



LGST 642x

Big Data, Big Responsibilities: The Law and Ethics of Business Analytics

Fall 2016 | MW 10:30am-12pm | JMHH F65

NOTE: This experimental course covers new and fast-moving issues. Some details of readings and session topics may change between now and the start date, or even during the course.

Overview

Significant technologies always have unintended consequences, and their effects are never neutral. A world of ubiquitous data, subject to ever more sophisticated collection, aggregation, analysis, and use, creates massive opportunities for both financial gain and social good. It also creates dangers in areas such as privacy and discrimination, as well as simple hubris about the effectiveness of management by algorithm. This course introduces students to the legal, policy, and ethical dimensions of big data, predictive analytics, and related techniques. It then examines responses—both private and governmental—that may be employed to address these concerns.

Instructor

Associate Professor Kevin Werbach
Department of Legal Studies and Business Ethics
673 Huntsman Hall
(215) 898-1222
werbach@wharton.upenn.edu (best way to reach me)

Office Hours: Monday 12:30-2pm, or by appointment

Learning Objectives

Good data-driven decision-making means not just generating solutions, but understanding how to use them. Some of the most sophisticated firms in terms of data science expertise have already gotten into trouble over concerns about privacy, security, manipulation, and discrimination. Failure to anticipate such issues can result in ethical lapses, public relations disasters, regulatory sanctions, and even legal liability. My goal is to help you develop the skills needed to use analytics in the most responsible way, while remaining focused on your business objectives.

After completion of the course, you should be able to:

1. Identify where algorithms depend on human judgments or assumptions.
2. Describe legal rules and regulatory obligations, in the U.S. and elsewhere, that may apply to business analytics.
3. Evaluate claims that applications of business analytics violate privacy, result in discrimination, or cause other legal, ethical, and public policy problems.
4. Develop thoughtful responses to concerns about the uses of data science.
5. After graduation from Wharton, don't destroy the world, crash the economy, go to jail, or all of the above. (Money-back guarantee not available.)

The course is non-quantitative, although a basic familiarity with data science techniques is assumed. Many of the questions it covers do not have simple answers, either because concepts such as privacy cannot be reduced to binary choices, or because the scenarios involved are still unfolding. Therefore, a major goal of the course is to help you develop your own well-grounded viewpoints.

Materials

All readings are contained in a coursepack available through Study.net, or are available for free online at the hyperlinks provided in this syllabus.

Course Requirements and Grading

Case Study Analysis (25%)

Pick a specific example from one of the optional readings. (E.g., in the article for class #4, the discussion about retailers tracking customer movements in stores). Imagine you are at the company or companies involved. Management asks you for recommendations. How would you assess the legal and ethical risks? What steps could be taken to mitigate them? Should the company move forward, or take other actions?

You have the choice of doing this assignment on either the Data Protection section of the course (due following class #8) or the Fairness section (due following class #12). You may also submit both assignments by the deadlines, in which case you will receive the higher of the two grades.

You may choose the format for your submission. It could be a paper or memo (up to 2500 words), a slide deck, a screencast with audio narration, a video, an animation, a public blog post, or another way to frame your analysis. (If you have an idea, feel free to check with me before starting.) I value creativity highly.

Regardless of the format, the same grading criteria will apply, taking into account effort and originality. I am looking for work that is well-organized, well-written, and well-reasoned. It should be persuasive, but should take counter-arguments or contrary examples into consideration. The most successful submissions will demonstrate critical engagement with the course materials, and will illustrate their points with concepts introduced in the course. You are also welcome to incorporate examples from your own research. Any statement that quotes or relies upon a source, whether from the course readings or outside materials, should include either a footnote or an in-line (parenthetical) reference with a bibliography.

Group Activity (15%)

We will do a group negotiation during class #13. You will take on roles and attempt to develop a consensus to address concerns about business analytics. You will receive credit as follows:

- Individual pre-negotiation prep sheet (5%)
- Group proposal, or explanation of your failure to reach consensus (10%)

Exam (40%)

There will be one closed-book exam, which will be administered during the final session. It will test your comprehension of the required readings and the concepts developed in class.

Participation (20%)

You will be assessed on the overall quality of your contributions to the course. Attendance is one factor in the participation score, but not the primary one.

Classroom Expectations and Participation

- Please arrive on time.
- Bring a name tent and display it for each class.
- Turn off all electronic devices not used for active note-taking.
- Be prepared to contribute to the class discussion.

Participation is your responsibility. Prepare for class. Raise your hand, and make comments that show engagement with the readings and course concepts. Don't be afraid to make a mistake or take a stand. That's how learning happens.

The course is interactive, so it won't be as successful (nor will you) if students are frequently absent. That being said, we all face tradeoffs in life. I recognize that students will sometimes miss class for understandable but unexcused reasons such as recruiting. Your grade will reflect your overall level of participation; active, high-quality contributions may counterbalance an occasional absence.

There is no way to "make up" a class. However, you may miss (or show up late by more than 10 minutes) one class session during the quarter for any reason, without penalty.

If you cannot attend a class, I encourage you to get notes from another student and/or watch the class recording. Video recordings of all sessions will be available on Canvas. You don't need to tell me in advance, or contact me afterwards to explain an un-excused absence (i.e., not involving a medical issue, obligatory university/athletic trip, or family/personal emergency.) If you believe your absence should be excused, please email me an explanation and, where possible, documentation.

Instructor Bio

Professor Kevin Werbach is an expert on the business, legal, and social implications of internet and communications technologies. He has published numerous academic and popular works on internet policy, telecommunications regulation, and legal aspects of emerging technologies. He is a pioneer in the emerging field of gamification, and the co-author of *For the Win: How Game Thinking Can Revolutionize Your Business*, which has been translated into six languages. Over 300,000 students worldwide have registered for his Coursera Massive Open Online Course (MOOC). He has received the Wharton MBA Program Teaching Innovation and Curricular Innovation Award, and was named Wharton's first-ever "Iron Prof".

Professor Werbach co-led the review of the Federal Communications Commission (FCC) for the Obama Administration's Presidential Transition Team in 2008. For nine years he hosted Supernova, a leading technology executive conference. Prior to joining the Wharton faculty, he was the editor of *Release 1.0*, a renowned technology report for senior executives, and co-organizer, with Esther Dyson, of the annual PC Forum. He also served as Counsel for New Technology Policy at the FCC during the Clinton Administration. Werbach is a *magna cum laude* graduate of Harvard Law School, where he was Publishing Editor of the Law Review, and he graduated *summa cum laude* from the University of California at Berkeley.

Syllabus

All readings are found in the coursepack, and hyperlinks are provided below to online versions. Where there are questions listed under a reading, be prepared to address them in class discussion.

FOUNDATIONS

1. THE PROMISE AND THE PERIL

How might data science change the relationships among firms, customers, employees, other firms, and governments? What are some of the legal or ethical concerns that may arise?

The White House, [Big Data: Seizing Opportunities, Preserving Values](#) (2014), pp. 1-10

- *What are the key innovations of big data and business analytics?*

Leo Hickman, [How Algorithms Rule the World](#), The Guardian, July 1, 2013

- *Algorithms are cool! Why should we worry if they rule the world?*

2. “IT’S JUST MATH”

Algorithms rely on human decisions about how data are collected, analyzed, and used. Failure to appreciate this can lead to problems.

Neil Richards and Jonathan King, [Three Paradoxes of Big Data](#), 66 Stanford Law Review Online 41 (2013)

- *What practical problems arise from the three paradoxes the authors identify?*

Zeynep Tufekci, [The Real Bias Built In at Facebook](#), N.Y. Times, May 19, 2016

- *Why was Facebook criticized for its Trending Topics?*
- *Why does Tufekci say that algorithms are not neutral? What does that even mean?*
- *If algorithms are inherently biased, does that undermine the value of analytics in business?*

Optional

Tarleton Gillespie, [Algorithms, Clickworkers, and the Befuddled Fury Around Facebook Trends](#), NiemanLab, May 19, 2016

Tyler Woods, [“Mathwashing” and the Zeitgeist of Data Worship](#), Technical.ly Brooklyn, June 8, 2016

3. LIMITS OF ALGORITHMS

The first step to responsible use of analytics is to appreciate its limitations and known statistical issues.

Gary Marcus & Ernest Davis, [8 \(No, 9!\) Problems with Big Data](#), N.Y. Times, April 6, 2014

- *What are some of the common themes in the authors’ list of problems?*

David Lazer et al, [The Parable of Google Flu: Traps in Big Data Analysis](#), Science, March 14, 2014

- *Why was Google Flu Trends so accurate initially, and not subsequently?*
- *Should the failure of Google Flu make us skeptical about the potential of business analytics?*

Paul Ohm, [The Underwhelming Benefits of Big Data](#), 161 Univ. of Penn. Law Review Online 339 (2013)

- *How does Ohm think we should assess the value of analytics? Do you agree?*
- *What is your reaction to the quote by Facebook’s Jeff Hammerbacher in footnote 28: “The best minds of my generation are thinking about how to make people click ads.... That sucks.”*

Optional

Alexis Madrigal, [The Deception That Lurks in Our Data-Driven World](#), Fusion, October 6, 2015

Gilad Lotan, [Apple, Apps and Algorithmic Glitches](#), Medium, January 16, 2015

DATA PROTECTION

4. PRIVACY IN A BIG DATA WORLD

Are there limits on how data should be collected, used, and shared?

Neil Richards & Jonathan King, [Big Data and the Future for Privacy](#) (2016), pp. 3-15

- *What is privacy? Why is it important?*
- *How does big data change the way we think about privacy?*

Solon Barocas and Helen Nissenbaum, [Big Data's End Run Around Procedural Privacy Protections](#), Communications of the ACM (November 2014)

- *What do the authors believe that transparency and consent are insufficient?*

Elizabeth Weise and Jessica Guynn, [Uber Tracking Raises Privacy Concerns](#), USA Today, Nov. 19, 2014

- *Do you find any of Uber's actions described in this article troubling?*

Optional

Alessandro Acquisti et al, [Privacy and Human Behavior in the Age of Information](#), Science, January 30, 2015

David Streitfeld, N.Y. Times, [Data-Crunching Is Coming to Help Your Boss Manage Your Time](#), August 17, 2015

Stephanie Clifford and Quentin Hardy, [Attention, Shoppers: Store is Tracking Your Cell](#), N.Y. Times, July 14, 2013

5. PRIVACY LAW

Privacy is the subject of many legal and regulatory regimes in the U.S. and elsewhere. How well do those rules apply to big data and business analytics?

The White House, [Big Data: Seizing Opportunities, Preserving Values](#) (2014), pp. 15-21

- *What are the key elements of the U.S. approach to privacy law?*
- *Do you think the U.S. legal framework is effective in general? Will it be effective for the novel challenges of big data and business analytics?*

Paul M. Schwartz, [Data Protection Law and the Ethical Use of Analytics](#), Centre for Information Policy Leadership (2010), pp. 18-26

- *How does the European approach to privacy differ from the U.S.?*
- *How should companies respond to the legal requirements around privacy?*

Kiel Brennan-Marquez, [The Supreme Court's Big Data Problem](#), Points, June 29, 2016

- *How is the growth of business analytics changing the legal boundaries of privacy?*

Optional

David Streitfeld, [Google Concedes That Drive-By Prying Violated Privacy](#), N.Y. Times, March 12, 2013

Andrea Peterson, [Snapchat Agrees to Settle FTC Charges That it Deceived Users](#), Washington Post, May 8, 2014

Woodrow Hartzog and Daniel Solove, [The FTC as Data Security Regulator: FTC v. Wyndham and Its Implications](#), Bloomberg BNA Privacy and Security Law Report, April 14, 2014

6. PERILS OF PREDICTION

The ability to make predictions about individual behavior based on models means that “private” information, including very sensitive facts, can be inferred without ever having access to personal data. Can we even talk about privacy if personally identifiable information isn’t being shared?

Charles Duhigg, [How Companies Learn Your Secrets](#), N.Y. Times Magazine, Feb. 16, 2012

- *How does Target analyze customer data to make inferences about customers?*
- *In your opinion, is the Target system an intrusion on privacy? Why or why not?*
- *Do Target’s actions violate any legal rules?*
- *Do Target’s actions violate any ethical norms?*
- *Should Target do anything differently?*

Kate Crawford and Jason Schultz, [Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms](#), 55 Boston College Law Review pp. 93-99, 121-28 (2014)

- *How is the “due process” approach different from existing privacy laws we’ve discussed?*
- *Will this approach be effective?*

Optional

Michal Kosinski et al., [Private Traits and Attributes Are Predictable From Digital Records of Human Behavior](#), 110 Proceedings of the National Academy of Sciences 5802 (2013)

Troy Raeder, Brian Dalessandro, and Claudia Perlich, [Considering Privacy in Predictive Modeling Applications](#), KDD 2014 Data Ethics Workshop (2014)

Jennifer Doleac, [The Unintended Consequences of 'Ban the Box'](#), Real Clear Markets, May 31, 2016

Rachel Emma Silverman, [Bosses Tap Outside Firms to Predict Which Workers Might Get Sick](#), Wall St. Journal, February 17, 2016

7. INFLUENCING USERS

To what extent does analysis itself influence behavior? And what are the limits on using analytics not merely to understand and predict customer actions, but to shape them?

Zeynep Tufekci, [Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency](#), J. on Telecomm. & High Tech L. (2015), pp. 203-209

- *What was Facebook trying to achieve in its emotional contagion study?*
- *Why were Facebook’s actions controversial?*
- *What is “algorithmic gatekeeping”? Why does Tufekci believe it is a concern?*

Rebecca Rosen, [Is This the Grossest Advertising Strategy of All Time?](#), The Atlantic, October 3, 2013

- *What exactly does the author find new and objectionable about this marketing approach?*

Evgeny Morozov, [Your Social Networking Credit Score](#), Slate, January 30, 2013

- *Why does Morozov see a moral hazard in the use of non-traditional analytics for credit scoring? Is his argument convincing?*

Optional

Lamont Wood, [Algorithms and Experiments Make Strange Bedfellows at SXSW](#), ComputerWorld, May 14, 2016

Robert Epstein, [How Google Could Rig the 2016 Election](#), Politico, August 19, 2015

Tristan Harris, [How Technology Hijacks People's Minds—from a Magician and Google's Design Ethicist](#), Medium, May 18, 2016

FAIRNESS

8. ALGORITHMIC MARKET POWER

What issues arise when analytics are used for dynamic pricing? Should we be concerned about algorithmic monopolies or other anti-competitive practices?

Salil Mehra, [Antitrust and the Robo-Seller: Competition in the Time of Algorithms](#), 100 Minnesota Law Review (2015), pp. 1-5, 13-18, 43-55

- *What does Mehra mean by "robo-sellers"?*
- *Why are robo-sellers a concern for antitrust law?*
- *How should we respond to possible anticompetitive applications of algorithmic pricing?*

Ryan Calo, [Digital Market Manipulation](#), 82 George Washington Law Review 995 (2014), pp. 1003-1012, 1020-24

- *How do online intermediaries, in Calo's account, engage in forms of manipulation?*

Optional

The White House, [Big Data and Differential Pricing](#) (2015), pp. 8-19

Ann Carrns, [When Good Drivers Pay More for Insurance Than Bad Ones](#), N.Y. Times Bucks Blog, January 29, 2013

Jennifer Valentino-Devries, et al, [Websites Vary Prices, Deals Based on Users' Information](#), Wall St. Journal, December 24, 2012

9. ALGORITHMIC DISCRIMINATION

When is a differential effect a neutral reflection of the state of the world, and when is it tantamount to illegitimate discrimination? The use of analytics has the potential both to counteract and to reinforce systematic biases against women, people of color, and other groups.

Solon Barocas & Andrew Selbst, [Big Data's Disparate Impact](#), California Law Review (2015), pp. 677-93

- *What are the ways that the authors believe big data can produce discrimination?*

Optional

Latanya Sweeney, [Discrimination in Online Ad Delivery](#), ACM Queue, April 2, 2013

Julia Angwin et al, [Machine Bias](#), ProPublica, May 23, 2016

Jennifer Stark and Nicholas Diakopoulos, [Uber Seems to Offer Better Service in Areas With More White People. That Raises Some Tough Questions](#), Washington Post, March 10, 2016

10. ADDRESSING DISPARATE IMPACT

The primary legal theory for unintentional discrimination is known as “disparate impact” in the United States. Can it apply to cases of algorithmic discrimination? Should it? Are there technical alternatives?

Solon Barocas & Andrew Selbst, [Big Data's Disparate Impact](#), California Law Review (2015), pp. 694-712

- *How well does the legal regime for indirect discrimination apply to the effects of analytics? Where does it fail?*

Suresh Venkatsubramanian, [Algorithmic Fairness: From Social Good to a Mathematical Framework](#), LSE Media Policy Project Blog, June 14, 2016

- *Do you think technical approaches such as those described by the author can be effective?*

Optional

Julia Angwin, Surya Mattu, and Jeff Larson, [The Tiger Mom Tax: Asians Are Nearly Twice as Likely to Get a Higher Price from Princeton Review](#), ProPublica, September 1, 2015

Don Peck, [They're Watching You at Work](#), The Atlantic, November 20, 2013

Jeremy Kun, [One Definition of Algorithmic Fairness: Statistical Parity](#), October 19, 2015

Jack Clark, [Artificial Intelligence Has a "Sea of Dudes" Problem](#), Bloomberg Technology, June 23, 2016

11. ALGORITHMIC INEQUALITY

As information and insight from data become more evaluable, should be concerned that their benefits are not universally available? Will analytics worsen the gap between haves and have-nots?

Jonas Lerman, [Big Data and its Exclusions](#), 66 Stanford Law Review Online 55, September 3, 2013

- *How are the author's concerns about digital exclusion different from, and similar to, the issues of digital discrimination?*

Kate Crawford, [The Hidden Biases in Big Data](#), Harvard Business Review, April 1, 2013

- *What did the services the author describes do wrong?*

Peter Swire, [Lessons From Fair Lending Law for Fair Marketing and Big Data](#) (white paper), September 11, 2014, Part III (pp. 6-11)

- *Does fair lending law provide a good model to address the “data divide”?*

Optional

David Ingold and Spencer Soper, [Amazon Doesn't Consider the Race of Its Customers. Should It?](#), Bloomberg, April 21, 2016,

Jen Schradie, [Big Data Not Big Enough? How the Digital Divide Leaves People Out](#), MediaShift, July 31, 2013

John Markoff, [Microsoft Finds Cancer Clues in Search Queries](#), N.Y. Times, June 6, 2016

RESPONSES

12. GUEST LECTURE: GERARD (JERRY) LEWIS, COMCAST

Gerard Lewis is Senior Vice President, Deputy General Counsel & Chief Privacy Officer for Comcast where he advises company management and business units on intellectual property, technology, e-commerce, Internet, privacy, and security law matters.

13. MOVING FORWARD

How to balance the various interests involved to develop responsible approaches to business analytics. Can industry be part of the solution, and not just the problem?

Neil Richards & Jonathan King, [Big Data and the Future for Privacy](#), Handbook of Research on Digital Transformations (2016), pp. 16-21

- *What is “soft regulation” and how is it different from ordinary regulation?*

Solon Barocas and Andrew Selbst, [Big Data's Disparate Impact](#), California Law Review (2015), pp. 729-32

- *How do these author's recommendations (focused on discrimination) compare to those of Richard and King (focused on privacy)?*

Christian Sandvig et al, [An Algorithm Audit](#) (2014)

- *How effective do you think the audit method described in the paper can be at identifying and helping to address problems with algorithms?*

14. EXAM AND COURSE WRAP-UP

The closed-book examination will be done in class, and will cover all prior sessions. Additional information on the format of the exam will be provided during the course.