Ethics of Big Data
DS 6002, Fall 2022

Doing Good Data Science

What does it mean to do good data science? It means more than having skills in math, statistics, and programming. It means thinking carefully about how to use data to serve society. Good data science relies on careful consideration of what you're doing with data, why, and what your work might mean for society. In this course, we'll investigate the perils and promises that data science offers, such as the ongoing erosion of privacy, the dangerous perpetuation of our biases and discrimination through algorithmic decision-making, the life-saving insights gained from big data in epidemiology, and the opportunities to hold governments accountable and demand policy improvements thanks to open civic data. We'll apply theories from ethics and social science to interpret how data science and society affect each other and, crucially, make arguments for how we can best use data science to achieve social benefits. In particular, we'll evaluate your role in data-inspired ethical dilemmas and help you prepare for your ethical responsibilities as professional data scientists. You're learning how to work with data in your other courses; in this course, you'll learn how to make decisions about that work in ways that evaluate its social implications and optimize its social good.

What You'll Learn

At the end of this course, you should be able to:

- Understand data as a product of human work and decisions
- Identify and evaluate ethical dilemmas in data science by applying analytical frameworks and theories
- Choose and justify practical responses to ethical dilemmas
- Take responsibility for the social implications of your work as a data scientist

Whom You'll Learn With: Renée Cummings

I am an AI Ethicist, Data Activist and Criminologist and the first Data Activist in Residence, at UVA's School of Data Science. More info here: https://datascience.virginia.edu/people/renee-cummings

Contact info: xme4kh@virginia.edu,
Office hours: Email me to make an appointment.

What You'll Need

- Access to Collab for all our readings and assignments

How This Class Will Work

On Mondays, I'll give you a high-level synthesis of the week's topic. Next, 2-3 students will serve as discussion leaders by posing open-ended questions for the class to discuss. On

Wednesdays, we'll continue Monday's discussion and do an activity, or a group of students will present a case study. See the schedule below for more details.

How You'll Learn

Participation (25%): Ethics is best learned in a community. Your success at becoming an ethical data scientist depends on your preparation, presence, and willingness to share your ideas in class. To earn participation points, you'll answer weekly reading reflection questions (10%), complete in-class assignments and activities (10%) and serve as a discussion leader for one class (10%).
Group presentation (20%): In a group of 3-4 students, you'll present a data-related ethical dilemma of your choice to the class. Together, and in 15-20 minutes, you'll explain the dilemma, apply an analytical framework to help us understand it, and make an argument for the best way to address the dilemma. Then you'll respond to questions from your audience.
Critical review (20%): There's a lot going on in the field of data ethics. Please choose a high-quality resource—a book, a government report, a journal paper, a video of an expert panel discussion, etc.—on a data ethics topic that you'd like to learn more about. Be careful – there are a lot of problematic resources about data ethics. Consider the Recommended Sources listed on this syllabus or send me another resource by Sept. 28 so I can approve it for you. Then write a critical review (1000 words maximum) of your resource that summarizes and analyzes its argument. Due by 5pm on Oct. 10.
Ethical analysis assignment (30%): In a 1500-word (maximum) essay or a 15-minute (maximum) podcast/video, draw from our discussions and readings to identify, analyze, and propose a response to two ethical dilemmas drawn from your own experience in data science, such as your capstone project, coursework, job experiences, side projects, or other situations in which you've worked with data. You may work with a group (4 or fewer students please). Due at the end of our assigned final exam slot (we won't have an exam or meet during finals week).

Academic Integrity

I trust you to follow UVA's Honor Code. That means being honest with your classmates and me, as well as not giving or receiving unauthorized aid on assignments. You are welcome to give and receive critique on assignments for this class, by which I mean exchanging ideas and general suggestions. This is a powerful way to learn, so I define it as "authorized aid." You *may not* edit someone else's work, meaning make specific changes to it, nor have someone edit your work. For example, a classmate may read a draft of your critical review and tell you that you overlooked an important source, the third paragraph is confusing, your overall argument is unclear, and there are typos in the conclusion. That classmate *may not* contribute an interpretation of that source to your review, rewrite that confusing paragraph, add a thesis statement to clarify your argument, or correct your typos. Collaboration is valuable but doing each other's work does not lead to learning. To improve your work, please talk to me or the UVA Writing Center – only we are allowed to critique *and* edit your work. Also, if you claim that someone else's words or ideas are your own, you are plagiarizing. Plagiarism is both cheating and stealing, and therefore is an honor violation.

Information on plagiarism: https://honor.virginia.edu/plagiarism-supplement
Information on the Honor System: http://www.virginia.edu/honor/overview/

How to Succeed in This Class

- I recognize and value the many perspectives that you bring to the classroom. Many factors—social identities, visible and invisible disabilities, family circumstances, mental health, etc.—influence the experiences that each person has in class this and every semester. I am committed to building an environment to support your learning, one that encourages you to go out on a limb to communicate and defend your ideas so that we all learn from each other.
- Communication is key. Let me know how I can help you do your best work. If you have a disability or think you may have a disability, consider meeting with the Student Disability Access Center (SDAC, https://studenthealth.virginia.edu/sdac).
- Read the assigned sources carefully before class. If you haven't read and understood the readings, class will be boring for all of us.
- This is a discussion-based course. Anything that prevents you from actively listening and speaking in class is bad for you and for the rest of us. Help yourself learn by reducing distractions.
- Take notes. Discussions only happen once, and you'll want to remember them later. Studies show that writing things down helps you pay attention and remember information, even if you never read your notes.
- Turn in assignments on time. You know when the due dates are (hint: they're listed on the schedule below and on Collab), so plan accordingly. If you need an extension, ask.
- Things happen that cause you to miss class. I trust you to make good decisions about those situations, including not attending class when you have COVID-like symptoms. Attending class is crucial for your learning, so let me know if you need help catching up when absent.
- My office hours are for you. Please use them to discuss your questions and concerns with me. If you can't come at the scheduled times, email me to schedule a meeting.
- I am committed to reducing violence, harassment, and discrimination at UVA and in our community. Every person can play a part in reducing these problems. For more information, visit UVA's Office of Equal Opportunity and Civil Rights: https://eocr.virginia.edu

What We'll Do

Schedule			
Week	**Date**	**Topics**	**Activities due before class**
[1]	8/24	What are data? Why do they matter for ethics?	**Reading:** Syllabus **Suggested reading:** Kitchin 2014, Ch. 1, Conceptualizing Data

Schedule			
[2]	8/29 & 8/31	What is ethics? Why does it matter for data science?	**Reading:** Johnson 2020, chapter 3 (especially pp. 53-69); White and Arp 2008 **Suggested reading:** Specia 2020, assigning exam grades; Metcalf et al. 2019, Owning Ethics (under the Resources tab); Valor et al. 2020, more ethics theories **In class:** Choose presentation groups and topics. **Assignment:** Suggest ideas for group presentation topics in this discussion forum on Collab.
[3]	9/5 & 7	How do people create data? How should people create data?	**Reading:** Leonelli 2019, Data Governance; Meng 2019, Data Science; Leonelli 2019 (video), COVID data collection **Suggested reading:** Leonelli 2019, Data: from objects to assets **Discussion leader:**
[4]	9/12 & 14	How do data travel?	**Reading:** Borgman 2019, The Lives and Afterlives of Data; Leonelli and Tempini (2020): read Leonelli's intro ("Learning from data journeys" - focus on sections 1, 2, and 5) and then choose another chapter to read as an example of a data journey. **Suggested reading:** All the other chapters in Leonelli and Tempini (2020) **Discussion leader:** **Group presentation:**

Schedule			
[5]	9/19 & 21	Diversity, equity, and inclusion in data science	**Reading:** Schellmann 2021, bias in hiring algorithms **Suggested reading:** Johnson 2020, Ch. 8, Are engineers responsible for social justice? **Discussion leader:** **Group Presentation:** **Assessment:**
[6]	9/26 & 28	What does data science have to do with value(s)?	**Reading:** Meng 2021, What Are the Values of Data, Data Science, or Data Scientists?; choose one additional article to read from this special issue (the 9 papers listed under "Featured Discussion") **Suggested reading:** The other articles in this special issue **Discussion leader:** **Group Presentation:** **Assessment:** Check with me about your chosen resource for the critical review assignment.
[7]	10/3 & 5 **No Class on 10/03**	How should people make data into knowledge to inform policy? How should data inform policy in criminal justice?	**Reading:** ProPublica on recidivism algorithms (Angwin et al. 2016); the methods and data ProPublica used for this study (Larson et al. 2016); the Washington Post's ongoing count of people killed by police (Fatal Force 2020) **Suggested Reading:** O'Neil 2016, Ch. 5, predictive policing; the WaPo's description of their database methods (Tate et al. 2016)(The PDF is in Resources if you hit a paywall.) **Discussion leader:** **Group Presentation:** **Assignment:** Critical review due by 5pm on 10/10

Schedule			
[8]	10/10 & 12 **10/12 Guest Speaker**	Who owns data? What does it mean for data or data products to be "open"?	**Reading:** Wilkinson et al. 2016 on FAIR; Perens 1999 on open source software (skim everything and carefully read the section titled "Analysis of the Open Source Definition"); CHOOSE to read either Prainsack 2019 or Taddeo 2016 **Suggested reading:** Zuboff 2015 on surveillance capitalism **Discussion leader:** **Group Presentation:** **Assessment:**
[9]	10/17 & 19	How should we ethically assess algorithmic decision-making?	**Reading:** Zarsky 2016 and Taddeo and Floridi 2018 on analytical frameworks; Eubanks 2017 Ch. 5 on inequity in algorithmic policymaking **Discussion leader:** Drew **Group Presentation:** Krissy, Pat, "Facial recognition" (will present 2-2:30) **Assessment:**
[10]	10/24 & 26	TBA	**Reading:** **Suggested Reading:** **Discussion leader:** **Group Presentation:**
[11]	10/31 & 11/02 **10/31** **Guest Speaker**	Who is responsible for autonomous systems?	**Reading:** Awad et al. 2018 and MIT's Moral Machine scenario tool; Ananny and Crawford 2016 on transparency **Suggested Reading:** A critique published of Awad et al. 2018 (Bigman and Gray 2020), and Awad and coauthors' (2020) reply **Discussion leader:** **Group Presentation:**

Schedule			
[12]	11/7 & 09	The ethics of whistleblowing and data visualization: the case of COVID dashboards	**Reading:** Iati 2020 on data scientist and whistleblower Rebekah Jones, Florida's COVID dashboard; Jones' COVID dashboard; Vasquez 2020 on colleges' COVID dashboards **Suggested Reading:** Oliver 2003 on whistleblowing, Mazzei 2020 on updates on Jones' case **Discussion leader:** **Group Presentation:** **Assessment:**
[13]	11/14 & 16	TBA	**Guest Speakers**
[14]	11/21 11/23 - no class, Happy Thanksgiving!	TBA	
[15]	11/28 &	TBA	
[16]	12/5 (LAST CLASS)	Should there be a code of ethics for data science?	**Reading:** Johnson 2020 ch 2; Zook et al. 2017; Wylie 2020 (yes, this Wylie) **Discussion leader:** **Group Presentation:** **Assessment:** Course evaluation due (Evals are really important for me and for SDS. We would be grateful for your input. Thank you!)
	Our assigned final exam slot (TBA)		Ethical analysis due (we won't meet or have an exam)

Grading Scale

97.0-100	A+
93.0-96.9	A

90.0-92.9	A-
87.0-89.9	B+
83.0-86.9	B
80.0-82.9	B-

etc.

References for Course Readings

Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media and Society* 20 (3): 973–89. https://doi.org/10.1177/1461444816676645.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." *ProPublica*, May 23, 2016.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean François Bonnefon, and Iyad Rahwan. 2018. "The Moral Machine Experiment." *Nature* 563 (7729): 59–64. https://doi.org/10.1038/s41586-018-0637-6.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2020. "Reply to: Life and Death Decisions of Autonomous Vehicles." *Nature* 579 (7797): E3–5. https://doi.org/10.1038/s41586-020-1988-3.

Bigman, Yochanan E., and Kurt Gray. 2020. "Life and Death Decisions of Autonomous Vehicles." *Nature* 579 (7797): E1–2. https://doi.org/10.1038/s41586-020-1987-4.

Claibourn, Michele et al. 2019. "Charlottesville Foster Care Study." https://doi.org/10.18130/v3-ztey-hn31.

Crawford, Kate, and Vladan Joler. 2018. ""Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources." AI Now Institute and Share Lab. 2018. https://doi.org/10.1386/vcr_00008_7.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police and Punish the Poor*. St. Martin's Press.

Iati, Marisa. 2020. "Florida Fired Its Coronavirus Data Scientist. Now She's Publishing the Statistics on Her Own." *The Washington Post*, June 16, 2020. https://www.washingtonpost.com/nation/2020/06/12/rebekah-jones-florida-coronavirus/.

Johnson, Deborah G. 2020. *Engineering Ethics*. New Haven, CT: Yale University Press.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures, and Their Consequences*. London: SAGE Publications Ltd.

Larson, Jeff, Surya Mattu, Lauren Kirchner, and Julia Angwin. 2016. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, May 23, 2016.

Leonelli, Sabina. 2019. "Data - from Objects to Assets." *Nature* 574 (7778): 317–20. https://doi.org/10.1038/d41586-019-03062-w.

Leonelli, Sabina. 2019. "Data Governance Is Key to Interpretation: Reconceptualizing Data in Data Science." *Harvard Data Science Review*, 1–9. https://doi.org/10.1162/99608f92.17405bb6.

Leonelli, Sabina. 2020. "History of Science ON CALL: Sabina Leonelli." Max Planck Institute for the History of Science. 2020. https://www.mpiwg-berlin.mpg.de/video/history-science-call-sabina-leonelli.

Leonelli, Sabina, and Niccolò Tempini, eds. 2020. *Data Journeys in the Sciences*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7.

Metcalf, Jacob. 2019. "Owning Ethics: Corporate Logics." *An International Quarterly* 82 (2): 449–76. https://muse.jhu.edu/article/732185.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.

O'Neil, Cathy, and Rachel Schutt. 2014. *Doing Data Science*. Sebastopol, CA: O'Reilly Media, Inc.

Oliver, Douglas. 2003. "Whistle-Blowing Engineer" 129 (4): 246–56.

Perens, Bruce. 1999. "The Open Source Definition." In *Open Sources: Voices from the Open Source Revolution*. O'Reilly Media, Inc. https://www.oreilly.com/openbook/opensources/book/perens.html.

Prainsack, Barbara. 2019. "Logged out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons." *Big Data and Society* 6 (1): 1–15. https://doi.org/10.1177/2053951719829773.

Specia, Megan. 2020. "Parents, Students and Teachers Give Britain a Failing Grade Over Exam Results." *New York Times*, August 14, 2020.

Taddeo, Mariarosaria. 2016. "Data Philanthropy and the Design of the Infraethics for Information Societies." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083). https://doi.org/10.1098/rsta.2016.0113.

Taddeo, Mariarosaria, and Luciano Floridi. 2018. "How AI Can Be a Force for Good." *Science* 361 (6404): 751–53. http://arxiv.org/abs/1606.05718.

Tate, Julie, Jennifer Jenkins, Steven Rich, John Muyskens, Kennedy Elliott, Ted Mellnik, and Aaron Williams. 2016. "How The Washington Post Is Examining Police Shootings in the United States." *The Washington Post*, July 7, 2016. https://www.washingtonpost.com/national/how-the-washington-post-is-examining-police-shootings-in-the-united-states/2016/07/07/d9c52238-43ad-11e6-8856-f26de2537a9d_story.html.

The Washington Post. 2020. "Fatal Force," December 14, 2020.

Vallor, Shannon, Irina Raicu, and Brian Green. 2020. "Technology and Engineering Practice: Ethical Lenses to Look Through." Markkula Center for Applied Ethics. 2020.

Vasquez, Michael. 2020. "A Tool to Inform Too Often Confuses." *The Chronicle of Higher Education*, September 18, 2020.

White, Mark D., and Robert Arp. 2008. "Should Batman Kill the Joker?" *New York Times*, July 25, 2008.

Wilkinson, Mark D, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 1–9. https://doi.org/10.1038/sdata.2016.18.

Wylie, Caitlin D. 2020. “Who Should Do Data Ethics?” *Patterns* 1 (1): 100015. https://doi.org/10.1016/j.patter.2020.100015.

Zarsky, Tal. 2016. “The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making.” *Science Technology and Human Values* 41 (1): 118–32. https://doi.org/10.1177/0162243915605575.

Zigon, Jarrett. 2019. “Can Machines Be Ethical? On the Necessity of Relational Ethics and Empathic Attunement for Data-Centric Technologies.” *Social Research* 86 (4): 1001–22.

Zook, Matthew, Solon Barocas, danah boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, et al. 2017. “Ten Simple Rules for Responsible Big Data Research.” *PLoS Computational Biology* 13 (3): 1–10. https://doi.org/10.1371/journal.pcbi.1005399.

Zuboff, Shoshana. 2015. “Big Other: Surveillance Capitalism and the Prospects of an Information Civilization.” *Journal of Information Technology* 30: 75–89. https://doi.org/10.1057/jit.2015.5.

Recommended Sources

These are sources that offer more detail about particular topics, in addition to what we'll read for class. Feel free to use these sources to inform your group presentation or as the basis for your critical review. This is an incomplete list of excellent scholarship on data ethics, of course. If you find other interesting sources, please share them with me.

Acquisti, A, and R Gross. 2009. "Predicting Social Security Numbers from Public Data." *Proceedings of the National Academy of Sciences* 106 (27): 10975–80.

Agre, Philip E. 1994. "Surveillance and Capture: Two Models of Privacy." *The Information Society* 10 (2): 101–27.

Ananny, Mike, and Kate Crawford. 2018. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media and Society* 20 (3): 973–89. https://doi.org/10.1177/1461444816676645.

Ananny, Mike. 2016. "Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness." *Science, Technology, & Human Values* 41 (1): 93–117.

Barassi, V. 2020. *Child Data Citizen: How Tech Companies Are Profiling Us from before Birth.* MIT Press.

Barocas, Solon, and Andrew Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104 (3): 671–732.

Bellamy, Rachel K E, and Et Al. 2018. "AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias." *ArXiv*.

Benjamin, Ruha. 2019. *Race After Technology.* Polity Press.

Bolukbasi, Tolga, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings." *ArXiv*.

Borgman, Christine L. 2019. "The Lives and After Lives of Data." *Harvard Data Science Review*, 1–10. https://doi.org/10.1162/99608f92.9a36bdb6.

Borgman, Christine L. 2015. "Big Data, Little Data, No Data." *MIT Press*, 2015.

boyd, danah, Karen Levy, and Alice E Marwick. 2014. "The Networked Nature of Algorithmic Discrimination." *Data and Discrimination*, 53–57.

Brunton, Finn, and Helen Nissenbaum. 2016. *Obfuscation: A User's Guide for Privacy and Protest*. Cambridge, MA: MIT Press.

Burrell, Jenna. 2016. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data and Society*, 1–12. https://doi.org/10.1177/2053951715622512.

Coleman, E Gabriella. 2012. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton University Press.

Courtland, Rachel. 2018. "Bias Detectives: The Researchers Striving to Make Algorithms Fair News-Feature." *Nature* 558 (7710): 357–60. https://doi.org/10.1038/d41586-018-05469-3.

Crawford, Kate, and danah boyd. 2012. “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information, Communication & Society* 15 (5): 662–79.

Crawford, Kate, and Vladan Joler. 2018. ““Anatomy of an AI System: The Amazon Echo As An Anatomical Map of Human Labor, Data and Planetary Resources.” AI Now Institute and Share Lab. 2018. https://doi.org/10.1386/vcr_00008_7.

Davies, T, S Walker, M Rubinstein, and F Perini, eds. 2019. “Introduction.” In *The State of Open Data: Histories and Horizons*, 8–12. Cape Town and Ottawa: African Minds and International Development Research Centre. https://doi.org/10.1017/CBO9781107415324.004.

Ebeling, Mary F.E. 2016. *Healthcare and Big Data: Digital Specters and Phantom Objects*. New York: Palgrave Macmillan US. https://doi.org/10.1057/978-1-137-50221-6.

Edwards, Paul. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. MIT Press.

Edwards, Paul, Matthew Mayernik S., Archer L Batcheller, Geoffrey C Bowker, and Christine L Borgman. 2011. “Science Friction: Data, Metadata, and Collaboration.” *Social Studies of Science* 41: 667–90.

Enserink, Martin, and Gilbert Chin. 2015. “The End of Privacy.” *Science* 347 (6221): 490–91.

Ensmenger, Nathan L. 2010. *The Computer Boys Take over: Computers, Programmers, and the Politics of Technical Expertise.* MIT Press.

Favaretto, Maddalena, Eva De Clercq, and Bernice Simone Elger. 2019. “Big Data and Discrimination: Perils, Promises and Solutions. A Systematic Review.” *Journal of Big Data* 6 (12): 1–27. https://doi.org/10.1186/s40537-019-0177-4.

Fleming, L, N Tempini, G Nichols, C Sarran, P Vineis, B Golding, A Haines, et al. 2017. “Big Data in Environment and Human Health: Challenges and Opportunities.” *Oxford Encyclopedia of Environment and Human Health*. f

Ford, Martin. 2018. *Architects of Intelligence: The Truth about AI and the People Building It*.

Hogle, Linda F. 2016. “Data-Intensive Resourcing in Healthcare.” *BioSocieties* 11 (3): 372–93. https://doi.org/10.1057/s41292-016-0004-5.

Johnson, Deborah G. 2020. *Engineering Ethics*. New Haven, CT: Yale University Press.

Kelty, Christopher M. 2008. “Two Bits: The Cultural Significance of Free Software.” *Duke University Press*, 2008.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures, and Their Consequences*. London: SAGE Publications Ltd.

Kitchin, Rob, and Gavin McArdle. 2016. “What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets.” *Big Data and Society* 3 (1): 1–10. https://doi.org/10.1177/2053951716631130.

Kudina, Olya, and Peter-Paul Verbeek. 2019. “Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy.” *Science, Technology, & Human Values* 44 (2): 291–314. https://doi.org/10.1177/0162243918793711.

Leach, James, Dawn Nafus, and Bernhard Krieger. 2009. “Freedom Imagined: Morality and Aesthetics in Open Source Software Design.” *Ethnos* 74 (1): 51–71.

Leonelli, Sabina. 2016. “Locating Ethics in Data Science: Responsibility and Accountability in Global and Distributed Knowledge Production Systems.” *Phil.Trans.R. Soc. A* 374. https://doi.org/10.1098/rsta.2016.0122.

Leonelli, Sabina. 2019. “Data - from Objects to Assets.” *Nature* 574 (7778): 317–20. https://doi.org/10.1038/d41586-019-03062-w.

Leonelli, Sabina. 2019. “Data Governance Is Key to Interpretation: Reconceptualizing Data in Data Science.” *Harvard Data Science Review*, 1–9. https://doi.org/10.1162/99608f92.17405bb6.

Magnet, Shoshana. 2011. *When Biometrics Fail: Gender, Race, and the Technology of Identity*. Durham, NC: Duke University Press.

Metcalf, Jacob. 2019. “Owning Ethics: Corporate Logics.” *An International Quarterly* 82 (2): 449–76. https://muse.jhu.edu/article/732185.

Metcalf, Jacob, and Kate Crawford. 2016. “Where Are Human Subjects in Big Data Research? The Emerging Ethics Divide.” *Big Data and Society*, 1–14. https://doi.org/10.1177/2053951716650211.

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. “The Ethics of Algorithms: Mapping the Debate.” *Big Data and Society* 3 (2): 1–21. https://doi.org/10.1177/2053951716679679.

Noble, S. 2018. “Algorithms of Oppression: How Search Engines Reinforce Racism.” *New York University Press*, 2018.

O’Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.

Pasquale, Frank. 2015. “The Black Box Society: The Secret Algorithms That Control Money and Information.” *Harvard University Press*, 2015.

Plunkett, L. A. (2019). *Sharenthood: Why We Should Think before We Talk about Our Kids Online*. Cambridge, MA: MIT Press.

Prainsack, Barbara. 2019. “Logged out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons.” *Big Data and Society*, 1–15. https://doi.org/10.1177/2053951719829773.

Richards, M, R Anderson, S Hinde, J Kaye, A Lucassen, P Matthews, M Parker, et al. 2015. “The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issue.” *Nuffield Council on Bioethics*, 2015.

Roberts, Sarah T. 2019. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.

Rouvroy, Antoinette, and Thomas Berns. 2013. Algorithmic Governmentality and Prospects of Emancipation. Disparateness as a Precondition for Individuation through Relationships? *Reseaux*. Vol. 177. https://doi.org/10.3917/res.177.0163.

Selbst, Andrew D, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. "Fairness and Abstraction in Sociotechnical Systems." *Conference on Fairness, Accountability, and Transparency*, 59–68. https://doi.org/10.1145/3287560.3287598.

Selinger, Evan, and Woodrow Hartzog. 2016. "Facebook's Emotional Contagion Study and the Ethical Problem of Co-Opted Identity in Mediated Environments Where Users Lack Control." *Research Ethics* 12 (1): 35–43. https://doi.org/10.1177/1747016115579531.

Starosielski, N. 2015. *The Undersea Network*. Duke University Press.

Taddeo, Mariarosaria, and Luciano Floridi. 2018. "How AI Can Be a Force for Good." *Science* 361 (6404): 751–53. http://arxiv.org/abs/1606.05718.

Taylor, Linnet. 2017. "What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally." *Big Data and Society*, 1–14. https://doi.org/10.1177/2053951717736335.

Dijck, José van. 2014. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *Surveillance and Society* 12 (2): 197–208. https://doi.org/10.24908/ss.v12i2.4776.

Webb, M. 2020. *Coding Democracy: How Hackers Are Disrupting Power, Surveillance, and Authoritarianism*. MIT Press.

Wilkinson, Mark D, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 1–9. https://doi.org/10.1038/sdata.2016.18.

Zarsky, Tal. 2016. "The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making." *Science Technology and Human Values* 41 (1): 118–32. https://doi.org/10.1177/0162243915605575.

Zigon, Jarrett. 2019. "Can Machines Be Ethical? On the Necessity of Relational Ethics and Empathic Attunement for Data-Centric Technologies." *Social Research* 86 (4): 1001–22.

Zittrain, Jonathan. 2019. "The Hidden Costs of Automated Thinking." *The New Yorker*, 1–6. https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking.

Zook, Matthew, Solon Barocas, danah boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, et al. 2017. "Ten Simple Rules for Responsible Big Data

Research." *PLoS Computational Biology* 13 (3): 1–10. https://doi.org/10.1371/journal.pcbi.1005399.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30: 75–89. https://doi.org/10.1057/jit.2015.5.