and Nicholas Diakopoulos, "News Bots: Automating News and Information Dissemination on Twitter," *Digital Journalism* 4 no. 6 (2016).
22. Chad Edwards et al., "Is That a Bot Running the Social Media Feed? Testing the Differences in Perceptions of Communication Quality for a Human Agent and a Bot Agent on Twitter," *Computers in Human Behavior* 33 (2014).
23. Max Willens, "Bots Are the New Apps, only They Suck (for Now)," *Digiday*, October 2016, <https://digiday.com/media/bots-new-apps-suck-now/>.
24. The BBC is experimenting heavily with bots for many different purposes; see "Bots," BBC News Labs, <http://bbcnewslabs.co.uk/projects/bots/>, accessed October 26, 2018.
25. Keith Collins, "The Injuries Most Likely to Land You in an Emergency Room in America," *Quartz*, February 2016, <https://qz.com/609255/the-injuries-most-likely-to-land-you-in-an-emergency-room-in-america/>.
26. Jia Zhang, "Introducing censusAmericans, a Twitter Bot for America," *FiveThirtyEight*, July 2015, <https://fivethirtyeight.com/features/introducing-censusamericans-a-twitter-bot-for-america/>.

27. Eduardo Suarez, "How We Created Politibot, the Bot Who Reports on Telegram Messenger about the Spanish Election Campaign," *Medium* (blog), June 2016, <https://nohacefaltapapel.com/how-we-created-politibot-the-bot-who-reports-on-telegram-messenger-about-the-spanish-election-bb21242f6b4f/>; Mădălina Ciobanu, "More than 6,000 People are Talking to Politibot about the Upcoming Elections in Spain," *Journalism.co.uk*, June 2016, <https://www.journalism.co.uk/news/thousands-of-people-are-talking-to-politibot-about-the-upcoming-elections/s2/a649186/>.
28. Amanda Zamora, "T-Squared: Meet Paige, Our New Facebook Messenger Bot," *Texas Tribune*, March 20, 2017, <https://www.texastribune.org/2017/03/20/meet-paige-our-new-facebook-messenger-bot/>.
29. Jennifer Nelson, "What the Coloradoan Staff Is Learning while Experimenting with Bots," Reynolds Journalism Institute, April 2017, <https://www.rjionline.org/stories/what-the-coloradoan-staff-is-learning-while-experimenting-with-bots>.
30. Superkühle, *Westdeutscher Rundfunk*, 2017, <https://superkuehe.wdr.de/>. see also Bertram Weiß and Jakob Vicari, "Sensor Journalism: Combining Reporting with the Internet of Things," presented at the Algorithms, Automation, and News Conference, Munich, Germany, 2018.
31. In fact, Al Jazeera Interactive has a prototype of tool called "InterviewJS" that could enable such an experience; see "InterviewJS," AJInteractive, April 2018, <https://github.com/AJInteractive/InterviewJS>.
32. Julia Angwin, Ariana Tobin, and Madeleine Varner, "Have You Experienced Hate Speech on Facebook? We Want to Hear from You," ProPublica, August 2017, <https://www.propublica.org/article/have-you-experienced-hate-speech-on-facebook-we-want-to-hear-from-you>.
33. Ricardo Bilton, "With Its First Facebook Messenger Bot, ProPublica Is Collecting Reader Stories about Hate Speech," Nieman Lab, August 2017, <http://www.niemanlab.org/2017/08/with-its-first-facebook-messenger-bot-propublica-is-collecting-reader-stories-about-hate-speech/>.
34. Ariana Tobin, Madeleine Varner, and Julia Angwin, "Facebook's Uneven Enforcement of Hate Speech Rules Allows Vile Posts to Stay Up," ProPublica, December 2017, <https://www.propublica.org/article/facebook-enforcement-hate-speech-rules-mistakes>.
35. Teddy Amenabar et al., "We Made a Facebook Messenger Bot to Track People's Feelings about the Election. Here's What We Learned," *Washington Post* (blog), December 2016, <https://medium.com/thewashingtonpost/we-made-a-facebook-messenger-bot-to-track-peoples-election-feeling-here-s-what-we-learned-1752a889a849>.
36. Valerie Belair-Gagnon, Colin Agur, and Nicholas Frisch, "Mobile Sourcing: A Case Study of Journalistic Norms and Usage of Chat Apps," *Mobile Media & Communication* 3, no. 33 (2017).
37. Amanda Hickman and Westley Hennigh-Palermo, "Buzz Bot: What We Learned," BuzzFeed, September 2016, <https://www.buzzfeed.com/amandahickman/buzz-bot-what-we-learned>.
38. Keith Collins, "Watch as These Bitcoin Wallets Receive Ransomware Payments from the Ongoing Global Cyberattack," Quartz, May 2017, <https://qz.com/982993/watch-as-these-bitcoin-wallets-receive-ransomware-payments-from-the-ongoing-cyberattack/>.
39. For examples from environmental monitoring, see Weiß and Vicari, "Sensor Journalism."
40. Amar Toor, "This Twitter Bot Is Tracking Dictators' Flights in and out of Geneva," *Verge*, October 13, 2016, <https://www.theverge.com/2016/10/13/13243072/twitter-bot-tracks-dictator-planes-geneva-gva-tracker>.
41. Margaret Sullivan, "No, I Haven't Given Up on Anonymous Sources," *New York Times*, January 13, 2015, <https://publiceditor.blogs.nytimes.com/2015/01/13/new-york-times-public-editor-anonymous-sources/>; Steve Buttry, "New York Times Takes a Tougher Approach on Unnamed Sources," *The Buttry Diary* (blog), March 16, 2016, <https://stevebuttry.wordpress.com/2016/03/16/new-york-times-takes-a-tougher-approach-on-unnamed-sources/>.
42. Nicholas Diakopoulos, "Bots and the Future of Automated Accountability," *Columbia Journalism Review*, September 11, 2018. https://www.cjr.org/tow_center/prepare-to-welcome-our-accountability-bot-overlords.php.
43. Heather Ford, Elizabeth Dubois, and Cornelius Puschmann, "Keeping Ottawa Honest, One Tweet at a Time? Politicians, Journalists, Wikipedians and Their Twitter Bots," *International Journal of*

Communication 10 (2016).

44. Mark Sample, "A Protest Bot Is a Bot So Specific You Can't Mistake It for Bullshit," *Medium* (blog), May 2014, <https://medium.com/@samplereality/a-protest-bot-is-a-bot-so-specific-you-cant-mistake-it-for-bullshit-90fe10b7fbaa>.

45. "United States Foreign Intelligence Surveillance Court," Wikipedia, accessed September 16, 2017, https://en.wikipedia.org/wiki/United_States_Foreign_Intelligence_Surveillance_Court.

46. Jeremy B. Merrill, "What Is the Sound of PunditBot Yapping," *Source*, April 2016, <https://source.opennews.org/articles/punditbot-yapping/>.

47. Annalisa Merelli, "Introducing Quartz's Newest Bot: @TrumpOfYore," Quartz, June 2017, <https://qz.com/1006391/introducing-quartzs-new-bot-trump-of-yore/>.

48. "Automation Rules," Twitter, April 2017, <https://support.twitter.com/articles/76915>.

49. Amber Madison, "When Social-Media Companies Censor Sex Education," *Atlantic*, March 2015, <http://www.theatlantic.com/health/archive/2015/03/when-social-mediadors-sex-education/385576/>.

50. Colin Lecher, "Twitter Shuts Down a Site that Saved Politicians' Deleted Tweets," *Verge*, June 4, 2016, <https://www.theverge.com/2015/6/4/8731387/politwoops-sunlight-foundation-twitter>.

51. David Murphy, "Politicians Beware: Twitter Has Reinstated Politwoops," *PCMag*, December 2015, <https://www.pcmag.com/article2/0,2817,2497304,00.asp>.

52. For a critical perspective, see Shneiderman, "Direct Manipulation versus Agents."

53. Cassell, "Embodied Conversational Interface Agents."

54. Luger and Sellen, "Like Having a Really Bad PA."

55. Elizabeth Dwoskin, "The Next Hot Job in Silicon Valley Is for Poets," *Washington Post*, April 7, 2016, <https://www.washingtonpost.com/news/the-switch/wp/2016/04/07/why-poets-are-flocking-to-silicon-valley/>; John Paul Titlow, "Google Enlists Artists to Make Bots Feel like Friends," *Fast Company*, May 2016, <https://www.fastcompany.com/3060180/google-enlists-artists-to-make-bots-feel-like-friends>.

56. Gene Ball and Jack Breese, "Emotion and Personality in a Conversational Agent," in *Embodied Conversational Agents*, ed. Justine Cassell et al. (Cambridge, MA: MIT Press, 2000).

57. Jennifer Hill, W. Randolph Ford, and Ingrid G. Farreras, "Real Conversations with Artificial Intelligence: A Comparison between Human–Human Online Conversations and Human–Chatbot Conversations," *Computers in Human Behavior* 49 (2015).

58. Kevin Munger, "Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment," *Political Behavior* 39, no. 3 (2017); Amir Shevat, "Hard Questions about Bot Ethics," *Techcrunch*, September 16, 2016, <https://techcrunch.com/2016/09/16/hard-questions-about-bot-ethics/>.

59. Henriette Cramer and Jennifer Thom, "Moving Parts Surrounding Conversational UX," presented at the Conversational UX Design CHI Workshop, Denver, CO, 2017.

60. "Scripting Chatbots Is Hard. Here's How We Made It Easier for BBC Journalists," *BBC News Lab* (blog), June 2018, <https://medium.com/bbc-news-labs/bbc-botbuilder-ba8e09b6a2e9>.

61. Mark Stephen Meadows, "I'll Be Bot: Give 'Em Frikkin' License Plates," *Medium*, June 2016, <https://medium.com/@meadovian/ill-be-bot-give-em-frikkin-license-plates-343e05f82a24>.

62. Samuel Woolley and Philip Howard, "Political Communication, Computational Propaganda, and Autonomous Agents—Introduction," *International Journal of Communication* 10 (2016).

63. Alessandro Bessi and Emilio Ferrara, "Social Bots Distort the 2016 US Presidential Election Online Discussion," *First Monday* 21 no. 11 (2016); Samuel Woolley and Douglas Guilbeault, *Computational Propaganda in the United States of America: Manufacturing Consensus Online* (Oxford, UK: Computational Propaganda Research Project, Oxford Internet Institute, 2017).

64. Bence Kollanyi, Philip Howard, and Samuel Woolley, "Bots and Automation over Twitter during the Third US Presidential Debate," Oxford Internet Institute, 2016, <https://www.oii.ox.ac.uk/blog/bots-and-automation-over-twitter-during-the-third-u-s-presidential-debate/>.

65. Douglas Guilbeault and Samuel Woolley, "How Twitter Bots Are Shaping the Election," *Atlantic*, November 11, 2016, <https://www.theatlantic.com/technology/archive/2016/11/election-bots/506072/>.

66. Ryan Bort, "Nearly Half of Donald Trump's Twitter Followers Are Fake Accounts and Bots,"

Newsweek, May 2017, <http://www.newsweek.com/donald-trump-twitter-followers-fake-617873>.

67. Geoff Goldberg, “When Bots Attack,” *Medium*, March 2018, <https://medium.com/@geoffgolberg/when-bots-attack-af7f9f87b612>.

68. Luis Daniel, “Rise of the Peñabots,” *Points* (blog), February 2016, <https://points.datasociety.net/rise-of-the-pe%C3%B1abots-d35f9fe12d67>.

69. Lisa-Maria Neudert, Bence Kollanyi, and Philip Howard, “Junk News and Bots during the German Parliamentary Election: What Are German Voters Sharing over Twitter?” Computational Propaganda Project, 2017, <http://comprop.oii.ox.ac.uk/research/junk-news-and-bots-during-the-german-parliamentary-election-what-are-german-voters-sharing-over-twitter/>.

70. “Robotrolling,” NATO Stratcom Centre for Excellence, 2017, <http://stratcomcoe.org/download/file/fid/75496>.

71. Panagiotis Takis Metaxas and Eni Mustafaraj, “Social Media and the Elections,” *Science* 338, no. 6106 (2012).

72. For an example of bots used for prosocial activism, see Saiph Savage, Andrés Monroy-Hernández, and Tobias Höllerer, “Botivist: Calling Volunteers to Action Using Online Bots,” *Proceedings of Computer Supported Cooperative Work & Social Computing [CSCW]* (New York: ACM, 2016).

73. Isaac Arnsdorf, “Pro-Russian Bots Take Up the Right-Wing Cause after Charlottesville,” *ProPublica*, August 2017, <https://www.propublica.org/article/pro-russian-bots-take-up-the-right-wing-cause-after-charlottesville>.

74. Chengcheng Shao et al., “The Spread of Fake News by Social Bots,” arXiv.org, 2017, <http://arxiv.org/abs/1707.07592>.

75. Metaxas and Mustafaraj, “Social Media and the Elections.”

76. Shao et al., “The Spread of Fake News by Social Bots.”

77. Bjarke Mønsted et al., “Evidence of Complex Contagion of Information in Social Media: An Experiment Using Twitter Bots,” *PloS one* 12, no. 9 (2017).

78. Woolley and Guilbeault, *Computational Propaganda in the United States*.

79. Bessi and Ferrara, “Social Bots Distort the 2016 US Presidential Election Online Discussion.”

80. Klint Finley, “Pro-Government Twitter Bots Try to Hush Mexican Activists,” *Wired*, August 2015, <https://www.wired.com/2015/08/pro-government-twitter-bots-try-hush-mexican-activists/>.

81. Abokhodair, Yoo, and McDonald, “Dissecting a Social Botnet.”

82. Henk van Ess and Jane Lytvynenko, “This Russian Hacker Says His Twitter Bots Are Spreading Messages to Help Germany’s Far Right Party in the Election,” *Buzzfeed News*, September 2017, <https://www.buzzfeed.com/henkvaness/these-russian-hackers-say-theyre-using-twitter-bots-to-help>.

83. “#BotSpot: The Intimidators,” *Digital Forensic Research Lab* (blog), August 2017, <https://medium.com/dfrlab/botspot-the-intimidators-135244bfe46b>.

84. Fenwick McKelvey and Elizabeth Dubois, *Computational Propaganda in Canada: The Use of Political Bots* (Oxford, UK: Computational Propaganda Research Project, Oxford Internet Institute, 2017), <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Canada.pdf>.

85. Woolley and Guilbeault, *Computational Propaganda in the United States*.

86. “Fakes, Bots, and Blockings in Armenia,” *Digital Forensic Research Lab* (blog), April 2017, <https://medium.com/dfrlab/fakes-bots-and-blockings-in-armenia-44a4c87ebc46>; Brian Krebs, “Twitter Bots Use Likes, RTs for Intimidation,” *Krebs on Security* (blog), August 2017, <https://krebsonsecurity.com/2017/08/twitter-bots-use-likes-rt-for-intimidation/>. For details on how Twitter approaches bots on the platform, see Colin Crowel, “Our Approach to Bots & Misinformation,” *Twitter* (blog), June 2017, https://blog.twitter.com/official/en_us/topics/company/2017/Our-Approach-Bots-Misinformation.html.

87. John Herrman, “Not the Bots We Were Looking For,” *New York Times*, November 1, 2017, <https://www.nytimes.com/2017/11/01/magazine/not-the-bots-we-were-looking-for.html>; Meg Heckman, “Used Carefully, Chatbots Can Be an Asset to Newsrooms,” *Columbia Journalism Review*, March 2018, <https://www.cjr.org/innovations/chatbots.php>.

88. Carolina Alves de Lima Salge and Nicholas Berente, “Is That Social Bot Behaving Unethically?”

Communications of the ACM 60, no. 9 (2017).

89. Jacob Ratkiewicz et al., “Detecting and Tracking Political Abuse in Social Media,” *Proceedings of the International Conference on Web and Social Media [ICWSM]* (Palo Alto: The AAAI Press, 2011).

90. V. S. Subrahmanian et al., “The DARPA Twitter Bot Challenge,” *Computer* 49, no. 6 (2016).

91. Ferrara et al., “The Rise of Social Bots.”

92. Grimme et al., “Social Bots.”

93. Nic Dias, “The Era of Whatsapp Propaganda Is upon Us,” *Foreign Policy*, August 17, 2017, <http://foreignpolicy.com/2017/08/17/the-era-of-whatsapp-propaganda-is-upon-us/>.

94. Jen Weedon, William Nuland, and Alex Stamos, “Information Operations and Facebook,” Facebook, 2017, <https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf>. see also: “Update on Twitter’s Review of the 2016 U.S. Election,” *Twitter* (blog), January 2018, https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html.

95. Nicholas Diakopoulos, “The Bots Beat: How Not to Get Punked by Automation,” *Columbia Journalism Review*, April 2018, https://www.cjr.org/tow_center/bots-manipulate-trends.php.

96. See Botometer, <https://botometer.iuni.iu.edu/#/>, accessed October 26, 2018.

97. Schweizer Radio und Fernsehen (SRF) used a bot classifier to quantify fake followers on Instagram: “Fake It ’til You Make It: Influencer in der Schweiz,” Schweizer Radio und Fernsehen, October 2017, <https://www.youtube.com/watch?v=K90QjwJSZXY>. The *New York Times* identified patterns in followers of various “influencers” suggesting influencers had bought fake followers; see Nicholas Confessore et al., “The Follower Factory,” *New York Times*, January 27, 2018, <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>.

5. DIGITAL PAPERBOYS: ALGORITHMS IN NEWS DISTRIBUTION

1. Source concentration is well-documented for Google News as well; see Mario Haim, Andreas Graefe, and Hans-Bernd Brosius, “Burst of the Filter Bubble?” *Digital Journalism* 6, no. 3 (2018); Roland Schroeder and Moritz Kralemann, “Journalism ex Machina—Google News Germany and Its News Selection Processes,” *Journalism Studies* 6, no. 2 (2006).

2. Although accounting for 3 percent of all links observed, it was really just one story from the *Buffalo News* that gained any exposure from the search engine: Jerry Zremski, “Tax Reform Bill Could Cost Buffalo Schools \$12.2 Million, Cut Medicare,” *Buffalo News*, November 18, 2017, <http://buffalonews.com/2017/11/18/tax-reform-bill-could-cost-buffalo-schools-12-2-million/>.

3. Daniel Trielli and Nicholas Diakopoulos, “How Google Shapes the News You See about the Candidates,” *Slate*, November 2016, http://www.slate.com/articles/technology/future_tense/2016/11/how_google_shapes_the_news_you_see_about_the_candidates.html.

4. Panagiotis Takis Metaxas and Yada Pruksachatkun, “Manipulation of Search Engine Results during the 2016 US Congressional Elections,” presented at the International Conference on Internet and Web Applications and Services, Venice, Italy, 2017.

5. “The Personal News Cycle: How Americans Choose to Get Their News,” Media Insight Project, March 2017, <https://www.americanpressinstitute.org/publications/reports/survey-research/how-americans-get-news/>.

6. Nic Newman et al., *Reuters Digital News Report* (Oxford, UK: Reuters Institute for the Study of Journalism, 2017), https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf.

7. Jeffrey Gottfried and Elisa Shearer, “News Use across Social Media Platforms 2017,” Pew Research Center, September 7, 2017, <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/>.

8. Referrer Dashboard for Parse.ly Customers, https://www.parse.ly/resources/data-studies/referrer-dashboard/#google__search,facebook.com__social, accessed January 15, 2018.

9. Efrat Nechushtai, “Could Digital Platforms Capture the Media through Infrastructure?” *Journalism* 89, no. 1 (2017); Emily Bell, “The Dependent Press: How Silicon Valley Threatens Independent

Journalism,” in *Digital Dominance: Implications and Risks*, ed. Martin Moore and Damian Tambini (Oxford: Oxford University Press, 2018).

10. Will Oremus, “The Great Facebook Crash,” *Slate*, June 2018. <https://slate.com/technology/2018/06/facebooks-retreat-from-the-news-has-painful-for-publishers-including-slate.html>.

11. Shan Wang, “When a Facebook Test Moves News Stories to a Separate Feed, Traffic—and Public Discourse—Are at Stake,” Nieman Lab, October 2017, <http://www.niemanlab.org/2017/10/when-a-facebook-test-moves-news-stories-to-a-separate-feed-traffic-and-public-discourse-are-at-stake/>.

12. “The Strange Tale of USA Today’s Facebook Page,” EzyInsights, May 2017, <https://ezyinsights.com/the-strange-tale-of-usa-todays-facebook-page/>.

13. Emily Bell and Taylor Owen, *The Platform Press: How Silicon Valley Reengineered Journalism* (New York: Tow Center for Digital Journalism, 2017).

14. Philip Napoli and Robyn Caplan, “Why Media Companies Insist They’re Not Media Companies, Why They’re Wrong, and Why It Matters,” *First Monday* 22, no. 5 (2017).

15. Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media* (New Haven, CT: Yale University Press, 2018).

16. Tarleton Gillespie, “Regulation of and by Platforms,” in *The SAGE Handbook of Social Media*, ed. Jean Burgess, Alice Marwick, and Thomas Poell (London: Sage, 2018).

17. “Expanding Our Work against Abuse of Our Platform,” YouTube Blog, December 2017, <https://youtube.googleblog.com/2017/12/expanding-our-work-against-abuse-of-our.html>; “More Information, Faster Removals, More People—An Update on What We’re Doing to Enforce YouTube’s Community Guidelines,” YouTube Blog, April 2018, <https://youtube.googleblog.com/2018/04/more-information-faster-removals-more.html>; Annalee Newitz, “The Secret Lives of Google Raters,” *Ars Technica*, April 2017, <https://arstechnica.com/features/2017/04/the-secret-lives-of-google-raters/>.

18. Amanda Hess, “How YouTube’s Shifting Algorithms Hurt Independent Media,” *New York Times*, April 17, 2017, <https://www.nytimes.com/2017/04/17/arts/youtube-broadcasters-algorithm-ads.html>.

19. Amber Madison, “When Social-Media Companies Censor Sex Education,” *Atlantic*, March 2015, <http://www.theatlantic.com/health/archive/2015/03/when-social-media-censors-sex-education/385576/>.

20. Nicholas Diakopoulos et al., “I Vote For—How Search Informs Our Choice of Candidate,” in *Digital Dominance: Implications and Risks*, ed. Martin Moore and Damian Tambini (Oxford: Oxford University Press, 2018).

21. Micah Sifry, “Facebook Wants You to Vote on Tuesday. Here’s How It Messed with Your Feed in 2012,” *Mother Jones*, October 2014, <http://www.motherjones.com/politics/2014/10/can-voting-facebook-button-improve-voter-turnout/>.

22. Bing Pan et al., “In Google We Trust: Users’ Decisions on Rank, Position, and Relevance,” *Journal of Computer-Mediated Communication* 12, no. 3 (2007); Eugene Agichtein et al., “Learning User Interaction Models for Predicting Web Search Result Preferences,” *Proceedings of the Conference on Research and Development in Information Retrieval (SIGIR)* (New York: ACM, 2006).

23. Robert Epstein and Ronald Robertson, “The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections,” *Proceedings of the National Academy of Sciences* 112, no. 33 (2015).

24. Taina Bucher, “Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook,” *New Media & Society* 14, no. 7 (2014); Tarleton Gillespie, “The Relevance of Algorithms,” in *Media Technologies: Essays on Communication, Materiality, and Society*, ed. Tarleton Gillespie, Pablo Boczkowski, and Kirstin Foot (Cambridge, MA: MIT Press, 2014); Martin Moore, *Tech Giants and Civic Power* (London: Centre for the Study of Media Communication & Power, 2016), <https://www.kcl.ac.uk/sspp/policy-institute/cmcp/tech-giants-and-civic-power.pdf>.

25. Zeynep Tufekci, “Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency,” *Colorado Technology Law Journal* 13 (2015).

26. Engin Bozdag and Jeroen van den Hoven, “Breaking the Filter Bubble: Democracy and Design,” *Ethics and Information Technology* 17, no. 4 (2015).

27. Motahhare Eslami et al., "I Always Assumed that I Wasn't Really that Close to [Her]: Reasoning about Invisible Algorithms in News Feeds," *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM, 2015); Elia Powers, "My News Feed Is Filtered?" *Digital Journalism* 5, no. 10 (2017).
28. Engin Bozdog, "Bias in Algorithmic Filtering and Personalization," *Ethics and Information Technology* 15, no. 3 (2013).
29. Philip Napoli, "Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers," *Telecommunications Policy* 39, no. 9 (2015); Mike Ananny and Kate Crawford, "A Liminal Press," *Digital Journalism* 3, no. 2 (2015); Matthew Weber and Allie Kosterich, "Coding the News," *Digital Journalism* 6, no. 3 (2018); Matt Carlson, "Facebook in the News," *Digital Journalism* 6, no. 1 (2018).
30. Rasmus Kleis Nielsen and Sarah Anne Ganter, "Dealing with Digital Intermediaries: A Case Study of the Relations between Publishers and Platforms," *New Media & Society* 20, no. 4 (2017).
31. Robyn Caplan and Danah Boyd, "Isomorphism through Algorithms: Institutional Dependencies in the Case of Facebook," *Big Data & Society* 5, no 1 (2018).
32. Caitlin Petre, *The Traffic Factories: Metrics at Chartbeat, Gawker Media, and the New York Times* (New York: Tow Center for Digital Journalism, 2015), <https://academiccommons.columbia.edu/catalog/ac:kd51c59zvx>.
33. Hong Tien Vu, "The Online Audience as Gatekeeper: The Influence of Reader Metrics on News Editorial Selection," *Journalism* 15, no. 8 (2013).
34. Philip Napoli, "Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption," *Communication Theory* 24, no. 3 (2014).
35. Eun-Ju Lee and Edson Tandoc, "When News Meets the Audience: How Audience Feedback Online Affects News Production and Consumption," *Human Communication Research* 43, no. 4 (2017); Kasper Welbers et al., "News Selection Criteria in the Digital Age: Professional Norms versus Online Audience Metrics," *Journalism* 17, no. 8 (2015).
36. Alicja Piotrkowicz et al., "Headlines Matter: Using Headlines to Predict the Popularity of News Articles on Twitter and Facebook," *Proceedings of the International Conference on Web and Social Media* (Palo Alto, CA: AAAI, 2017).
37. Shailesh Prakash, "Journalism & Technology: Big Data, Personalization & Automation," keynote presented at the Computation + Journalism Symposium, Evanston, IL, 2017, <https://www.youtube.com/watch?v=PqMvxo89AQ4&index=1&list=PL3kjMDQ990L556im9F1qozenDo2rMpccH>.
38. Shuguang Wang, Eui-Hong (Sam) Han, and Alexander M. Rush, "Headliner: An Integrated Headline Suggestion System," presented at the Computation + Journalism Symposium, Palo Alto, CA, 2016.
39. Yaser Keneshloo et al., "Predicting the Shape and Peak Time of News Article Views," presented at the International Conference on Big Data, Washington, DC, 2016.
40. Shan Wang, "The New York Times Built a Slack Bot to Help Decide which Stories to Post to Social Media," Nieman Lab, August 2015, <http://www.niemanlab.org/2015/08/the-new-york-times-built-a-slack-bot-to-help-decide-which-stories-to-post-to-social-media/>.
41. Carlos Castillo et al., "Characterizing the Life Cycle of Online News Stories Using Social Media Reactions," *Proceedings of the Conference on Computer Supported Cooperative Work & Social Computing* (New York: ACM, 2014).
42. Alexander Spangher, "Building the Next New York Times Recommendation Engine," *New York Times*, August 11, 2015, <https://open.blogs.nytimes.com/2015/08/11/building-the-next-new-york-times-recommendation-engine/>.
43. Prakash, "Journalism & Technology."
44. Jiahui Liu, Peter Dolan, Elin Rønby Pedersen, "Personalized News Recommendation Based on Click Behavior," *Proceedings of the International Conference on Intelligent User Interfaces* (New York: ACM, 2010).

45. Shan Wang, "The New York Times Is Trying to Narrow the Distance between Reporters and Analytics Data," *Nieman Lab*, July, 2016, <http://www.niemanlab.org/2016/07/the-new-york-times-is-trying-to-narrow-the-distance-between-reporters-and-analytics-data/>.
46. Sheila Doshi, "How BuzzFeed Uses Real-Time Machine Learning to Choose Their Viral Content," *Domino Data Lab* (blog), October 2016, <https://blog.dominodatalab.com/buzzfeed-uses-real-time-machine-learning-choose-viral-content/>.
47. Federica Cherubini and Rasmus Kleis Nielson, *Editorial Analytics: How News Media Are Developing and Using Audience Data and Metrics* (Oxford, UK: Reuters Institute for the Study of Journalism, 2016).
48. Michael A. DeVito, "From Editors to Algorithms," *Digital Journalism* 5, no. 6 (2017); Kelly Cotter, Janghee Cho, and Emilee Rader, "Explaining the News Feed Algorithm: An Analysis of the 'News Feed FYI' Blog," *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (New York: ACM, 2017).
49. Chris Wiggins, "Data Science @ The New York Times," keynote presented at the Computation + Journalism Symposium, New York, 2015, <https://youtu.be/jgurTmCyAA?t=22231>.
50. Shan Wang, "After Years of Testing, The Wall Street Journal Has Built a Paywall That Bends to the Individual Reader," *Nieman Lab*, February 2018, <http://www.niemanlab.org/2018/02/after-years-of-testing-the-wall-street-journal-has-built-a-paywall-that-bends-to-the-individual-reader/>; Liam Corcoran, "Not All News Site Visitors Are Created Equal. Schibsted is Trying to Predict the Ones Who Will Pay Up," *Nieman Lab*, February 2018, <http://www.niemanlab.org/2018/02/not-all-news-site-visitors-are-created-equal-schibsted-is-trying-to-predict-the-ones-who-will-pay-up/>; Lucinda Southern, "How Swiss News Publisher NZZ Built a Flexible Paywall Using Machine Learning," *Digiday*, June 2018, <https://digiday.com/media/swiss-news-publisher-nzz-built-flexible-paywall-using-machine-learning>.
51. Nir Grinberg, "Identifying Modes of User Engagement with Online News and Their Relationship to Information Gain in Text," *Proceedings of the World Wide Web Conference* (New York: ACM, 2018).
52. Shan Wang, "Die Welt's Analytics System De-Emphasizes Clicks and Demystifies What It Considers a 'Quality' Story," *Nieman Lab*, May 2016, <http://www.niemanlab.org/2016/05/die-welts-analytics-system-de-emphasizes-clicks-and-demystifies-what-it-considers-a-quality-story/>.
53. Edson Tandoc and Ryan Thomas, "The Ethics of Web Analytics," *Digital Journalism* 3, no. 2 (2014).
54. Elia Powers, "Selecting Metrics, Reflecting Norms," *Digital Journalism* 6, no. 4 (2018).
55. Pablo Boczkowski and Eugenia Mitchelstein, *The News Gap: When the Information Preferences of the Media and the Public Diverge* (Cambridge, MA: MIT Press, 2013).
56. Krishna Bharat, Tomonari Kamba, and Michael Albers, "Personalized, Interactive News on the Web," *Multimedia Systems* 6, no. 5 (1998).
57. See work from the Center for Investigative Reporting on their Impact Tracker: Lindsay Green-Barber, "CIR's Impact Tracker: How to Use It and Why You Need It," Center for Investigative Reporting, September 2016, <https://www.revealnews.org/article/cirs-impact-tracker-how-to-use-it-and-why-you-need-it/>.
58. Michael Schudson, *The Sociology of News*, 2nd ed. (New York: W. W. Norton, 2011).
59. Felipe Rodrigues, "Meet the Swedish Newspaper Editor Who Put an Algorithm in Charge of His Homepage," *Storybench*, March 2017, <http://www.storybench.org/meet-swedish-newspaper-editor-put-algorithm-charge-homepage/>.
60. Balázs Bodó, "Means, Not an End (of the World)—The Customization of News Personalization by European News Media," presented at the Algorithms, Automation, and News Conference, Munich, Germany, 2018.
61. René Pfitzner, "Data Science for a Smart News Experience," *Medium*, October 2017, <https://medium.com/@RenePfitznerZH/data-science-for-a-smart-news-experience-35d316846d04>.
62. Mike Vlad Cora, "Detecting Trustworthy Domains," *Flipboard Engineering* (blog), April 2017, <http://engineering.flipboard.com/2017/04/domainranking>.

63. For instance, Facebook relies on user surveys to help define what it sees as a “high-quality” post. See Varun Kacholia, “News Feed FYI: Showing More High Quality Content,” *Facebook News Feed FYI* (blog), August 2013, <https://www.facebook.com/business/news/News-Feed-FYI-Showing-More-High-Quality-Content>.

64. See Nicholas Diakopoulos, “Picking the NYT Picks: Editorial Criteria and Automation in the Curation of Online News Comments,” *#ISOJ Journal* 6, no. 1 (2015); Yinfei Yang and Ani Nenkova, “Combining Lexical and Syntactic Features for Detecting Content-Dense Texts in News,” *Journal of Artificial Intelligence Research* 60, no. 1 (2017); and Frederic Filloux, “Scoring Stories to Make Better Recommendation Engines for News,” *Monday Note (blog)*, October 2017, <https://mondaynote.com/scoring-stories-to-make-better-recommendation-engines-c3c73a596893>.

65. Sounel Park et al., “NewsCube: Delivering Multiple Aspects of News to Mitigate Media Bias,” *Proceedings of the Conference on Human Factors in Computing Systems* (New York: ACM, 2009).

66. C. W. Anderson, “Deliberative, Agonistic, and Algorithmic Audiences: Journalism’s Vision of Its Public in an Age of Audience Transparency,” *International Journal of Communication* 5 (2011); Bozdag and van den Hoven, “Breaking the Filter Bubble.”

67. Mike Ananny, “Networked News Time,” *Digital Journalism* 4 no. 4 (2016).

68. Kjerstin Thorson and Chris Wells, “Curated Flows: A Framework for Mapping Media Exposure in the Digital Age,” *Communication Theory* 26, no. 3 (2015); Adrienne LaFrance, “What If You Could Subscribe to Somebody Else’s Facebook Feed?” *Atlantic*, August 2014, <https://www.theatlantic.com/technology/archive/2014/08/what-if-people-could-subscribe-to-different-facebook-algorithms/378925/>.

69. Roger McNamee, “How to Fix Facebook—Before It Fixes Us,” *Washington Monthly*, January–March 2018, <https://washingtonmonthly.com/magazine/january-february-march-2018/how-to-fix-facebook-before-it-fixes-us/>.

70. Jonah Engel Bromwich and Matthew Haag, “Facebook Is Changing. What Does That Mean for Your News Feed?” *New York Times*, January 12, 2018, <https://www.nytimes.com/2018/01/12/technology/facebook-news-feed-changes.html>.

71. See an early study of the rollout of the technology at Omni in Taina Bucher, “‘Machines Don’t Have Instincts’: Articulating the Computational in Journalism,” *New Media & Society* 12, no. 5 (2016).

72. Carlson, “Facebook in the News.”

6. ALGORITHMIC ACCOUNTABILITY REPORTING

1. This scenario is fictitious, but informed by real data from an individual who was scored by the Chicago Police Department.

2. Chicago Police Department, *Custom Notifications in Chicago*, October 2015, <http://directives.chicagopolice.org/directives/data/a7a57bf0-1456faf9-bfa14-570a-a2debf33c56ae59.html>.

3. Rob Arthur, “We Now Have Algorithms to Predict Police Misconduct,” *FiveThirtyEight*, March 2016, <https://fivethirtyeight.com/features/we-now-have-algorithms-to-predict-police-misconduct/>.

4. Matt Stroud, “Chicago’s Predictive Policing Tool Just Failed a Major Test,” *Verge*, August 19, 2016, <https://www.theverge.com/2016/8/19/12552384/chicago-heat-list-tool-failed-rand-test>; Yana Kunichoff and Patrick Sier, “The Contradictions of Chicago Police’s Secretive List,” *Chicago Magazine*, August 2017, <http://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/>. For more on how predictive policing can create feedback loops, see Kristian Lum and William Isaac, “To Predict and Serve?” *Significance* 13, no. 5 (2015).

5. Mick Dumke and Frank Main, “A Look Inside the Watch List Chicago Police Fought to Keep Secret,” *Chicago Sun Times*, May 18, 2017, <https://chicago.suntimes.com/chicago-politics/what-gets-people-on-watch-list-chicago-police-fought-to-keep-secret-watchdogs/>.

6. Jeff Asher and Rob Arthur, “Inside the Algorithm that Tries to Predict Gun Violence in Chicago,” *New York Times*, June 13, 2017, <https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html>.

7. The code for the model was supplied to the author upon request, and the results were replicated. A

generalized linear model (GLM) with eight variables explained 94.6 percent of the variance in the SSL score; in statistical terms $R^2 = 0.946$.

8. Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016); Frank Pasquale. *The Black Box Society: The Secret Algorithms that Control Money and Information* (Cambridge, MA: Harvard University Press, 2015); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin’s Press, 2018).

9. Batya Friedman and Helen Nissenbaum, “Bias in Computer Systems,” *ACM Transactions on Information Systems* 14, no. 3 (1996); Peter Neumann, *Computer-Related Risks* (New York: ACM Press, 1995).

10. Nicholas Diakopoulos, *Algorithmic Accountability Reporting: On the Investigation of Black Boxes* (New York: Tow Center for Digital Journalism, January, 2014). The phrase “algorithmic accountability” was originally coined in Nicholas Diakopoulos, “Sex, Violence, and Autocomplete Algorithms,” *Slate*, August 2013, http://www.slate.com/articles/technology/future_tense/2013/08/words_banned_from_bing_and_google_s_autocomplete_algorithms.html, and elaborated in Nicholas Diakopoulos, “Rage against the Algorithms,” *Atlantic*, October 2013, <https://www.theatlantic.com/technology/archive/2013/10/rage-against-the-algorithms/280255/>.

11. Mark Bovens, Thomas Schillemans, and Robert E. Goodin, “Public Accountability,” in *The Oxford Handbook of Public Accountability*, ed. Mark Bovens, Robert E. Goodin, and Thomas Schillemans (Oxford: Oxford University Press, 2014).

12. Dillon Reisman et al., *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability* (New York, AI Now Institute, 2018), <https://ainowinstitute.org/aiareport2018.pdf>.

13. For stories in this series, see “Machine Bias: Investigating Algorithmic Injustice,” *ProPublica*, <https://www.propublica.org/series/machine-bias>, accessed October 26, 2018.

14. Diakopoulos, “Sex, Violence, and Autocomplete Algorithms”; Jennifer Stark and Nicholas Diakopoulos, “Uber Seems to Offer Better Service in Areas with More White People. That Raises Some Tough Questions,” *Washington Post*, March 10, 2016, <https://www.washingtonpost.com/news/wonk/wp/2016/03/10/uber-seems-to-offer-better-service-in-areas-with-more-white-people-that-raises-some-tough-questions/>.

15. Nick Seaver, “Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems,” *Big Data & Society* 4, no. 2 (2017); *Algorithmic Accountability: Applying the Concept to Different Country Contexts*, World Wide Web Foundation, 2017, http://webfoundation.org/docs/2017/07/WF_Algorithms.pdf; Mike Ananny, “Toward an Ethics of Algorithms,” *Science, Technology & Human Values* 41, no. 1 (2015).

16. Taina Bucher, *If ... Then: Algorithmic Power and Politics* (New York: Oxford University Press, 2018).

17. In March 2014 researchers in the United States uncovered discrepancies in the output emissions of certain Volkswagen diesel engines when tested under normal road conditions. But this only turned into national news about eighteen months later, when it became clear that VW had used a software algorithm to deceive emissions testing and that engineers had tried to cover their tracks to evade regulators. This *intentional* deception made the story much more powerful than had the explanation been “it was a bug” or “it emerged from unanticipated contexts of use.”

18. The notion of sociotechnical blindness comes from Deborah Johnson and Mario Verdicchio, “AI Anxiety,” *Journal of the Association for Information Science and Technology* 68, no. 9 (2017).

19. Julia Angwin et al., “Machine Bias,” *ProPublica*, May 2016, <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

20. Jeff Larson et al., “How We Analyzed the COMPAS Recidivism Algorithm,” *ProPublica*, May 2016, <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm/>.

21. William Dieterich, Christina Mendoza, and Tim Brennan, “COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity,” Northpointe Research Report, July 2016.

22. Ongoing research has further critiqued the algorithm by pointing out that crowd workers were able to predict recidivism at the same accuracy level, and that a simplified linear model using two features (age and total number of previous convictions) performs just as well as the algorithm. See Julia Dressel and Hany Farid, “The Accuracy, Fairness, and Limits of Predicting Recidivism,” *Science Advances* 4, no. 1 (2018).
23. Mirror Mirror: Reflections on Quantitative Fairness, accessed August 2, 2018, <https://speak-statistics-to-power.github.io/fairness/>. See also Arvind Narayanan, “21 Fairness Definitions and Their Politics,” tutorial presented at the Fairness, Accountability, and Transparency Conference, New York, 2018, <https://www.youtube.com/watch?v=jIXIuYdnyyk>.
24. Bruno Lepri et al., “Fair, Transparent, and Accountable Algorithmic Decision-Making Processes,” *Philosophy & Technology* 84, no. 3 (2017).
25. Richard Berk et al., “Fairness in Criminal Justice Risk Assessments: The State of the Art,” *Sociological Methods & Research* (2018).
26. “Unfairness by Algorithm: Distilling the Harms of Automated Decision-Making,” Future of Privacy Forum, December 2017, <https://fpf.org/wp-content/uploads/2017/12/FPF-Automated-Decision-Making-Harms-and-Mitigation-Charts.pdf>. For a treatment of representational harms to individuals and society, see Kate Crawford, “The Trouble with Bias,” keynote presentation at the Neural Information Processing Systems Conference, Long Beach, CA, 2017, https://www.youtube.com/watch?v=fMym_BKWQzk.
27. Meghan E. Irons, “Caught in a Dragnet,” *Boston Globe*, July 17, 2011, http://archive.boston.com/news/local/massachusetts/articles/2011/07/17/man_sues_registry_after_license_mistakenly_revoked/.
28. Sapna Maheshwari, “On YouTube Kids, Startling Videos Slip Past Filters,” *New York Times*, November 4, 2017, https://www.nytimes.com/2017/11/04/business/media/youtube-kids-paw-patrol.html?_r=0.
29. Diakopoulos, “Sex, Violence, and Autocomplete Algorithms.”
30. Simon Scott, “With a Hurricane Approaching Florida, Airline Algorithms Show No Sympathy,” NPR, September 9, 2017, <https://www.npr.org/2017/09/09/548853940/with-a-hurricane-approaching-florida-airline-algorithms-show-no-sympathy>.
31. “Sydney Siege Sees Uber Raise Prices before Backtracking,” BBC, December 2014, <http://www.bbc.com/news/technology-30478008>.
32. Michal Kosinski, David Stillwell, and Thore Graepel, “Private Traits and Attributes Are Predictable from Digital Records of Human Behavior,” *Proceedings of the National Academies of Sciences* 110, no. 15 (2013).
33. Kate Crawford and Jason Schultz, “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms,” *Boston College Law Review* 55, no. 1 (2014).
34. Kashmir Hill, “How Facebook Outs Sex Workers,” Gizmodo, October 2017, <https://gizmodo.com/how-facebook-outs-sex-workers-1818861596>.
35. Kashmir Hill, “How Facebook Figures Out Everyone You’ve Ever Met,” Gizmodo, November 2017, <https://gizmodo.com/how-facebook-figures-out-everyone-youve-ever-met-1819822691>.
36. Ryan McNeill and Janet Roberts, “Exclusive: Ready for Sandy, NJ Transit Erred in Modeling Storm,” Reuters, January 11, 2013, <http://www.reuters.com/article/us-storm-sandy-newjerseytransit/exclusive-ready-for-sandy-nj-transit-erred-in-modeling-storm-idUSBRE90B00I20130112>.
37. Friedman and Nissenbaum, “Bias in Computer Systems.” See also the notion of “transfer context bias” in David Danks and Alex John London, “Algorithmic Bias in Autonomous Systems,” presented at the International Joint Conference on Artificial Intelligence, Melbourne, Australia, 2017.
38. “Machine Bias with Jeff Larson,” Data Stories Podcast, October 2016, <http://datastori.es/85-machine-bias-with-jeff-larson/>.
39. Daniel Trielli and Nicholas Diakopoulos, “How to Report on Algorithms Even If You’re Not a Data Whiz,” *Columbia Journalism Review*, May 2017, https://www.cjr.org/tow_center/algorithms-reporting-algorithmtips.php.
40. Eldad Eilam, *Reversing: Secrets of Reverse Engineering* (Indianapolis: Wiley, 2005).
41. Lennart Ljung, “System Identification,” in *Signal Analysis and Prediction*, ed. Ales Procházka et al.

(Boston, MA: Birkhäuser, 1998).

42. Bucher, *If ... Then: Algorithmic Power and Politics*.

43. S. Michael Gaddis, "An Introduction to Audit Studies in the Social Sciences," in *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, ed. S. Michael Gaddis (Cham, Switzerland: Springer, 2017).

44. Christian Sandvig et al., "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms," presented at the International Communication Association Preconference on Data and Discrimination Converting Critical Concerns into Productive Inquiry, Seattle, WA, 2014.

45. Jennifer Valentino-DeVries, Jeremy Singer-Vine, and Ashkan Soltani, "Websites Vary Prices, Deals Based on Users' Information," *Wall Street Journal*, December 24, 2012, <https://www.wsj.com/articles/SB1000142412788732377204578189391813881534>.

46. Kashmir Hill and Surya Mattu, "Keep Track of Who Facebook Thinks You Know with This Nifty Tool," *Gizmodo*, January 2018, <https://gizmodo.com/keep-track-of-who-facebook-thinks-you-know-with-this-ni-1819422352>.

47. In early 2018 Algorithm Watch launched an initiative to gather data donations that would enable an audit of the Schufa credit score in Germany; see <https://www.startnext.com/openschufa>, accessed October 26, 2018.

48. Jeremy B. Merrill, "Why Facebook Showed You That Ad for the Candidate You Hate," *New York Times*, November 8, 2016, <https://www.nytimes.com/2016/11/08/us/politics/facebook-ads-campaign.html>; Julia Angwin, Surya Mattu, and Terry Parris Jr., "Facebook Doesn't Tell Users Everything It Really Knows About Them," *ProPublica*, December 2016, <https://www.propublica.org/article/facebook-doesnt-tell-users-everything-it-really-knows-about-them>. For information on Algorithm Watch's Google audit, see "Datenspende," Algorithm Watch, 2017, <https://datenspende.algorithmwatch.org/en/index.html>.

49. Bucher, *If ... Then: Algorithmic Power and Politics*.

50. Katie Notopoulos, "How I Cracked Facebook's New Algorithm and Tortured My Friends," *BuzzFeed News*, February 2018, <https://www.buzzfeed.com/katienotopoulos/how-i-cracked-facebooks-new-algorithm-and-tortured-my>.

51. Sandvig et al., "Auditing Algorithms."

52. Danielle Citron, "Technological Due Process," *Washington University Law Review* 85 (2017). For a recent example, see Colin Lecher, "What Happens When an Algorithms Cuts your Health Care," *Verge*, March 21, 2018, <https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy>.

53. Brendt Mittelstadt, "Auditing for Transparency in Content Personalization Systems," *International Journal of Communication*, 10 (2016).

54. Lauren Kirchner, "ProPublica Seeks Source Code for New York City's Disputed DNA Software," *ProPublica*, September 2017, <https://www.propublica.org/article/propublica-seeks-source-code-for-new-york-city-disputed-dna-software>.

55. Shefali Patil, Ferdinand Vieider, and Philip Tetlock, "Process versus Outcome Accountability," in *The Oxford Handbook of Public Accountability*, ed. Mark Bovens, Robert E. Goodin, and Thomas Schillemans (Oxford: Oxford University Press, 2014); Karen Yeung, "Algorithmic Regulation: A Critical Interrogation," *Regulation & Governance* (2017).

56. Crawford and Schultz, "Big Data and Due Process."

57. Carolina Alves de Lima Salge and Nicholas Berente, "Is That Social Bot Behaving Unethically?" *Communications of the ACM* 60, no. 9 (2017).

58. Sentiments like this motivated our development of [Algorithmtips.org](http://algorithmtips.org), a website that seeks to lower the cost of finding newsworthy leads about the use of algorithms in government by providing an easily searchable database. See Daniel Trielli, Jennifer Stark, and Nicholas Diakopoulos, "Algorithm Tips: A Resource for Algorithmic Accountability in Government," presented at the Computation + Journalism Symposium, Evanston, IL, October, 2017; and see the Algorithm Tips site online at <http://algorithmtips.org/>, accessed October 26, 2018.

59. Nicholas Diakopoulos and Michael Koliska, “Algorithmic Transparency in the News Media,” *Digital Journalism* 5, no. 7 (2017).
60. Albert Meijer, Mark Bovens, and Thomas Schillemans, “Transparency,” in *The Oxford Handbook of Public Accountability*, ed. Mark Bovens, Robert. E. Goodin, and Thomas Schillemans (Oxford: Oxford University Press, 2014).
61. Nicholas Diakopoulos et al., “I Vote For—How Search Informs Our Choice of Candidate,” in *Digital Dominance: Implications and Risks*, ed. Martin Moore and Damian Tambini (Oxford: Oxford University Press, 2018).
62. Gaddis, “An Introduction to Audit Studies in the Social Sciences”
63. Nicholas Diakopoulos, “Accountability in Algorithmic Decision Making,” *Communications of the ACM* 59, no. 2 (2016).
64. Diakopoulos et al., “I Vote For—How Search Informs Our Choice of Candidate.”
65. Walt Hickey, “Be Suspicious of Online Movie Ratings, Especially Fandango’s,” *FiveThirtyEight*, October 2015, <https://fivethirtyeight.com/features/fandango-movies-ratings/>.
66. Julia Angwin and Surya Mattu, “Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn’t,” *ProPublica*, September 2016, <https://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its-pricing-algorithm-doesnt>.
67. Reubin Binns, “Algorithmic Accountability and Public Reason,” *Philosophy & Technology* (2017).
68. Universal Declaration of Human Rights, 1948, United Nations, <http://www.un.org/en/universal-declaration-human-rights/>. For an additional treatment of how AI could more broadly impact on human rights see: Mark Latonero, *Governing Artificial Intelligence: Upholding Human Rights & Dignity* (New York: Data & Society, 2018).
69. Two federal agencies, the Department of the Treasury (DoT) and the Social Security Administration, explicitly claim that software or source code are not “agency records”—meaning that they are not subject to public records requests. The guidance from DoT in the Code of Federal Regulations states unequivocally that “proprietary (or copyrighted) software is not an agency record.” See Code of Federal Regulations § 1.1 General, United States, <https://www.law.cornell.edu/cfr/text/31/1.1>.
70. Nicholas Diakopoulos, “We Need to Know the Algorithms the Government Uses to Make Important Decisions about Us,” *Conversation*, May 2016, <https://theconversation.com/we-need-to-know-the-algorithms-the-government-uses-to-make-important-decisions-about-us-57869>.
71. For the specific language we used in our requests see: <http://algorithmtips.org/resources/>, accessed October 27, 2018.
72. Katherine Fink, “Opening the Government’s Black Boxes: Freedom of Information and Algorithmic Accountability,” *Information, Communication, & Society* 21, no. 10 (2018).
73. Robert Brauneis and Ellen Goodman, “Algorithmic Transparency for the Smart City,” *Yale Journal of Law & Technology* 20 (2018).
74. Esha Bhandari and Rachel Goodman, “Data Journalism and the Computer Fraud and Abuse Act: Tips for Moving Forward in an Uncertain Landscape,” presented at the Computation + Journalism Symposium, Evanston, IL, 2017.
75. Tarleton Gillespie, “The Relevance of Algorithms,” in *Media Technologies: Essays on Communication, Materiality, and Society*, ed. Tarleton Gillespie, Pablo Boczkowski, and Kirstin Foot (Cambridge, MA: MIT Press, 2014).
76. Gillespie, “The Relevance of Algorithms”; Mike Ananny, “Toward an Ethics of Algorithms,” *Science, Technology & Human Values* 41, no. 1 (2015).
77. Susanne Fengler and Stephan Russ-Mohl, “The (Behavioral) Economics of Media Accountability,” in *Journalists and Media Accountability: An International Study of News People in the Digital Age*, ed. Susanne Fengler et al. (New York: Peter Lang, 2014).
78. Michael Ananny and Kate Crawford, “Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability,” *New Media & Society* 20, no. 3 (2018).
79. Kelly McBride and Tom Rosenstiel, *The New Ethics of Journalism: Principles for the 21st Century*

(Thousand Oaks, CA: CQ Press, 2013).

80. Mark Deuze, "What Is Journalism?: Professional Identity and Ideology of Journalists Reconsidered," *Journalism* 6, no. 4 (2015).

81. Martje ter Hoeve et al., "Do News Consumers Want Explanations for Personalized News Rankings?" presented at the Fairness, Accountability, and Transparency in Recommender Systems (FATREC) Workshop, Como, Italy, 2017.

82. Stephen J. A. Ward, "The Magical Concept of Transparency," in *Ethics for Digital Journalists: Emerging Best Practices*, ed. Laurie Zion and David Craig (New York: Routledge, 2015).

83. Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* 3, no. 1 (2016).

84. Diakopoulos and Koliska, "Algorithmic Transparency in the News Media."

85. Tal Montal and Zvi Reich, "I, Robot. You, Journalist. Who Is the Author?" *Digital Journalism* 5, no. 7 (2017).

86. Ward, "The Magical Concept of Transparency."

87. For an extensive accounting of factors influencing journalistic responsibility, see Denis McQuail, *Journalism and Society* (London: Sage, 2013).

88. For a list of dozens of Github repositories maintained by newsrooms, see <https://github.com/silva-shih/open-journalism>, accessed October 26, 2018. For examples of specific repositories that support editorial projects, see Jennifer Stark and Nicholas Diakopoulos, "Towards Editorial Transparency in Computational Journalism," presented at Computation + Journalism Symposium, Palo Alto, CA, 2016; Ryann Grochowski Jones and Charles Ornstein, "Matching Industry Payments to Medicare Prescribing Patterns: An Analysis," ProPublica, March 2016, <https://static.propublica.org/projects/d4d/20160317-matching-industry-payments.pdf?22>.

89. Stuart Myles, "How Can We Make Algorithmic News More Transparent?," presented at the Algorithms, Automation, and News Conference, Munich, Germany, 2018.

90. Nicholas Diakopoulos, "Enabling Accountability of Algorithmic Media: Transparency as a Constructive and Critical Lens," in *Transparent Data Mining for Big and Small Data*, ed. Tania Cerquitelli, Daniele Quercia, and Frank Pasquale (Cham, Switzerland: Springer International, 2017).

91. Nicholas Diakopoulos, "BuzzFeed's Pro Tennis Investigation Displays Ethical Dilemmas of Data Journalism," *Columbia Journalism Review*, November 2016, https://www.cjr.org/tow_center/transparency_algorithms_buzzfeed.php.

92. Tamar Charney, Michael Oreskes, and Thomas Hjelm, "The Secret Sauce behind NPR One: An Editorially Responsible Algorithm," NPR, December 21, 2016, <https://www.npr.org/sections/npr-extra/2016/12/21/505315422/secret-sauce-npr-one-algorithm>.

CONCLUSION: THE FUTURE OF ALGORITHMIC NEWS MEDIA

1. Nicholas Diakopoulos, "There Are a Lot of Rote Tasks a Good AI Interviewer Could Do for You," *Columbia Journalism Review*, June 2018, https://www.cjr.org/tow_center/artificial-intelligence-reporting-interviews.php.

2. Astrid Gynnild and Turo Uskali, eds., *Responsible Drone Journalism* (New York: Routledge, 2018).

3. Pietro Passarelli, "An Open Source Tool for Enabling Faster, Easier Editing of Video Interviews," *Product* (blog), Vox, November 22, 2016, <https://product.voxmedia.com/2016/11/22/13669486/faster-video-editing>.

4. Lucas Graves, "Understanding the Promise and Limits of Automated Fact-Checking" (Oxford, UK: Reuters Institute for the Study of Journalism, 2018), <https://reutersinstitute.politics.ox.ac.uk/our-research/understanding-promise-and-limits-automated-fact-checking>.

5. For details on how to apply user research methods in practice, see Kathy Baxter, Catherine Courage, and Kelly Caine, *Understanding Your Users: A Practical Guide to User Research Methods*, 2nd ed. (Waltham, MA: Morgan Kaufman, 2015). For examples of the user centered design process applied to automated journalism, see Hanna Zoon, "Designing for Automated Journalism in the Netherlands: First

Steps and No Way Back,” presented at the Algorithms, Automation, and News Conference, Munich, Germany, 2018, <https://hannazon.files.wordpress.com/2018/05/abstract-design-for-automated-journalism-in-nl-zoon-van-dongen-alves-lino.pdf>.

6. See also Nikki Usher, *Interactive Journalism: Hackers, Data, and Code* (Urbana: University of Illinois Press, 2016).

7. Philip Meyer, *Precision Journalism: A Reporter’s Introduction to Social Science Methods*, 4th ed. (Lanham, MD: Rowman & Littlefield, 2002).

8. Charles Berret and Cheryl Phillips, “Teaching Data and Computational Journalism” (New York: Columbia Journalism School, 2016), https://journalism.columbia.edu/system/files/content/teaching_data_and_computational_journalism.pdf.

9. Bahareh R. Heravi, “3WS of Data Journalism Education,” *Journalism Practice* (2018).

10. For a critique of price, see Felix Salmon’s contribution to Bill Grueskin, Felix Salmon, and Alexandria Neason, “Do We Need J-Schools?,” *Columbia Journalism Review*, Spring / Summer 2018, https://www.cjr.org/special_report/do-we-need-j-schools.php/.

11. For a treatment of several worrisome scenarios see: Robert Chesney and Danielle Keats Citron, “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security,” *California Law Review* 107 (2019).

12. For an overview of media forensic techniques, see Hany Farid, *Photo Forensics* (Cambridge, MA: MIT Press, 2016).

13. V. Conotter, E. Bodnari, G. Boato, and H. Farid, “Physiologically-Based Detection of Computer Generated Faces in Video,” presented at the International Conference on Image Processing (ICIP), Paris, France, 2014.

14. For a list of digital forensics tools, see Bellingcat’s Digital Toolkit, <https://docs.google.com/document/d/1BfLPJpRtyq4RFtHJoNpvWQjmGnyVkfE2HYoICKOGguA/edit>, accessed October 27, 2018.

15. See Media Forensics (MediFor), Defense Advanced Research Projects Agency (DARPA), <https://www.darpa.mil/program/media-forensics>; Media Forensics Challenge 2018, National Institute of Standards and Technology (NIST), <https://www.nist.gov/itl/iad/mig/media-forensics-challenge-2018>; InVID—In Video Veritas, <http://www.invid-project.eu/>, accessed October 27, 2018.

16. Such guides are beginning to be developed and standardized. See First Draft, Resources, <https://firstdraftnews.org/en/education/curriculum-resources/>, accessed October 27, 2018; see also Alfred Hermida, “Filtering Fact from Fiction: A Verification Framework for Social Media,” in *Ethics for Digital Journalists: Emerging Best Practices*, ed. Lawrie Zion and David Craig (New York: Routledge, 2015).

ACKNOWLEDGMENTS

The ideas in this book have been simmering for a while, some since as far back as 2006. As I've worked through the material over the years, and particularly in 2017 and 2018 as I wrote this book, I've gotten a lot of help along the way. Colleagues, friends and family, mentors, teachers, students, and editors have all played important roles in shaping the project.

As a computer scientist trained at the Georgia Institute of Technology, I never thought I would actually *write* a book. It's not something that computer scientists typically do. But here we are. Although I had been thinking about it off and on, it wasn't until 2016 that I jumped in wholeheartedly. Sarah Oates deserves special thanks for encouraging me, as does Ben Shneiderman, who brought me into the fold on his own book project and showed me it wouldn't be so daunting. And it was Irfan Essa, who sidled up to my desk in 2006 and remarked something like, "I just got back from CNN and we started talking about this idea of *computational journalism*. We should talk about what that might mean." Without that spark I probably never would have started down the path that led to this book. Thanks, Irfan. Over the years there were others who encouraged me too, in both big and small ways, and I'm appreciative.

On the specific development of this book I'm particularly thankful for early help from Karin Assmann, a research assistant who identified resources, interview candidates, and other leads for the book. Speaking of interviewees, they also deserve a special acknowledgment since this book never would have happened without their generosity in spending their time talking to me. I'm also thankful to the scholars and students who offered valuable feedback on various draft versions of chapters, including Peng Ao, Scott Cambo, Diego Gómez-Zarà, Hanlin Li, Carl-Gustav Linden, Wiebke Loosen, Jacob Nelson, Chelsea Peterson-Salahuddin, Cornelius Puschmann, Jan Schmidt, Daniel Trielli Paiva da Silva, Nikki Usher, Nicholas Vincent, and Yixue Wang. And I'm indebted to my editor, Jeff Dean, who taught me a thing or two about writing a book as a crossover. Finally, the peer reviewers of this book had a number of insightful comments that really pushed me to make the book better. Thank you!

I've been lucky to have the intellectual support of many fine institutions over the years. Leaders at both the Tow Center for Digital Journalism at Columbia University and the Tow-Knight Center for Entrepreneurial Journalism at CUNY were early believers in the value of thinking critically and constructively about the role of computation in the news media. I learned a lot during the fellowships I spent at both places. I am also grateful to the University of Bergen in Norway, which saw something in my work perhaps before even I had fully grasped it. A particular thanks goes to Dag Elgesem there for beginning an intellectual exchange with the Nordic countries that is impressed in this book. And thanks to Astrid Gynnild for both continuing that exchange and helping to financially support some of my time writing the book in the context of her VisMedia project. I've also been lucky to have some of the research

featured in the book financially supported by institutions including the Knight Foundation and the National Science Foundation. I'm grateful to the University of Maryland, College Park, for hosting me as I began this book and to Northwestern University for supporting me as I finished the book—the vibrancy of the campus community dedicated to computational journalism at NU is inspiring and nourishing every day.

I couldn't have written this text without the support of my family and friends. My father was a longtime visual journalist and editor. Having grown up in a household where there was a newspaper on the kitchen table every morning and where family dinner conversation frequently steered toward issues in the news, perhaps it's no surprise that I was primed to care about journalism. At home now I'm grateful for the support of Risa Chubinsky, who has over the years been a reader, critic, and sounding board for ideas while also helping provide a balance in life. And I have to thank Minerva (Minnie), my thirteen-pound Siberian forest cat, for “big hugs” before (and often interjected between) intense bouts of writing.