

# LIS 640: Data Ethics

Information School  
University of Wisconsin-Madison  
Summer 2019

Dorothea Salo (please call me “Dorothea”)  
Office address: 4261 Helen C. White Hall  
Course link page: <https://pinboard.in/u:dsalo/t:dataethicscourse>

salo@wisc.edu  
Office hours: by appointment  
Special course attributes: none



To the extent possible under law, the person who associated CCo with this work has waived all copyright and related or neighboring rights to this work. This work is published from: United States.

## Course Description and Objectives

Students completing this course will earn one credit hour. One credit is the learning that takes place in at least 45 hours of learning activities, which include time in lectures or class meetings, in person or online, labs, exams, presentations, tutorials, reading, writing, studying, preparation for any of these activities, and any other learning activities.

This course has no prerequisites or co-requisites. Course objectives:

- ◆ Basic awareness of major Western schools of ethics and how to use their precepts to resolve ethical dilemmas
- ◆ Ability to define “big data” and recognize its use across industries
- ◆ Broad awareness of collection, retention, use, and transmission of big data about individuals and groups of people across industries
- ◆ Ability to recognize potential and actual harms in the collection, retention, use, and transmission of big data about individuals and groups of people, including but not limited to bias, loss of autonomy, loss of privacy, loss of free expression, manipulation, redlining, security breaches
- ◆ Ability to recognize and respond to common “ethics smells”—poor rationalizations/justifications—for collection, retention, use, and transmission of person-specific big data across industries

This course is designed to assess student progress in the following iSchool program-level learning outcomes: 1, 6, 7.

## Course Policies

**I aim to make this course as accessible as possible to all students. Students seeking accommodations must provide instructor with a McBurney Center VISA within the first week of class. For more information on obtaining a McBurney Center VISA, see <http://mcburney.wisc.edu/students/howto.php>.**

**Preferred name/pronouns:** It is sometimes the case that a student’s legal name or gender assigned at birth are reported to me on official documents in a form not in keeping with that student’s preferred name or gender expression. Please let me know, as you are comfortable, about your preferences. My pronouns are she/her/hers. UW-Madison also permits students to indicate a preferred name: [https://registrar.wisc.edu/preferred\\_name.htm](https://registrar.wisc.edu/preferred_name.htm)

**Academic Honesty:** I follow the academic standards for cheating and plagiarism set forth by the University of Wisconsin.

### Readings

There are no required textbooks or software purchases for this course. **THE SYLLABUS IS REQUIRED READING.**

### Contacting me

**READ THE SYLLABUS** before asking a question, please; the syllabus may answer it! If it does not, please ask in the class’s Canvas help forums. Please use email rather than the Canvas instructor-contact form; email me **ONLY** with confidential individual concerns or to set up a synchronous appointment. Any email question that could be answered on the forums will not receive a reply. I check class forums daily during the workweek, and do my best to answer email within two business days. I am not available weekends. If you can answer a classmate’s question in any Canvas forum, please do so.

If you see dead links (it does happen, usually with no notice), weird due dates, or other syllabus or Canvas problems, please post them to the “Syllabus/Canvas questions and problems” forum on Canvas. I will do my best to resolve them promptly. I do not allow my own mistakes to cause harm to anyone’s grade.

## Course week, due dates, and assignment policies

Our course “week” runs from Monday to Monday for convenience; course content will open Sunday, however. I try to open course content one week early for students who need to work ahead; if you need more at some point during the term, I will try to accommodate you. Late assignments will be penalized one final-grade percentage point per day or fraction thereof late. McBurney accommodations aside, I will allow revision/resubmission at my discretion and on my schedule only; student resistance will remove the opportunity.

## Module Readings

### Module 1: Schools-of-ethics speed-dating

*Topics: Consequentialism. Deontology. Virtue ethics. Confucian ethics. How each of these ethics schools approaches ethical dilemmas; how they can shore one another up. Professional ethics codes; how they come to be. “Ethics smells.” Workplace ethics; consequences of poor workplace ethics; ethics activism in workplaces.*

*“I think there’s incredible value to teaching people how to grapple with ethical questions even when shared ethics have not been clearly established. The same way there’s important value in learning to think critically even when there is no prescribed conclusion/opinion.” —Karen Hao, [https://twitter.com/\\_KarenHao/status/1102082083990749190](https://twitter.com/_KarenHao/status/1102082083990749190)*

Vallor, Green, Raicu. “Conceptual frameworks in technology and engineering practice.” <https://www.scu.edu/ethics-in-technology-practice/conceptual-frameworks/>

Jing, Doorn. “Engineers’ moral responsibility: a Confucian perspective.” <https://doi.org/10.1007/s11948-019-00093-4>  
“A framework for ethical decision making.” <https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/a-framework-for-ethical-decision-making/>

Floridi, Taddeo. “What is data ethics?” <http://dx.doi.org/10.1098/rsta.2016.0360>

Kampfe, Vallor. “On data ethics: an interview with Shannon Vallor.” <https://www.scu.edu/ethics/internet-ethics-blog/on-data-ethics-an-interview-with-shannon-vallor/>

Johnson, Lichfield, et al. “Hey, Google, sorry you lost your ethics council, so we made one for you.” <https://www.technologyreview.com/s/613281/google-cancels-ateac-ai-ethics-council-what-next/>

Warzel, Mac. “Mark Zuckerberg’s biggest problem.” <https://www.buzzfeednews.com/article/charliwarzel/facebooks-tensions-zuckerberg-sandberg>

Kowitz. “Inside Google’s civil war.” <http://fortune.com/longform/inside-googles-civil-war/>

Wong. “Uber’s ‘hustle-oriented’ culture becomes a black mark on employee résumés.” <https://www.theguardian.com/technology/2017/mar/07/uber-work-culture-travis-kalanick-susan-fowler-controversy>

**At least one of** (pick the one most apropos for you; whichever you choose, pay special attention to how and why ethics-related systems arise and where and how they fall short):

Witt. “The evolution of privacy within the American Library Association, 1906-2002.” <https://doi.org/10.1353/lib.2017.0022>

Vitak et al. “Ethics regulation in social computing research.” <https://doi.org/10.1177%2F1556264617725200>

Crichton. “The demise and rebirth of the ethical engineer.” <https://techcrunch.com/2019/04/28/the-demise-and-rebirth-of-the-ethical-engineer/>

Rogaway. “The moral character of cryptographic work.” <https://web.cs.ucdavis.edu/~rogaway/papers/moral.pdf>  
(This is brilliant. If you can’t decide what to read, consider reading this one.)

**At least one of** (as appropriate to you; if I missed your intended/actual profession, ask me about its ethics code):

American Association of University Professors. “Statement on professional ethics.” <https://www.aaup.org/report/statement-professional-ethics>

American Library Association. “Code of ethics.” <http://www.ala.org/advocacy/sites/ala.org.advocacy/files/content/proethics/codeofethics/Code%20of%20Ethics%20of%20the%20American%20Library%20Association.pdf>

Association of Computing Machinery. “Code of ethics and professional conduct.” <https://www.acm.org/code-of-ethics>

Association of Internet Researchers. “Ethical decision-making and Internet research.” <https://aoir.org/reports/ethics2.pdf>

Society of American Archivists. “SAA core values statement and code of ethics.” <https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics>

Society of Professional Journalists. “SPJ code of ethics.” <https://www.spj.org/ethicscode.asp>

N.b. there is no established ethics code for data scientists, as best I can tell!

## Module 2: Harms of Big Data collection

*Topics: Defining “big data.” The uses and individual/societal value of privacy. Notice-and-consent and its ethical problems. Contextual integrity and how it is violated; context collapse. “Anonymization;” reidentification. Government surveillance. Surveillance capitalism; data brokers. Internalization of surveillance. Peer-to-peer surveillance. Surveillance by law enforcement. Doxxing/involuntary outing.*

*“Privacy for marginalized populations has never been, and will never be an abstract. Being surveilled, whether by private actors, or the state, is often the gateway to very tangible harms—violence in the form of police brutality, incarceration, or deportation. And there can be more subliminal, insidious impacts, too.” —Chris Gilliard, <https://www.fastcompany.com/90323529/privacy-is-not-an-abstraction>*

Ward. “Undefined by data: a survey of big data definitions.” <https://arxiv.org/pdf/1309.5821.pdf> (Ask yourself: why are people absent from these definitions, and what are the ethical ramifications of that absence?)

Melendez and Pasternack. “Here are the data brokers quietly buying and selling your personal information.” <https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information>

Gilliard. “Privacy is not an abstraction.” <https://www.fastcompany.com/90323529/privacy-is-not-an-abstraction>

Zuboff. “The surveillance threat is not what Orwell imagined.” <https://time.com/5602363/george-orwell-1984-anniversary-surveillance-capitalism/>

Lubarsky. “Re-identification of ‘anonymized’ data.” <https://georgetownlawtechreview.org/re-identification-of-anonymized-data/GLTR-04-2017/>

Berinato and Nissenbaum. “Stop thinking about consent: it isn’t possible and it isn’t right.” <https://web.archive.org/web/20180927010432/https://hbr.org/2018/09/stop-thinking-about-consent-it-isnt-possible-and-it-isnt-right>

Ng. “Amazon’s helping police build a surveillance network with Ring doorbells.” <https://www.cnet.com/features/amazons-helping-police-build-a-surveillance-network-with-ring-doorbells/>

Hill. “‘People you may know’: a controversial Facebook feature’s ten-year history.” <https://gizmodo.com/people-you-may-know-a-controversial-facebook-features-1827981959>

Pomeroy. “How community members in Ramsey County stopped a big-data plan from flagging students as at-risk.” <https://www.tcdailyplanet.net/how-community-members-in-ramsey-county-stopped-a-big-data-plan-from-flagging-students-as-at-risk/>

## Module 3: Bias

*Topics: The false “technology is neutral” assumption. Training-set bias. Bias and lack of representation in data governance. Past biases fueling predictive analytics. False positives and false negatives. Redlining.*

*“But the problem here isn’t only one of biased datasets or unfair algorithms and of unintended consequences. It’s also indicative of a more persistent problem of researchers actively reproducing ideas that damage vulnerable communities and reinforce current injustices.” —Anna Hoffmann, <https://medium.com/s/story/data-violence-and-how-bad-engineering-choices-can-damage-society-39e44150e1d4>*

O’Shea. “Technology is as biased as its makers.” <https://longreads.com/2019/05/14/technology-is-as-biased-as-its-makers/>

D’Ignazio and Klein. “Chapter seven: the power chapter.” <https://bookbook.pubpub.org/pub/7ruegkt6>

Suresh and Guttag. “A framework for understanding unintended consequences in machine learning.” [http://harinisuresh.com/img/bias\\_framework.pdf](http://harinisuresh.com/img/bias_framework.pdf) (Skim the math; your goal is to understand the sources of bias.)

Hoffmann. “Data violence and how bad engineering choices can damage society.” <https://medium.com/s/story/data-violence-and-how-bad-engineering-choices-can-damage-society-39e44150e1d4>

Adler-Bell. “Privacy for whom?” <https://thenewinquiry.com/privacy-for-whom/> (I’d assign the books if we had time, but since we don’t...)

Biddle. “Facebook’s ad algorithm is a race and gender stereotyping machine.” <https://theintercept.com/2019/04/03/facebook-ad-algorithm-race-gender/>

Mayson. “Bias in, bias out.” [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3257004](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3257004)

Levin. “‘Bias deep inside the code’: the problem with AI ‘ethics’ in Silicon Valley.” <https://www.theguardian.com/technology/2019/mar/28/big-tech-ai-ethics-boards-prejudice>

O’Neil. “How can we stop algorithms telling lies?” <https://www.theguardian.com/technology/2017/jul/16/how-can-we-stop-algorithms-telling-lies>

“Address data-driven discrimination, protect civil rights.” <https://civilrights.org/resource/address-data-driven-discrimination-protect-civil-rights/>

## Module 4: Manipulation through data

*Topics: Bots, sock puppets, and political manipulation; Twitter. Manipulating search engines through data voids. Radicalization through recommender systems; YouTube. “Dark patterns” in web and web service design. Manipulating children and other vulnerable users. Using big data to assess and target emotion/mood.*

*“We can’t control every way that people use the tools we build, but that doesn’t make us any less responsible for them. I realize you didn’t build these monsters on your own, but regardless of how many people’s parts your monster is made from, the fingerprints will always be your own.” —Mike Monteiro, <https://www.buzzfeednews.com/article/mikemonteiro/we-built-a-broken-internet-now-we-need-to-burn-it>*

- Freed. “The tech industry’s war on kids.” <https://medium.com/@richardnfreed/the-tech-industrys-psychological-war-on-kids-c452870464ce> (Pay attention to how this manipulation relies on data collection.)
- boyd. “The fragmentation of truth.” <https://points.datasociety.net/the-fragmentation-of-truth-3c766ebb74cf>
- Gabrielle. “The dark side of gamifying work.” <https://www.fastcompany.com/90260703/the-dark-side-of-gamifying-work> (Pay special attention to the data collection here!)
- Bell. “How ethical is it for advertisers to target your mood?” <https://www.theguardian.com/media/commentisfree/2019/may/05/how-ethical-is-it-for-advertisers-to-target-your-mood>
- Lapowsky. “Here’s how much bots drive conversation during news events.” <https://www.wired.com/story/new-tool-shows-how-bots-drive-conversation-for-news-events/>
- Lewis. “‘Fiction is outperforming reality’: how YouTube’s algorithm distorts truth.” <https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth>
- Machkovech. “Report: Facebook helped advertisers target teens who feel ‘worthless.’” <https://arstechnica.com/information-technology/2017/05/facebook-helped-advertisers-target-teens-who-feel-worthless/>
- Smith. “High-tech redlining: AI is quietly upgrading institutional racism.” <https://www.fastcompany.com/90269688/high-tech-redlining-ai-is-quietly-upgrading-institutional-racism>
- Simonite. “The people trying to make internet recommendations less toxic.” <https://www.wired.com/story/people-trying-make-internet-recommendations-less-toxic/>
- Caulfield. “Four moves.” <https://webliteracy.pressbooks.com/chapter/four-strategies/>

## Module 5: Oppression through data

*Topics: Workplace oppression through employee surveillance. Facial recognition and its abuses. School surveillance; learning analytics; the Library Value Agenda. Social-media surveillance. Research surveillance; IRBs. Stalkerware.*

*“In a time when so much cultural energy is devoted to the self, maximizing, promoting, decorating and sharing it, it has become hard to think with anything like the scrutiny required about how our actions impact others... Despite the rhetoric of sharing, we are doing our best to turn away from examining how our actions impact others. Our world could do with a lot more, rather than less, of that kind of thinking.” —David Golumbia, <http://www.uncomputing.org/?p=1658>*

- Ajunwa et al. “Limitless worker surveillance.” [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2746211](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2746211)
- Herold. “Schools are deploying massive digital surveillance systems.” <https://www.edweek.org/ew/articles/2019/05/30/schools-are-deploying-massive-digital-surveillance-systems.html>
- Baron. “Classroom technology is indoctrinating students into a culture of surveillance.” <https://www.forbes.com/sites/jessicabaron/2019/01/29/classroom-technology-is-indoctrinating-students-into-a-culture-of-surveillance/>
- O’Neil. “How big data transformed applying to college.” <https://slate.com/business/2016/09/how-big-data-made-applying-to-college-tougher-crueler-and-more-expensive.html>
- Grimmelmann. “Do you consent?” <https://slate.com/technology/2015/05/facebook-emotion-contagion-study-tech-companies-need-irb-review.html>
- Metcalf. “‘The study has been approved by the IRB: Gayface AI, research hype and the pervasive data ethics gap.’” <https://medium.com/pervade-team/the-study-has-been-approved-by-the-irb-gayface-ai-research-hype-and-the-pervasive-data-ethics-ed76171b882c>
- Patel et al. “Social media monitoring.” [https://www.brennancenter.org/sites/default/files/publications/2019\\_DHS-SocialMediaMonitoring\\_FINAL.pdf](https://www.brennancenter.org/sites/default/files/publications/2019_DHS-SocialMediaMonitoring_FINAL.pdf)
- Meyer. “An American university is spying on students to predict dropouts.” <http://fortune.com/2018/03/13/university-arizona-catcard-big-data-dropouts/> (Combine this with...)
- Svrluga. “University president allegedly says struggling freshmen are bunnies that should be drowned.” <https://www.washingtonpost.com/news/grade-point/wp/2016/01/19/university-president-allegedly-says-struggling-freshmen-are-bunnies-that-should-be-drowned-that-a-glock-should-be-put-to-their-heads/> (... and consider the Arizona researcher’s ethical responsibility for giving the Simon Newmans of this world a new tool.)
- Greenberg. “Hacker Eva Galperin has a plan to eradicate stalkerware.” <https://www.wired.com/story/eva-galperin-stalkerware-kaspersky-antivirus/> (Content alert: domestic violence and abuse)

## ASSIGNMENT INFORMATION

All due dates are in Canvas. (Maintaining them both in this syllabus and in Canvas led to vastly too many due-date errors on my part.) **PLEASE IGNORE CANVAS'S PERCENTAGE-GRADE CALCULATIONS**; they are scaremongery garbage, but I can't make Canvas turn them off. Available points in this course total to 100, so the best way to keep track of your grade is to track how many points you lose on assignments.

Final grade scale: 100-93.5 A; 93.4-89.5 AB; 89.4-83.5 B; 83.4-79.5 BC; 79.4-73.5 C, 69.5-73.4 D, below 69.5 F

No extra credit opportunities are available in this class.

### Weekly responses (15 points apiece, total 75)

Each week of the course, you will respond to a provocation from me in a designated Canvas discussion forum. Aim for something roughly the length and style of a thoughtful blog post or short-to-medium-length trade-press editorial or explainer (something you might read in *Gizmodo* or *Ars Technica*). You own the copyright to anything you write for this class, of course—I encourage you to write pieces that could be expanded or repurposed in whole or in part for publication or other public dissemination!

**Module 1:** Read Moshe Y. Vardi's "Are we having an ethical crisis in computing?" <https://doi.org/10.1145/3292822> Evaluate its ethical stance from the point of view of at least two of the schools of ethics discussed this week. Then, since this was published in *Communications of the ACM*, evaluate it against the ACM's own ethics code. Finally, note any ethics smells this piece contains (including any I didn't mention that I should have).

**Module 2:** Reread Ward and Barker's piece on definitions of big data. Where are people in these definitions? (Spoiler: they're largely absent!) What are the ethical ramifications of this absence? Write a response updating this piece (which is from 2013) and its definitions to incorporate people (in any or all relevant roles). You are welcome, of course, to seek more people-centric definitions of big data in appropriate research/theory/practice literatures as well as on the web more broadly.

**Module 3:** The Unizin Consortium, through which we have Canvas, is working on the "Unizin Data Platform," a data warehouse containing data on students from Unizin colleges and universities. Read Unizin's description of UDP at <https://unizin.org/solutions/data/data-management/udp/>. Also look at what we know so far about what data might live in UDP: Canvas Data at <https://guides.instructure.com/m/4214/1/449098-what-is-canvas-data> and the database-column list for the Unizin Common Data Model at <http://ucdm.unizin.org/columns.byTable.html>. List student data present in these sources that could be part of biased decisionmaking. How should Unizin and its participating organizations guard against such bias? (You may suggest technical fixes, social/organizational fixes, or both. A fix need not fix the entirety of the problems you notice for you to suggest it.)

**Module 4:** Read Desouza and Smith, "Predictive analytics: nudging, shoving, and smacking behaviors in higher education" at <https://er.educause.edu/articles/2016/8/predictive-analytics-nudging-shoving-and-smacking-behaviors-in-higher-education/> Taking into consideration data collection, data governance, data security, bias, and any other ethical considerations that occur to you, and using any and all questions from applicable ethics codes and schools of ethics, answer: When, if at all, are nudging, shoving, and smacking ethical behaviors from college/university officials? How can unethical uses of these behaviors be prevented?

**Module 5:** Read Wu and Zhang "Responses to critiques on machine learning of criminality perceptions," <https://arxiv.org/pdf/1611.04135v3.pdf> As you did with Vardi's piece in Module 1, formulate an ethics-based response relying on what you have learned about schools of ethics, relevant ethics codes, and ethics smells. (In this case, the AAUP's or AoIR's ethics codes are likely more relevant than ACM's.) Finally, explain how and by whom, in your opinion, Wu and Zhang's further research should be evaluated for ethics, and what should happen should their research design not pass evaluation.

### "Holy !@#\$\$#@, I didn't know that! Now what?" (25 points)

While you do course readings, pay attention to your own reactions to them. Choose a reading or phenomenon that both surprised you (that is, you were not previously aware of it) and repelled you (that is, your immediate response was strongly negative). Please avoid any readings I use in the weekly provocations, since you will already be writing about those.

Doing any additional research you think warranted (don't overload yourself; it's a five-week course!), write a response containing:

- As best you can tell, why you reacted so negatively (I will not be judging whatever your reasons are, just to be clear! There is no right or wrong answer to this; I'd just like you to reflect on it.)

- Where you stand (your “subject position”) with respect to this phenomenon. Does it target you, or others you care about? Make you complicit in something you find unethical? Offend your sense of justice?
- Insofar as you can find out, how do those responsible for this phenomenon justify it ethically?
- Do you find that justification adequate? Why or why not?
- What would those responsible have to do to square the phenomenon with your sense of ethics, with any applicable ethics codes, and with one or more of the schools of ethics we have discussed?
- (Optional, but at least consider:) What do you plan to do about this phenomenon now?

This is due at course end, though of course you may start earlier and turn it in early if you wish.

## Program-level learning outcomes table

Course learning objective	Related to iSchool Program-Level Outcome(s)	Assignments providing evidence of Program-Level Outcome(s)
Basic awareness of major Western schools of ethics and how to use their precepts to resolve ethical dilemmas	1. Students demonstrate understanding of societal, legal, policy or ethical information issues.  6. Students apply theory to professional practice.	1. Weekly responses for Modules 1, 4, and 5. Final assignment.  6. Weekly responses for Modules 3, 4, and 5.
Ability to define “big data” and recognize its use across industries	1. Students demonstrate understanding of societal, legal, policy or ethical information issues.	Weekly response for Module 2.
Broad awareness of collection, retention, use, and transmission of big data about individuals and groups of people across industries	1. Students demonstrate understanding of societal, legal, policy or ethical information issues.  7. Students demonstrate understanding of issues surrounding marginalized communities and information.	1. Weekly responses for Modules 2 and 3.  7. Weekly response for Modules 3 and 5. Final assignment, potentially.
Ability to recognize potential and actual harms in the collection, retention, use, and transmission of big data about individuals and groups of people, including but not limited to bias, loss of autonomy, loss of privacy, loss of free expression, manipulation, redlining, security breaches	1. Students demonstrate understanding of societal, legal, policy or ethical information issues.  7. Students demonstrate understanding of issues surrounding marginalized communities and information.	1. Weekly response for Modules 2, 3, 4, and 5. Final assignment.  7. Weekly response for Modules 3 and 5. Final assignment, potentially.

Course learning objective	Related to iSchool Program-Level Outcome(s)	Assignments providing evidence of Program-Level Outcome(s)
<p>Ability to recognize and respond to common “ethics smells”—poor rationalizations/justifications—for collection, retention, use, and transmission of person-specific big data across industries</p>	<p>1. Students demonstrate understanding of societal, legal, policy or ethical information issues.</p> <p>7. Students demonstrate understanding of issues surrounding marginalized communities and information.</p>	<p>1. Weekly responses for Modules 1, 2, 3, 4, and 5. Final assignment.</p> <p>7. Weekly response for Modules 3 and 5. Final assignment, potentially.</p>

*Do not be daunted  
by the enormity  
of the world's grief.  
Do justly, now.  
Love mercy, now.  
Walk humbly, now.  
You are not obligated  
to complete the work,  
but neither are you free  
to abandon it.  
—Talmud*