



## Faculty and Staff

Professor: Salman Azhar, Ph.D.

Office: W309 | E-mail: [salman.azhar@duke.edu](mailto:salman.azhar@duke.edu) | Phone: +1.408.806.3500

Google/Skype: dsazhar; Facebook Messenger: sazhar

Office hours: M 4:30-6pm, Tu 4-4:30pm, and by appointment.

*I'm usually available without an appointment in the classroom right after my class on Mon & Thu 11:15-11:45am and 2:45-3:15am.*

*I'm available with an appointment, Mon to Fri 9am-5pm unless I'm teaching a class, in a meeting, or meditating.*

Faculty Academic Assistant: Emily Dysart

Office: A322B | E-mail: [emily.dysart@duke.edu](mailto:emily.dysart@duke.edu) | Phone: +1.919.660.1927

Generally available M-F 9am-4pm

There will be a few Teaching Assistants to provide feedback on your work under my direction. They have been trained in a similar course.

## Catalog Description

This course provides an introduction to the legal, policy, and ethical implications of data. The course will examine legal, policy, and ethical issues that arise throughout the full lifecycle 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.

## Motivation

Ideally, all of us would like data analytics to maximize our benefit from personalization without compromising any privacy. This course will provide frameworks to analyze this trade-off and other ethical and legal trade-offs related to data analytics. You will learn about ownership of data, privacy, consent, fairness, and related concepts. You will apply these concepts to develop your own ethical code and get a deeper understanding of yourself and your values.

This course does not promote any particular ethical and legal points of view. Rather students will use the principles and tools to create their own personal ethical and legal codes. These codes will be tested through class discussion and homework against a wide range of examples from work and life. You are likely to go through some mind-bending experiences as you acquire a better understanding of yourself (and your values), and learn about others (and their values), during the process of developing ethical sensitization and legal awareness. You will master the fundamental legal and ethical distinctions necessary to think clearly about legal and ethical issues and will be able to exercise disciplined decision-making skills to choose wisely based on your personal code.



## Pre-Requisites

1. DECISION 520Q: Data Science for Business
2. MANAGEMENT 542Q: Critical thinking, Communication, and collaboration

## Goals

The primary goals of this class are:

1. Ethical: Develop a personal code of ethics
2. Legal: Identify legal and regulatory issues related to data analytics, especially in the US.
3. Social: Understand when data analytics algorithms depend on human bias or opinions, and how that impacts social issues.
4. Analytical: Assess when data analytics raise ethical, legal, and bias concerns by viewing decisions through the lens of ethical, legal, and social aspects.
5. Practical: Apply the principles above to real and simulated situations so you can sleep better at night (with a clear conscience and not in jail).

## Faculty Bio

Salman Azhar is an entrepreneur in residence and a faculty member at Duke University's Fuqua School of Business. He serves on the board of a few startups. Salman has more than 25 years in industry and academia, during which he has led talented teams in developing and launching state-of-the-art software systems.

Salman co-founded DecisionStreet, Swans International, SoftWeb, and over 30 other startups in the areas of predictive computing, business intelligence, data analytics, and decision support. His current and former clients include Honda, Toyota, Lamborghini, Audi, Sony, SAP, HP, ADP, i2, Western Digital, Oracle, Agile, Outcome, and several startups.

Salman's individual expertise lies in the areas of product management, predictive computing, natural language processing, financial systems, automotive, and health information systems.

## Class Preparation

The course website is on Canvas at <https://fuqua.instructure.com>. All course material will be available there. To maximize your learning experience:

1. Complete readings before class time.
2. Submit writing assignments before deadline. (See Canvas for assignments)
3. Synthesize learning from different sources and think about developing an understanding of the topic, not just the reading/writing assignment.
4. Be physically and mentally present in class and contribute to active discussions.
5. Focus on learning by actively listening and clearly expressing your ideas. Balance humility and confidence.
6. Be authentic and transparent with yourself (and, when appropriate, with others).



## Grading

This course follows the recommended grade distribution or “curve” for core courses (SPs (4.0) are given to the top 25 percent of the class, HPs to the next 40 percent (3.5), and Ps or below (3.0 or less) to the remaining 35 percent). The most effective way for you to get your desired grade is to demonstrate that you learned the material. My main responsibility is to instill long-term learning and to assign a fair grade at the end. You must make any regrade requests via Google Form here (link TBD) within one week of receiving back your graded work.

Your final letter grade will be based on the following weighted numeric score.

- Individual
  - 25% Final Exam
  - 15% Personal Ethical code
    - Including weekly progress
  - 20% Individual Assignments
  - 15% Class participation and activities
    - This is a vital part of your learning experience and includes sharing and learning from others
- Team:
  - 25% Team Assignments
    - 20% based on your team submissions
      - Same score for everyone in the team
    - 5% based on your peer review (semester-end)
      - Score based on individual quality & effort

I *may* scale the scores of any component to normalize data distribution (based mean and variance).

## Honor Code

Duke University’s Fuqua School of Business Honor Code applies at all times and can be viewed at [http://www.fuqua.duke.edu/student\\_resources/honor\\_code/](http://www.fuqua.duke.edu/student_resources/honor_code/). Please note:

1. All individual work *must* reflect only your own effort.
2. All team assignments *must* reflect ONLY with members of your team. You are *not* allowed to discuss the team assignments with anyone outside your own team before you turn in the assignments.

It is your responsibility to understand and ask for clarifications if you don’t fully understand anything related to the application of Honor Code in this class. I will spend as much time as needed to answer your questions. Trust is the foundation of our work together, and anyone who violates that trust will have to bear the consequences of anything that must be done to restore that trust.



### Course Plan

We will start with two parallel themes that will come together (along with other topics) later in the course.

1. Developing your personal legal and ethical code
2. Understanding the legal, ethical, and social aspects of Data Analytics

You can access assigned articles from the links provided in the syllabus. If you run into any issues, you can also access the readings using the following URL:

[https://getitatduke.library.duke.edu/ejp/?libHash=PM6MT7VG3J#/search/?searchControl=title&searchType=alternate\\_title\\_equals&criteria=new%20york%20times&language=en-US&titleType=JOURNALS](https://getitatduke.library.duke.edu/ejp/?libHash=PM6MT7VG3J#/search/?searchControl=title&searchType=alternate_title_equals&criteria=new%20york%20times&language=en-US&titleType=JOURNALS).

Date	Topics	Resources
1. Thu 18 Jan	<ol style="list-style-type: none"> <li>1. Start developing a personal code of ethics [Howard &amp; Korver, 2008] and [Anonymous, X]               <ol style="list-style-type: none"> <li>a. Pragmatic consequences of unethical behavior</li> <li>b. Prohibitions against lying, stealing, and harming</li> <li>c. Practical understanding of ethical desensitization (with examples)</li> </ol> </li> <li>2. Start understanding the legal, ethical, and social aspects of Data Analytics. [The White House, 2014]               <ol style="list-style-type: none"> <li>a. What is Big Data?</li> <li>b. The “5 Vs” of Big Data: Volume, Variety, Velocity, Value, and Veracity</li> <li>c. Personalization, identification, de-identification, and re-identification</li> </ol> </li> </ol>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 1-30 (Introduction, Chapter 1)</li> <li>• [Anonymous, X]<sup>1</sup> Entire webpage<sup>2</sup></li> <li>• [The White House, 2014] pages 1-10 &amp; skim the rest of the report<sup>3</sup></li> </ul>
2. Mon 22 Jan	<p>We will continue with the two parallel themes:</p> <ol style="list-style-type: none"> <li>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]               <ol style="list-style-type: none"> <li>a. Lesser-of-two-evils thinking (pages 32)</li> <li>b. Useful distinctions in evaluating ethical problems (prudential, legal, ethical)</li> </ol> </li> </ol>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] Pages 31-50 (Chapter 2)</li> <li>• [Hickman, 2013]</li> </ul>

<sup>1</sup> This webpage overviews [Howard & Korver, 2008] and will help you understand the book.

<sup>2</sup> Tip: Reading overview, then the source, and then the overview again, helps you internalize the reading. You should consider this approach to processing reading materials.

<sup>3</sup> Tip: Develop the ability to quickly get the gist of material when you don’t have time to read the entire source. You can go into details and/or pause to reflect based on your interest.



	<ul style="list-style-type: none"> <li>c. Distinctions of positive versus negative ethics (causing something good or causing something bad)</li> <li>d. Action-based versus consequence-based ethics</li> <li>e. Ethical reasoning versus rationalization (page 34)</li> <li>f. Dangers of rationalizations and justifications</li> <li>g. Tests and guidelines to identify reveal arguments as rationalizations</li> </ul> <p>2. Continue understanding the legal, ethical, and social aspects of Data Analytics. [Hickman, 2013] and [Minority Report, 2002]</p> <ul style="list-style-type: none"> <li>a. What is an algorithm?</li> <li>b. Does technology make “mistakes”? What kind of “mistakes”?</li> <li>c. How is big data is used in parole decisions?</li> <li>d. How do you distinguish between good and bad algorithms?</li> <li>e. Should we worry about AI ruling the world?</li> <li>f. Predictive policing using CRUSH (Criminal Reduction Utilizing Statistical History) and National Security Agency (NSA)’s data analytics algorithms. Problems setting the rules:             <ul style="list-style-type: none"> <li>i. Filter out remove false positives and false negatives (strength and accuracy)</li> <li>ii. Validation of results</li> </ul> </li> </ul> <p>3. Real world example of predictive policing</p> <ul style="list-style-type: none"> <li>a. FBI is asking to build software to scan social media online to predict crime.</li> </ul> <p>4. Discussion of Minority Report movie that dramatizes predictive policing [Minority Report, 2002]</p> <ul style="list-style-type: none"> <li>a. PreCrime police stops murderers before they act, reducing the murder rate to zero.</li> <li>b. PreCrime’s captain is predicted as a murderer who tries to recover the minority report to prove his innocence</li> </ul>	<ul style="list-style-type: none"> <li>• [Minority Report, 2002]</li> <li>• Optional: [Warman, 2012]</li> </ul>
<p>3. Thu 25 Jan</p>	<p>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]</p> <ul style="list-style-type: none"> <li>a. Shared ethical principles from various religions and cultures</li> <li>b. Prudential ethical principles, such as avoid intoxication (pages 53-54)</li> </ul>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 51-70 (Chapter 3)</li> </ul>

	<ul style="list-style-type: none"> <li>c. Positive ethics, such as “thou shalt...”, and negative ethics, such as “thou shalt not...” (page 56)</li> <li>d. Complications when ethical rules conflict... Example: Comprehending the Golden Rule example (pages 57-58)</li> <li>e. Ethical principles from cultural traditions and writings (people, universities, organizations, etc.)</li> </ul> <p>2. Overview of Decision Analysis [Stanford University School of Engineering, 2014], [Fox &amp; Howard 2014], [Keelin, Schoemaker, &amp; Spetzler, 2009], [Korver, 2012]</p> <ul style="list-style-type: none"> <li>a. Formal definition: The point of irrevocable assignment of resources - Applicable only if there are resource constraints</li> <li>b. Identify:             <ul style="list-style-type: none"> <li>1. Major factors</li> <li>2. Possible alternatives</li> <li>3. Available information</li> </ul> </li> <li>c. Determine Impact (possible outcomes under uncertainty of your decision)</li> <li>d. Understand your objective/preferences             <ul style="list-style-type: none"> <li>• Risk preference: The trade-off between greed &amp; fear</li> <li>• Time preference: The trade-off between greed &amp; impatience</li> </ul> </li> <li>e. SCOPED Decision Making Framework (Situation, Choices, Objectives, People, Evaluation, Decision)</li> <li>f. Modeling uncertainty</li> <li>g. Decision Examples</li> </ul>	<ul style="list-style-type: none"> <li>• [Stanford University School of Engineering, 2014] Short video</li> <li>• [Fox &amp; Howard 2014] Audio or Text</li> <li>• [Keelin, Schoemaker, &amp; Spetzler, 2009]</li> <li>• [Korver, 2012]</li> </ul>
<p>4. Mon 29 Jan</p>	<ul style="list-style-type: none"> <li>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]             <ul style="list-style-type: none"> <li>a. Construct an ethical code</li> <li>b. Identify and remove prudential principles from the list (more details on page 106 in Chapter 5)</li> <li>c. Challenges of positive ethics because they have “no bounds” and do not help to distinguish right/wrong</li> <li>d. Identify and sharpen the lines between ethical and unethical (pages 82-85).</li> <li>e. Remove language that is superficial, lofty, etc.</li> </ul> </li> <li>2. AIDA marketing model [Suggett, 2017]             <ul style="list-style-type: none"> <li>a. Attention</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 71-90 (Chapter 4)</li> <li>• [Suggett, 2017]</li> <li>• [Duhigg, 2012]</li> <li>• [mewrox99, 2014]</li> </ul>

	<ul style="list-style-type: none"> <li>b. Interest</li> <li>c. Desire</li> </ul> <p>3. Other data analytics algorithms, such as dating, trading, searching, marketing, ... [Duhigg, 2012]</p> <ul style="list-style-type: none"> <li>a. Marketers want to send specially designed ads to women in their second trimester so they can capture their business before others even know about it.</li> <li>b. Three step process that creates habits (cue, routine, reward).</li> </ul> <p>4. Understanding “False Positives” and “False Negatives” using Bayesian theorem [mewrox99, 2014]</p> <ul style="list-style-type: none"> <li>a. Example of Drug Testing and Breast Cancer.</li> <li>b. Action</li> </ul>	
<p>5. Thu 1 Feb</p>	<ul style="list-style-type: none"> <li>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008] <ul style="list-style-type: none"> <li>a. The hasty decision-making drug test</li> <li>b. Three simple guidelines for ethical decision-making <ul style="list-style-type: none"> <li>i. Clarify (frame) the issue</li> <li>ii. Create alternatives</li> <li>iii. Evaluate the alternatives</li> </ul> </li> <li>c. A consequentialist approach to ethics... Decision points, prudential/legal aspects (page 106)</li> <li>d. Google’s and China’s censorship demands to move servers to China and censor them (pages 106-111) <ul style="list-style-type: none"> <li>• Ethical costs &amp; Prudential benefits (utility) (page 108)</li> <li>• Time frame (short-term vs. long-term)</li> <li>• Information limitation and uncertainty</li> </ul> </li> </ul> </li> <li>2. Three Paradoxes of Big Data provides a framework that encompasses a lot of the legal/ethical considerations [Richards &amp; King, 2013] <ul style="list-style-type: none"> <li>a. Transparency Paradox <ul style="list-style-type: none"> <li>• Data analytics algorithms and operations of big data are kept secrets</li> <li>• Need to develop the right technical, commercial, ethical, and legal safeguards for big data and for individuals.</li> </ul> </li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 91-112 (Chapter 5)</li> <li>• [Richards &amp; King, 2013]</li> </ul>



	<ul style="list-style-type: none"> <li>b. Identity Paradox           <ul style="list-style-type: none"> <li>• Data analytics identifies patterns at the expense of individual and collective identity</li> <li>• Need to reduce the how personalized recommendations undermine our intellectual choices to the point we lose our identity. For example, our Facebook newsfeeds reinforce our biases. Our Netflix recommendations limit our exposure.</li> </ul> </li> <li>c. Power Paradox           <ul style="list-style-type: none"> <li>• Big Data has the power to transform society and only large government and corporate entities have the privilege of using it (at the expense of ordinary individuals).</li> <li>• We need a balance of power between those who generate the data and those who make inferences and decisions based on it, so that one doesn't come to unduly revolt or control the other.</li> </ul> </li> </ul>	
<p>6. Mon 5 Feb</p>	<ol style="list-style-type: none"> <li>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]       <ol style="list-style-type: none"> <li>a. Avoiding ethical risk, such as group memberships (pages 115-116)           <ul style="list-style-type: none"> <li>• Story of a leader in the Nazi party</li> <li>• Avoid ethical violations</li> </ul> </li> <li>b. Three guidelines for ethical decision making (page 116):           <ul style="list-style-type: none"> <li>• Finding the whole truth</li> <li>• Framing issues as relationships</li> <li>• Raising the reciprocity bar</li> </ul> </li> <li>c. Additional guidelines for ethical decision making           <ul style="list-style-type: none"> <li>• Promises and secrets</li> <li>• Reassessing acts of theft</li> <li>• Reevaluating harm</li> </ul> </li> </ol> </li> <li>2. Pricing       <ol style="list-style-type: none"> <li>a. 5Cs (Context, Competition, Customers, Collaborators, Company)</li> <li>b. Relevant (differential) cost           <ul style="list-style-type: none"> <li>• Incremental</li> <li>• Avoidable               <ul style="list-style-type: none"> <li>○ Fixed costs</li> </ul> </li> </ul> </li> </ol> </li> </ol>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 113-130 (Chapter 6)</li> <li>• [Useem, 2017]</li> </ul>



	<ul style="list-style-type: none"> <li>○ Opportunity Costs</li> <li>○ Pricing Playbook</li> <li>● Skimming: Start high, reduce later...</li> <li>● Penetration: Start low, increase later...</li> <li>● Buffering: Limit price cuts through weak demand and then raise slowly</li> </ul> <p>3. Sellers are comparison buyers using data analytics [Useem, 2017]</p> <ul style="list-style-type: none"> <li>a. Price hierarchy in cars (GM example)</li> <li>b. List price distraction</li> <li>c. Left digit bias</li> <li>d. Experiments to produce data</li> <li>e. Nash equilibrium</li> <li>f. Price anchoring/framing</li> </ul>	
<p>7. Thu 8 Feb</p>	<p>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]</p> <ul style="list-style-type: none"> <li>a. Ethical dilemma of corporate leaders</li> <li>b. The example of VC telling COO not to share cash crisis with employees</li> <li>c. Choosing where to work</li> <li>d. Pressures to conform to company expectations, professional guidelines, etc.</li> <li>e. Legal compliance doesn't cover many ethical decisions</li> </ul> <p>2. The Internet That Was (and Still Could Be) [Weinberger, 2015]</p> <ul style="list-style-type: none"> <li>a. Information source without reinforce biases</li> <li>b. Effects on brain</li> <li>c. Net neutrality</li> <li>d. Reasons for optimism</li> </ul> <p>3. Is the Internet Being Ruined? [Dubner, 2016]</p> <ul style="list-style-type: none"> <li>a. Leverage over our livelihoods, interactions, politics</li> <li>b. How Internet engagement is curated by all platforms</li> </ul> <p>4. Algorithmic Harms Beyond Facebook and Google [Tufekci, 2015]</p>	<ul style="list-style-type: none"> <li>● [Howard &amp; Korver, 2008] pages 131-150 (Chapter 7)</li> <li>● [Weinberger, 2015]</li> <li>● [Dubner, 2016]</li> <li>● [Tufekci, 2016]</li> </ul>

	<ol style="list-style-type: none"> <li>a. Facebook algorithms determine what more than a billion people see every day. Are the “trending topics” vetted by a small team of editors or “surfaced by an algorithm?”</li> <li>b. While these algorithms also use data, math and computation, they are a fountain of bias and slants — of a new kind.</li> <li>c. Machine learning is becoming pervasive and that allows computers to independently learn to learn, sometimes using techniques used by humans. We know the algorithms but we don’t know the knowledge gained using them, and how (due to indirect approach).</li> <li>d. Facebook’s goal is to maximize the amount of user engagement, sometimes superficially</li> <li>e. The newsfeed algorithm also values comments and sharing. This suits items that generate quantity, rather than quality.</li> <li>f. Google PageRank can have similar biases that influences viral aspects of webpages.</li> <li>g. Gatekeeper algorithms             <ul style="list-style-type: none"> <li>• Lack of visibility</li> <li>• Information asymmetry</li> <li>• Hidden influence</li> <li>• Privacy</li> <li>• Incorrect conclusions</li> <li>• Misuse of information</li> </ul> </li> <li>h. Case studies             <ul style="list-style-type: none"> <li>• Social Movements and the Civic Sphere</li> <li>• Elections and Algorithmic Harms</li> </ul> </li> </ol>	
<p>8. Mon 12 Feb</p>	<ol style="list-style-type: none"> <li>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]             <ol style="list-style-type: none"> <li>a. Forming habits that reinforce our ethical code</li> </ol> </li> <li>2. Limits of Analytics [Marcus &amp; Davis, 2014]             <ol style="list-style-type: none"> <li>a. Problems with Big Data                 <ul style="list-style-type: none"> <li>• Doesn’t tell which correlations are meaningful</li> <li>• Doesn’t replace scientific (supplements it)</li> <li>• Can be gamed</li> <li>• Results are less robust than they seem at first glance</li> <li>• Echo-chamber effect</li> </ul> </li> </ol> </li> </ol>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 151-174 (Epilogue &amp; Appendices A &amp; B)</li> <li>• [Marcus &amp; Davis, 2014]</li> <li>• [Lazer et al, 2014]</li> </ul>



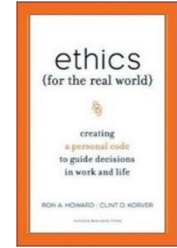
	<ul style="list-style-type: none"> <li>• Too many correlations can make meaningful correlations hard to identify</li> <li>• Scientific-sounding solutions to imprecise questions</li> <li>• Can't analyze things that are less common</li> <li>• People believe the hype</li> </ul> <p>b. Google Flu story [Lazer et al, 2014]</p> <ul style="list-style-type: none"> <li>• Big Data as substitute for traditional data collection and analysis</li> <li>• Predicting accuracy and inaccuracy</li> <li>• Google Flu algorithm methodology</li> <li>• Transparency and replicability</li> <li>• Granularity</li> <li>• Predicting the unknown</li> <li>• Using all data sources</li> </ul>	
9. Thu 15 Feb	<p>1. Continue developing a personal code of ethics [Howard &amp; Korver, 2008]</p> <p>a. Discussion on personal ethical codes</p> <p>2. Privacy Concerns [CBS, 2016] [Weise &amp; Guynn, 2014] [Barocas &amp; Nissenbaum, 2014]</p> <p>a. Uber Tracking</p> <ul style="list-style-type: none"> <li>• Tracking travel records without permission</li> <li>• #DeleteUber movement</li> </ul> <p>b. Procedural Privacy Protections</p> <ul style="list-style-type: none"> <li>• Procedural approaches to protecting privacy             <ul style="list-style-type: none"> <li>a. Informed Consent</li> <li>b. Anonymity</li> </ul> </li> <li>• Recognizing the limits of purely procedural approaches to protecting privacy - Privacy loses the trade-off with big data</li> </ul>	<ul style="list-style-type: none"> <li>• [Howard &amp; Korver, 2008] pages 175-186 (Our Messages)</li> <li>• [CBS, 2016]</li> <li>• [Weise &amp; Guynn, 2014]</li> <li>• [Barocas &amp; Nissenbaum, 2014]</li> </ul>
10. Mon 19 Feb	<p>1. US approach to privacy law ([U.S. Constitution] and [The White House, 2014] pages 15-21 starting from U.S. Privacy Law and International Privacy Frameworks)</p> <p>a. Focus on the fourth amendment</p> <p>b. Development of Privacy Law</p> <p>c. The Fair Information Practice Principles</p> <p>d. Sector-Specific Privacy Laws</p> <p>e. Consumer Privacy Bill of Rights</p>	<ul style="list-style-type: none"> <li>• [U.S. Constitution]</li> <li>• [The White House, 2014] pages 15-21</li> <li>• [Schwartz, 2010]</li> </ul>

	<p>f. Global Interoperability</p> <p>2. Data Protection Law and the Ethical Use of Analytics [Schwartz, 2010]</p> <p>a. The need for contextual examination of analytics</p> <ul style="list-style-type: none"> <li>• Risks posed to privacy and responsible processes</li> <li>• Evolving environment for data protection</li> <li>• The ethics of analytics</li> </ul> <p>b. Application of analytics</p> <ul style="list-style-type: none"> <li>• Multichannel marketing</li> <li>• Fraud detection and prevention</li> <li>• Protecting data security</li> <li>• Health Care Research (The story of Viagra)</li> <li>• Consumer products <ul style="list-style-type: none"> <li>○ Personalized (news feeds)</li> <li>○ Productive (financial, medical, ...)</li> </ul> </li> </ul> <p>c. The Different Stages of Analytics</p> <ol style="list-style-type: none"> <li>i. Collection</li> <li>ii. Integration and Analysis</li> <li>iii. Decision-making</li> <li>iv. Review and Revision</li> </ol> <p>d. Fair Information Practices (FIPs)</p> <ul style="list-style-type: none"> <li>• Automated Individual Decisions</li> <li>• Purpose Specification and Use Limitations</li> <li>• FIPs in the Ethical Use of Analytics</li> </ul> <p>3. The Supreme Court's Big Data Problem [Brennan-Marquez, 2016]</p> <ol style="list-style-type: none"> <li>a. Spokeo v. Robins case</li> <li>b. Breaking the law or breaking the law in a manner that injured anyone</li> <li>c. Use of inaccurate and accurate data</li> <li>d. The trade-off between discrimination or personalization?</li> </ol>	<ul style="list-style-type: none"> <li>• [Brennan-Marquez, 2016]</li> </ul>
<p>11. Thu 22 Feb</p>	<p>1. How Data Mining Discriminates [Barocas &amp; Selbst, 2016]</p> <ol style="list-style-type: none"> <li>a. Defining the “Target Variable” and “Class Labels”</li> <li>b. Training Data <ul style="list-style-type: none"> <li>• Labeling Examples</li> <li>• Data Collection</li> </ul> </li> <li>c. Feature Selection</li> </ol>	<ul style="list-style-type: none"> <li>• [Barocas &amp; Selbst, 2016] pages 677-693 (Section I only)</li> </ul>

	<ul style="list-style-type: none"> <li>d. Proxies</li> <li>e. Masking</li> </ul> <p>2. The Hidden Biases in Big Data [Crawford, 2013]</p> <ul style="list-style-type: none"> <li>a. Data fundamentalism <ul style="list-style-type: none"> <li>• Case study: Twitter and foursquare data during Hurricane Sandy</li> </ul> </li> <li>b. Address these weaknesses in big data science <ul style="list-style-type: none"> <li>• Validate data sources</li> <li>• Identify cognitive biases</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• [Crawford, 2013]</li> </ul>
<p>12. Mon 26 Feb</p>	<p>1. Redress Predictive Privacy Harms [Crawford &amp; Schultz, 2014]</p> <ul style="list-style-type: none"> <li>a. Expansion in use of PII (Personally Identifiable Information) using Big Data by <ul style="list-style-type: none"> <li>• Leveraging technology that maximizes computational power and accuracy</li> <li>• Using on a range of tools to clean and compare data.</li> <li>• Believing that large data sets are more reliable</li> </ul> </li> <li>b. Vulnerability of health records <ul style="list-style-type: none"> <li>• Electronic pharmaceutical records</li> <li>• Social media activity</li> <li>• Downloadable data</li> <li>• Differential privacy</li> </ul> </li> </ul> <p>2. Protecting Privacy [Richards &amp; King, 2016]</p> <ul style="list-style-type: none"> <li>a. Regulations <ul style="list-style-type: none"> <li>• Procedural: reinforcing transparency of processing, the basis for algorithmic decisions (“algorithmic accountability”), and providing choice to opt-out</li> <li>• Regulating decisions that threatens identity, equality, security, data integrity, and trust</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• [Crawford &amp; Schultz, 2014] pages 93-99</li> <li>• [Richards &amp; King, 2016]</li> </ul>

## Reading List

All the readings below are available online except the one(s) marked with an asterisk. This will save you money and save the earth a few trees. You can access assigned articles from the links provided in the syllabus. If you run into any issues, you can also access the readings using the following URL: [https://getitatduke.library.duke.edu/ejp/?libHash=PM6MT7VG3J#/search/?searchControl=title&searchType=alternate\\_title\\_equals&criteria=new%20york%20times&language=en-US&titleType=JOURNALS](https://getitatduke.library.duke.edu/ejp/?libHash=PM6MT7VG3J#/search/?searchControl=title&searchType=alternate_title_equals&criteria=new%20york%20times&language=en-US&titleType=JOURNALS).



1. \* [Howard & Korver, 2008] Howard, R. E. & Korver, C. D. (2008). *Ethics for the Real World: Creating a Personal Code to Guide Decisions in Work and Life*. Cambridge, MA: Harvard Business Review Press. (You'll need a physical or electronic copy of this complete book.) A preview at [https://books.google.com/books/about/Ethics\\_for\\_the\\_Real\\_World.html?id=OqnrQFfXb0C](https://books.google.com/books/about/Ethics_for_the_Real_World.html?id=OqnrQFfXb0C).
2. [Anonymous, X] Ethics (for the real world). (n.d.). Retrieved from the Utilitarianism Wiki: [http://utilitarianism.wikia.com/wiki/Ethics\\_\(for\\_the\\_real\\_world\)](http://utilitarianism.wikia.com/wiki/Ethics_(for_the_real_world))
3. [The White House, 2014] The White House. (2014). *Big Data: Seizing Opportunities, Preserving Values*, pp. 1-10. Retrieved from [https://obamawhitehouse.archives.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_may\\_1\\_2014.pdf](https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf)
4. [Hickman, 2013] Hickman, L. (2013, July 1). How algorithms rule the world. *The Guardian*. Retrieved from <https://www.theguardian.com/science/2013/jul/01/how-algorithms-rule-world-nsa>
5. [Warman, 2012] Warman, M. (2012, January 27). FBI to Use Twitter to Predict Crimes. *The Telegraph*. Retrieved from <http://www.telegraph.co.uk/technology/news/9043690/FBI-to-use-Twitter-to-predict-crimes.html>
6. [Minority Report, 2002] Minority Report (film). (n.d.). Retrieved from Wikipedia: [https://en.wikipedia.org/wiki/Minority\\_Report\\_\(film\)](https://en.wikipedia.org/wiki/Minority_Report_(film))
7. [Stanford University School of Engineering, 2014] (2014, October 30). *Stanford Professor Ron Howard: Decision Analysis Turns 50* [Video File]. Retrieved from [https://youtu.be/3MnsqBFpo\\_A](https://youtu.be/3MnsqBFpo_A)
8. [Fox & Howard, 2014] Fox, J. (Interviewer) & Howard, R. E. (Interviewee). (2014). *Making Good Decisions* [Interview audio file and transcript]. Retrieved from Harvard Business Review Web site: <https://hbr.org/2014/11/making-good-decisions>
9. [Keelin, Schoemaker, & Spetzler, 2009] Keelin, T., Schoemaker, P., & Spetzler, C. (2009). *Decision Quality: The Fundamentals of Making Good Decisions*. Decision Education Foundation. Retrieved from [http://media.wix.com/ugd/d3c97e\\_9452fec9f54e674f5c4e6f0dcd5edc63.pdf](http://media.wix.com/ugd/d3c97e_9452fec9f54e674f5c4e6f0dcd5edc63.pdf)
10. [Korver, 2012] Korver, C. (2012). Applying Decision Analysis to Venture Investing. *Kauffman Fellows Report, volume 3*. Retrieved from

[http://www.kauffmanfellows.org/journal\\_posts/applying-decision-analysis-to-venture-investing/](http://www.kauffmanfellows.org/journal_posts/applying-decision-analysis-to-venture-investing/)

11. [Suggett, 2017] Suggett, P. (2017, February 16). Get To Know, and Use, AIDA. *The Balance*. Retrieved from <https://www.thebalance.com/get-to-know-and-use-aida-39273>
12. [Duhigg, 2012] Duhigg, C. (2012, February 16). How Companies Learn Your Secrets. *The New York Times Magazine*. Retrieved from <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>
13. [mewrox99, 2014] [mewrox99]. (2014, September 1). *Bayes' Theorem - Drug Test Probability* [Video File]. Retrieved from <https://www.youtube.com/watch?v=Fp8tvqEGOJM>
14. [Richards & King, 2013] Richards, N. M., & King, J. H. (2013). Three Paradoxes of Big Data. *66 Stanford Law Review Online* 41. Retrieved from [https://review.law.stanford.edu/wp-content/uploads/sites/3/2016/08/66\\_StanLRevOnline\\_41\\_RichardsKing.pdf](https://review.law.stanford.edu/wp-content/uploads/sites/3/2016/08/66_StanLRevOnline_41_RichardsKing.pdf)
15. [Jerry Useem, 2017] Useem, J. (2017, May). How Online Shopping Makes Suckers of Us All. *The Atlantic*. Retrieved from <https://www.theatlantic.com/magazine/archive/2017/05/how-online-shopping-makes-suckers-of-us-all/521448/>
16. [Weinberger, 2015] Weinberger, D. (2015, June 22). The Internet That Was (and Still Could Be). *The Atlantic*. Retrieved from <https://www.theatlantic.com/technology/archive/2015/06/medium-is-the-message-paradise-paved-internet-architecture/396227/>
17. [Dubner, 2016] Dubner, S. J. (2016, July 13). Is the Internet Being Ruined? *Freakonomics Radio*. Podcast retrieved from <http://freakonomics.com/podcast/internet/>
18. [Tufekci, 2016] Tufekci, Z. (2016, May 19). The Real Bias Built In at Facebook. *The New York Times*. Retrieved from <https://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html>
19. [Marcus & Davis, 2014] Marcus, G., & Davis, E. (2014, April 6). Eight (No, Nine!) Problems with Big Data. *The New York Times*. Retrieved from <https://www.nytimes.com/2014/04/07/opinion/eight-no-nine-problems-with-big-data.html>
20. [Lazer et al., 2014] Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014, March 14). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203-1205. Retrieved from <https://gking.harvard.edu/files/gking/files/0314policyforumff.pdf> (also available at <http://science.sciencemag.org/content/343/6176/1203.full>).
21. [CBS, 2016] [CBS This Morning]. (2016, December 12). *Uber raises privacy concerns with policy to track location even after ride* [Video File]. Retrieved from <https://www.youtube.com/watch?v=HRjbFvnOPa8>
22. [Weise & Guynn, 2014] Weise, E., & Guynn, J. (2014, November 19). Uber tracking raises privacy concerns. *USA Today*. Retrieved from <https://www.usatoday.com/story/tech/2014/11/19/uber-privacy-tracking/19285481/>
23. [Barocas & Nissenbaum, 2014] Barocas, S., & Nissenbaum, H. (2014). Big Data's End Run Around Procedural Privacy Protections. *Communications of the ACM*, 57(11), 31-33. Retrieved from <https://pdfs.semanticscholar.org/e6cc/9aa6de6ff274421743a397f8da77490d7e7b.pdf>



24. [U.S. Constitution] U.S. Constitution (Amendments to the Constitution). Retrieved from <http://constitutionus.com/#amendments>
25. [Schwartz, 2010] Schwartz, P. M. (2010). Data Protection Law and the Ethical Use of Analytics. *Centre for Information Policy Leadership*, pp. 18-26. Retrieved from [https://iapp.org/media/pdf/knowledge\\_center/Ethical\\_Underpinnings\\_of\\_Analytics.pdf](https://iapp.org/media/pdf/knowledge_center/Ethical_Underpinnings_of_Analytics.pdf)
26. [Brennan-Marquez, 2016] Brennan-Marquez, K. (2016, June 29). The Supreme Court's Big Data Problem. *Points*. Retrieved from <https://points.datasociety.net/the-supreme-courts-big-data-problem-9401fa88a3e0>
27. [Barocas & Selbst, 2016] Barocas, S., & Selbst, A. (2016). Big Data's Disparate Impact. *California Law Review*, 104(671), 677-693. Retrieved from <http://www.californialawreview.org/2-big-data/> (PDF at <http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf>).
28. [Crawford, 2013] Crawford, K. (2013, April 1). The Hidden Biases in Big Data. *Harvard Business Review*. Retrieved from <https://hbr.org/2013/04/the-hidden-biases-in-big-data>
29. [Crawford & Schultz, 2014] Crawford, K., & Schultz, J. (2014). Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms. *Boston College Law Review*, 55(1), 93-99, 121-128. Retrieved from <http://lawdigitalcommons.bc.edu/cgi/viewcontent.cgi?article=3351&context=bclr>
30. [Richards & King, 2016] Richards, N. M., & King, J. H. (2016). Big Data and the Future for Privacy. *Handbook of Research on Digital Transformations*, 16-21. Retrieved from [https://papers.ssrn.com/sol3/Delivery.cfm/SSRN\\_ID2714496\\_code400644.pdf?abstractid=2512069&mirid=1](https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2714496_code400644.pdf?abstractid=2512069&mirid=1)

These readings just missed the final cut. I'm including them if you are interested in learning more.

1. Tufekci, Z. (2015). Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency. *Colorado Technology Law Journal*, 13, 203-209. Retrieved from <http://ctlj.colorado.edu/wp-content/uploads/2015/08/Tufekci-final.pdf>
2. Mozur, P., Scott, M., & Isaac, M. (2017, September 17). Facebook Faces a New World as Officials Rein in a Wild Web. *The New York Times*. Retrieved from <https://www.nytimes.com/2017/09/17/technology/facebook-government-regulations.html>
3. Maheshwari, S., & Stevenson, A. (2017, September 15). Google and Facebook Face Criticism for Ads Targeting Racist Sentiments. *The New York Times*. Retrieved from <https://www.nytimes.com/2017/09/15/business/facebook-advertising-anti-semitism.html>
4. Ehrlich, E. (2017, July 20). 'Neutrality' for Thee, but Not for Google, Facebook, and Amazon. *The Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/neutrality-for-thee-but-not-for-google-facebook-and-amazon-1500591612>
5. Rosen, R. J. (2013, October 3). Is This the Grossest Advertising Strategy of All Time? *The Atlantic*. Retrieved from <https://www.theatlantic.com/technology/archive/2013/10/is-this-the-grossest-advertising-strategy-of-all-time/280242/>
6. Schneier, B. (2017, January 27). Click Here to Kill Everyone: With the Internet of Things, we're building a world-size robot. How are we going to control it? *New York Magazine*.





Retrieved from <http://nymag.com/selectall/2017/01/the-internet-of-things-dangerous-future-bruce-schneier.html>

7. Calo, R., & Rosenblat, A. (2017). The Taking Economy: Uber, Information, Power<sup>4</sup>. *Columbia Law Review*, 117. Retrieved from [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2929643](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2929643)
8. Thaler, R. H., & Tucker, W. (2013). Smarter Information, Smarter Consumers. *Harvard Business Review*. Retrieved from <https://hbr.org/2013/01/smarter-information-smarter-consumers><sup>5</sup>

---

<sup>4</sup> This paper talks about Information Asymmetry, “Leveraging access to information about users and their control over the user experience to mislead, coerce, or otherwise disadvantage sharing economy participants”, Consumer Protection law/legal

<sup>5</sup> An interesting argument for the share-the-data approach – and how it could improve decision making on the societal scale