Behind the Data: Humans and Values

Summer 2019

Instructors:	Morgan Ames	morganya@berkeley.edu	slack: @morganya
	Deirdre Mulligan	dkm@ischool.berkeley.edu	
Teaching Assistant:	Jared Maslin	jaredmaslin@gmail.com	slack: @mids_jared_m
Class Times:	Section W-231-1	Monday 4:00-5:30 PT	
	Section W-231-2	Tuesday 4:00-5:30 PT	

Course Pages:

1. Course Github: https://github.com/UC-Berkeley-I-School/w231

2. Course Blog: https://blogs.ischool.berkeley.edu/w231/

3. Course Slack Channel: #w231-announce

4. Course Signup sheets:

   - Introduction Survey: https://goo.gl/forms/Yht3jSSFUxXT8hKq2
   - Public Advocacy Blog Assignment Sign-Ups: https://goo.gl/6uZY62
   - Final Project Presentation Sign-Ups: https://goo.gl/krXaCd
   - Anonymous Course Feedback: https://forms.gle/yxEHyfshm8FGLgqz8

Office Hours:
Dr. Ames: The 30 minutes after class or by appointment in Zoom.

Course Description: This course provides an introduction to the legal, policy, and ethical implications of data. The course will examine how these issues arise throughout the full life cycle of data science from collection, to storage, processing, analysis and use including, privacy, surveillance, security, classification, discrimination, decisional autonomy, and duties to warn or act. Case studies will be used to explore these issues across various domains such as criminal justice, national security, health, marketing, politics, education, automotive, employment, athletics, and development. Attention will be paid to legal and policy constraints and considerations that attach to specific domains as well as particular data-types, collection methods, and institutions. Technical, legal, and market approaches to mitigating and managing discrete and compound sets of concerns will be introduced, and the strengths and benefits of competing and complementary approaches will be explored.

Prerequisites: Students must complete MIDS courses 201 and 203 before enrolling in this course.

Academic Calendar: https://www.ischool.berkeley.edu/intranet/students/mids/calendar

Class Calendar:

SECTION I INTRODUCTION TO ETHICAL PRINCIPLES

May 6-7	Unit 1	Information Policy and Research Ethics: Foundations	(Ames)
		DUE Friday, May 10: Survey, Blog Post Sign-ups	
May 13-14	Unit 4	Intent, Consent, and Research Design	(Ames)
May 20-21	Unit 2	Philosophy, Ethics, and Countercultures of Science	(Ames)
		DUE Friday, May 24: Application of Belmont Principles	
5/28, 6/17	Unit 3	Why Do Categories Matter? Sociocultural Considerations	(Ames)

SECTION II BIASES, RIGHTS OF INDIVIDUALS, AND NEW KNOWLEDGE

June 3-4	Unit 5	Principles of Privacy	(Mulligan)
June 10-11	Unit 6	Data Subjects' Rights, Rights of Others, and Institutional Obligations and Duties	(Mulligan)
		DUE Friday, June 21: Privacy Policy Assignment (slides, doc)	
June 24-25	Unit 7	In-class Presentations of Privacy Policy Assignment	(Ames)
		DUE Friday, June 28: Final Project Outline	
July 1-2	Unit 8	Biases In Data and Algorithms	(Ames)
		DUE Friday, July 5: Peer Feedback Assignment	

SECTION III MITIGATION AND MANAGEMENT

July 8-9	Unit 10	Due Process and Big Data	(Ames)
		DUE Friday, July 12: Final Project Presentation Sign-ups	
July 15-16	Unit 12	In-class Presentations of Group Legal/Ethical Analysis	
		DUE Friday, July 19: Group Legal/Ethical Analysis (slides, doc)	
July 22-23	Unit 11	Copyright and Other Protections	(Ames)

FINAL PROJECTS

July 26-27	NONE	Final Project Presentations (EXTENDED SESSION)	
Aug 5-6		**NO CLASS - final project completion**	
		DUE Friday, August 9: Final Project (slides, doc)	

Assignments: Descriptions are in the `Assignments` folder in the class git repository. All assignments are due 11:59pm Pacific Time (PT) on the date noted, in ISVC unless otherwise indicated in the assignment description. If you cannot make a deadline, contact us as soon as you can to make a plan for an alternate due date. We offer five floating late days penalty-free, but beyond those five days late assignments may be docked up to 5 points per day late.

Assignment Formatting: Please convert docx and pptx files to pdf.

Course Evaluation:

- Participation (25%)

  - Synchronous session discussion (15%). You are expected to attend and actively participate in every synchronous session.
  - Asynchronous participation (10%). This class relies on readings and asynch content. Both are necessary for participating in the synchronous session, and are important for the course. Complete both before the synchronous session each week. NOTE that completing asynch content includes answering the discussion questions in asynch. As an alternative to answering these questions, you may instead send the teaching team three *substantive* questions about the readings and asynch material via Slack each week, due by midnight the night before your synch class session.

- Assignments (40%)

  - Application of Belmont Principles (8%)
  - Privacy Policy Assignment (8%)
  - Peer Feedback Assignment (8%)
  - Group Legal/Ethical Analysis (8%)
  - Public Advocacy Blog Post (8%)

- Final Project (35%)

  - Final Project Outline (5%)
  - Final Project Write-Up (30%)

Statement of Diversity and Inclusion: As we discuss in this class, science (and knowledge creation more generally) is subjective and is historically built on a small subset of privileged voices. Integrating a diverse set of experiences is important for a more comprehensive understanding of data science.

In this class, we will make an effort to read papers from a diverse group of scholars and actively address questions of privilege, marginality, and inclusion in data science and beyond. I would moreover like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities (including race, gender, class, sexuality, age, religion, ability, etc.).

To help accomplish this:

- Please let me know in the initial survey what your preferred name and pronouns are.

- You should always strive to honor the diversity of your classmates.

- If you feel like your performance in the class is being impacted by your experiences outside of class, please come and talk with me or submit anonymous feedback via the w231 feedback survey (which may lead to me making an announcement to the class, if this is necessary to address your concerns). If you prefer to speak with someone outside of the course, the Academic Director of Data Science, the Assistant Dean of Academic Programs, and the I-School diversity and inclusion resource page are good places to start.

- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone, including me) that made you feel uncomfortable, please talk to me about it. (Again, anonymous feedback is always an option via the w231 feedback survey.

SECTION I INTRODUCTION TO ETHICAL PRINCIPLES

1 *Information Policy and Research Ethics: Foundations*

Weekly Readings and Asynch Unit 1

[1] Rebecca Skloot. "The Immortal Life of Henrietta Lacks, the Sequel". In: *The New York Times* (Mar. 2013). URL: http://www.nytimes.com/2013/03/24/opinion/sunday/the-immortal-life-of-henrietta-lacks-the-sequel.html.

[2] The Commission. *Appendix, the Belmont Report: ethical principles and guidelines for the protection of human subjects in research.* The Commission, 1978.

[3] *Records, Computers and the Rights of Citizens (SKIM)*. Government Printing Office, 1973. URL: http://www.justice.gov/sites/default/files/opcl/docs/rec-com-rights.pdf.

[4] *Henrietta Lacks Foundation.* URL: https://www.henriettalacksfoundation.org.

2 *Philosophy, Ethics, and Countercultures of Science*

Weekly Readings and Asynch Unit 2

[1] Sandra G. Harding. *Whose Science? Whose Knowledge?: Thinking from Women's Lives (chapter 1).* Cornell Univ. Press, 1991.

[2] Nathan Jurgenson. *View From Nowhere.* Apr. 2017. URL: http://thenewinquiry.com/essays/view-from-nowhere/.

[3] Kate Crawford. "Hidden Biases in Big Data". In: *Harvard Business Review* (2013). URL: https://hbr.org/2013/04/the-hidden-biases-in-big-data.

[4] Florangela Davila. "USDA disqualifies three Somalian markets from accepting federal food stamps". In: *The Seattle Times* (Apr. 2002). URL: http://community.seattletimes.nwsource.com/archive/?date=20020410&slug=somalis10m.

[5] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. "Experimental evidence of massive-scale emotional contagion through social networks". In: *Proceedings of the National Academy of Sciences* 111.24 (2014), pp. 8788–8790. DOI: 10.1073/pnas.1320040111. eprint: http://www.pnas.org/content/111/24/8788.full.pdf. URL: http://www.pnas.org/content/111/24/8788.abstract.

[6] Mark A Rothstein and Abigail B Shoben. "Does consent bias research?" In: *American Journal of Bioethics* 13.4 (Apr. 2013). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2244990.

3 *What Do Categories Matter? Sociocultural Considerations*

Weekly readings and Asynch Unit 3

[1] Guy Stuart. "Databases, Felons, and Voting: Errors and Bias in the Florida Felons Exclusion List in the 2000 Presidential Elections". In: *SSRN Electronic Journal* (2002). DOI: 10.2139/ssrn.336540.

[2]   Geoffrey C. Bowker and Susan Leigh Star. *Sorting Things Out: Classification and its Consequences (chapter 7)*. MIT Press, 1999.

[3]   D. Valentine. "The Categories Themselves". In: *GLQ: A Journal of Lesbian and Gay Studies* 10.2 (Jan. 2004), pp. 215–220. DOI: 10.1215/10642684-10-2-215.

[4]   David I. Kertzer. *Census and identity: the politics of race, ethnicity, and language in national censuses*. Cambridge University Press, 2002.

[5]   Mona Chalabi. *Why We Don't Know The Size Of The Transgender Population*. Mar. 2016. URL: https://fivethirtyeight.com/features/why-we-dont-know-the-size-of-the-transgender-population/.

[6]   Solon Barocas and Andrew D. Selbst. *Big Data's Disparate Impact*. Aug. 2014. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899.

4   *Intent, Consent, and Research Design*

Weekly Readings and Asynch Unit 4

[1]   *OK Cupid: We Experiment on Human Beings*. URL: Formerly%20blog.okcupid.com/index.php/we-experiment-on-human-beings/.

[2]   *Ok Cupid Study Reveals Perils of Big Data Science*. URL: https://www.wired.com/2016/05/okcupid-study-reveals-perils-big-data-science/.

[3]   *Ok Cupid Users Data Release*. URL: http://www.vox.com/2016/5/12/11666116/70000-okcupid-users-data-release.

[4]   John Kropf. *Public information and privacy in a global society*. Bloomberg BNA, Mar. 2014. URL: http://privacylaw.bna.com/pvrc/7057/split_display.adp?fedfid=43383440&vname=pvlrnotallissues&wsn=499496500&searchid=22477163&doctypeid=6&type=oadate4news&mode=doc&split=0&scm=7057&pg=0.

[5]   Office of Civil Rights. *OCR HIPAA Privacy Research*. Vol. 512. Washington: Health and Human Services, Dec. 2003. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html.

[6]   *In re Google Inc. Street View Electronic Communications Litigation: Radio Communications and Privacy by Convention*. June 2011.

[7]   David Kravets. "An intentional mistake: The anatomy of Google's Wi-Fi sniffing debacle". In: *Wired* (May 2012). URL: https://www.wired.com/2012/05/google-wifi-fcc-investigation/.

[8]   Paul M. Schwartz and Daniel J. Solove. "Reconciling Personal Information in the United States and European Union". In: *SSRN Electronic Journal* (May 2013). DOI: 10.2139/ssrn.2271442. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2271442.

[9]   State of California. *Chapter 22.1. Privacy rights for California minors in the digital world [22580-22582]*. 2014. URL: http://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=BPC%EF%BF%BD%EF%BF%BDionNum=22581.

[10]   Office of Civil Rights. *45 CFR 164.514 - Other Requirements Relating to Uses and Disclosures of Protected Health Information*. Health and Human Services, 2010.

[11]   Executive Office of the President. *Big Data: Seizing Opportunities, Preserving Values (excerpts)*. URL: https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

[12] Office of the Federal Register. *Children's Online Privacy Protection Rule ("COPPA"): final rule: 16 CFR part 312*. Office of the Federal Register, 2013. URL: https://www.law.cornell.edu/cfr/text/16/part-312.

SECTION II - BIASES, RIGHTS OF INDIVIDUALS, AND NEW KNOWLEDGE

5 *Principles of Privacy*

Weekly Readings and Asynch Units 5

[1] Iain Thomson. "Hello Barbie: Hang on, this Wi-Fi doll records your child's voice?" In: *The Register* (Feb. 2015). URL: http://www.theregister.co.uk/2015/02/19/hello_barbie/.

[2] Daniel J Solove. "Conceptualizing Privacy (NOTE: READ 1099 to 1123)". In: *California Law Review* 90.4 (July 2002). DOI: https://doi.org/10.15779/Z382H8Q.

[3] Daniel J Solove. "Taxonomy of privacy (NOTE: READ 479 to 483, 490 to 491)". In: *University of Pennsylvania Law Review* 154.3 (Jan. 2006).

[4] Helen Nissenbaum. "A Contextual Approach to Privacy Online". In: *Daedalus* 140.4 (2011), pp. 32–48. DOI: 10.1162/daed_a_00113.

[5] Deirdre K Mulligan, Colin Koopman, and Nick Doty. "Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy". In: *Phil. Trans. R. Soc. A* 374.2083 (2016), p. 20160118.

6 *Data Subjects' Rights, Rights of Others, and Institutional Obligations and Duties*

Weekly Readings and Asynch Units 6

[1] A. Acquisti, L. Brandimarte, and G. Loewenstein. "Privacy and human behavior in the age of information". In: *Science* 347.6221 (Jan. 2015), pp. 509–514. DOI: 10.1126/science.aaa1465.

[2] Alessandro Acquisti, Curtis R Taylor, and Liad Wagman. "The economics of privacy". In: *Journal of Economic Literature* 52.2 (2016).

[3] *Federal Trade Commission Act Section 5: Unfair or Deceptive Acts or Practices*. URL: https://www.federalreserve.gov/boarddocs/supmanual/cch/ftca.pdf.

[4] *[SKIM] OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. URL: http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflows htm.

[5] *GDPR Data Portability FAQ*. URL: http://ec.europa.eu/information_society/newsroom/image/document/2016-51/wp242_annex_en_40854.pdf.

[6] *GDPR Guidelines on the Right to Data Portability*. URL: http://ec.europa.eu/information_society/newsroom/image/document/2016-51/wp242_en_40852.pdf.

[7] *In re Facebook, Complaint, FTC File No. 092 3184*. 2012.

[8] *In the Matter of ELI LILLY and COMPANY, a corporation*. 2002.

7   *Science of Privacy*

Weekly readings and Asynch Unit 7

[1] Fida Kamal Dankar and Khaled El Emam. "[OPTIONAL] Practicing differential privacy in health care: A review". In: *Trans. Data Privacy* 6.1 (2013), pp. 35–67.

[2] Cynthia Dwork. "Differential Privacy". In: *33rd International Colloquium on Automata, Languages and Programming, part II (ICALP 2006)*. Vol. 4052. Venice, Italy: Springer Verlag, July 2006, pp. 1–12. ISBN: 3-540-35907-9. URL: https://www.microsoft.com/en-us/research/publication/differential-privacy/.

[3] Bernardo A Huberman, Eytan Adar, and Leslie R Fine. "Valuating Privacy". In: *IEEE Security & Privacy* (Oct. 2005), pp. 22–25. DOI: 10.1109/MSP.2005.137.

[4] Serge Egelman, Adrienne Porter Felt, and David Wagner. "Choice Architecture and Smartphone Privacy: There's a Price for That". In: *The Economics of Information Security and Privacy* (2013), pp. 211–236. DOI: 10.1007/978-3-642-39498-0_10.

[5] Janice Y. Tsai et al. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study". In: *Information Systems Research* 22.2 (June 2011), pp. 254–268. DOI: 10.1287/isre.1090.0260.

[6] Sarah Spiekermann, Jens Grossklags, and Bettina Berendt. "E-privacy in 2nd generation E-commerce". In: *Proceedings of the 3rd ACM conference on Electronic Commerce - EC 01* (Oct. 2001). DOI: 10.1145/501158.501163.

[7] Ryan Calo and Alex Rosenblat. "The Taking Economy: Uber, Information, and Power". In: (2017).

8   *Biases in Data and Algorithms*

Weekly Readings and Asynch Unit 8

[1] Tarleton Gillespie. "The Relevance of Algorithms". In: *Media Technologies* (2014), pp. 167–194. DOI: 10.7551/mitpress/9780262525374.003.0009.

[2] Felicitas Kraemer, Kees van Overveld, and Martin Peterson. "Is there an ethics of algorithms?" In: *Ethics and Information Technology* 13.3 (Sept. 2011), pp. 251–260.

[3] Jenna Burrell. "How the machine thinks: Understanding opacity in machine learning algorithms". In: *Big Data & Society* 3.1 (2016). DOI: 10.1177/2053951715622512.

[4] Tad Hirsch et al. "Designing Contestability: Interaction Design, Machine Learning, and Mental Health". In: *Proceedings of the 2017 Conference on Designing Interactive Systems* (2017), pp. 95–99. DOI: 10.1145/3064663.3064703.

[5] FTC. *Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues (READ: executive summary and questions for legal compliance)*. URL: https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report.

[6] Jonas Lehrman. "Big Data and Its Exclusions". In: *Stanford Law Review* 66 (55 2013). DOI: 10.2139/ssrn.2293765.

[7] Cynthia Dwork and Deirdre K Mulligan. "It's not privacy, and it's not fair". In: *Stanford Law Review* (Sept. 2013). URL: https://www.stanfordlawreview.org/online/privacy-and-big-data-its-not-privacy-and-its-not-fair/.

[8]   *United States of America (for the Federal Trade Commission) v. Spokeo Inc., Civ. No. CV12-05001*. 2012.

SECTION III - MITIGATION AND MANAGEMENT

9   *Due Process and Big Data*

Weekly Readings and Asynch Unit 10

[1]   NIST. *Privacy Framework Factsheet*. 2018. URL: https://www.nist.gov/sites/default/files/documents/2018/09/04/privacyframeworkfactsheet-sept2018.pdf.

[2]   Danielle Keats Citron. "Technical Due Process". In: *Washington University Law Review* 85 (2007). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360.

[3]   Kate Crawford and Jason Schultz. "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms". In: *Boston College Law Review* 55.93 (2014). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2325784.

[4]   Privacy Office. *Privacy Impact Assessment*. Department of Homeland Security, 2014.

[5]   TSA and Robin Kane. *Privacy Impact Assessment Update for TSA Whole Body Imaging*. Department of Homeland Security, 2009.

[6]   Joe Sharkey. "A Farewell to 'Nudity' at Airport Checkpoints". In: *The New York Times* (Jan. 2013). URL: http://www.nytimes.com/2013/01/22/business/a-farewell-to-nudity-at-airport-checkpoints.html?_r.

[7]   Katie Rogers. "T.S.A. Defends Treatment of Transgender Air Traveler". In: *The New York Times* (Sept. 2015). URL: https://www.nytimes.com/2015/09/23/us/shadi-petosky-tsa-transgender.html?_r=0.

[8]   Charles Raab and David Wright. "Surveillance: Extending the Limits of Privacy Impact Assessment". In: *Privacy Impact Assessment* (Sept. 2012), pp. 363–383. DOI: 10.1007/978-94-007-2543-0_17.

10   *Additional Legal Limitations for Data: Copyright, Contracts, and Database Rights*

Weekly Readings and Asynch Unit 11

[1]   *Facebook v. Vachani & Power Ventures*. July 2016. URL: https://cdn.ca9.uscourts.gov/datastore/opinions/2016/07/12/13-17102.pdf.

[2]   *Feist Pubs., Inc. v. Rural Tel. Svc. Co., Inc. 499 U.S. 340 (1991)*. 1991.

[3]   *ProCD, INC. v. ZEIDENBERG, 86 F.3d 1447 (7th. Cir. 1996)*. 1996.

[4]   *Copyright Act 102(a), 106*.

[5]   *CalECPA (state bill 178)*. URL: https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB178.

[6]   *Computer Fraud and Abuse Act, 18 USC*. URL: https://www.law.cornell.edu/uscode/text/18/1030.

11 *Organizational: Professionals and Professional Associations*

Weekly Readings and Asynch Unit 12

[1] Alex Fowler. *Do Not Track: It's the user's voice that matters*. May 2012. URL: https://blog.mozilla.org/netpolicy/2012/05/31/do-not-track-its-the-users-voice-that-matters/.

[2] *Tracking Preference Expression (DNT)*. URL: https://www.w3.org/TR/tracking-dnt/.

[3] Allison Grande. "Do-Not-Track Group Finally Nails Down Tech Standard". In: *Law360 - The Newswire for Business Lawyers* (2014). URL: https://www.law360.com/articles/531445.

[4] Elon Musk. *A Most Peculiar Test Drive*. Feb. 2013. URL: https://www.tesla.com/blog/most-peculiar-test-drive.

[5] Kashmir Hill. *The Big Privacy Takeaway From Tesla vs. The New York Times*. Feb. 2013. URL: http://www.forbes.com/sites/kashmirhill/2013/02/19/the-big-privacy-takeaway-from-tesla-vs-the-new-york-times/.

[6] Tesla Motors. *Privacy Statement*. June 2013. URL: https://www.tesla.com/sites/default/files/pdfs/tmi_privacy_statement_external_6-14-2013_v2.pdf.