



LGST 6420

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

Fall 2023 Quarter 2

Ubiquitous data and increasingly powerful machine learning technologies create massive opportunities for both financial gain and social good. They also create dangers such as privacy violations 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 artificial intelligence, big data, predictive analytics, and related techniques of algorithmic decision-making.

Note: This is a fast-evolving subject area. The course is revised frequently. Make sure to check Canvas for the most up-to-date readings and materials.

Instructor

Kevin Werbach (werbach@wharton.upenn.edu)

Liem Sioe Liong/First Pacific Company Professor and Chair, Department of Legal Studies and Business Ethics

Learning Objectives

Good data-driven decision-making means not just generating solutions, but understanding how to use them. Sophisticated firms have frequently gotten into trouble over privacy, security, manipulation, and bias. Failure to anticipate such issues can result in ethical lapses, public relations disasters, regulatory sanctions, and legal liability.

My goal is to help you develop the skills to use analytics in a responsible way, while remaining focused on your business objectives. After completion of the course, you should be able to:

1. Identify the limitations and flaws of algorithmic decision-making systems.
2. Anticipate legal or ethical controversies arising from applications of business analytics.
3. Evaluate mechanisms to promote algorithmic accountability from a variety of perspectives.
4. Avoid destroying the world, crashing the economy, or going to jail. (*Money-back guarantee not available.*)

Course Structure

This is a fast-paced, interactive course. It is only 14 sessions long, and there are a number of assignments you need to complete. These are part of the learning experience, not just tests. I have designed in significant flexibility to make the workload manageable, but it's your responsibility to stay on top of the requirements. Active participation is important for a good learning experience, especially in a remote environment.

Course components are organized in Canvas Modules. For most class sessions, I've recorded several short videos. This is the "lecture" component, to complete on your own prior to class, with embedded quiz questions to check your comprehension.

Class time is for interaction—live discussions, groupwork, and activities. Because you will have already watched the lecture videos beforehand, I will often end class sessions in less than 90 minutes, remaining available afterwards for questions.

Policy on Use of Generative AI

You may use AI tools such as ChatGPT to assist you, such as for research, understanding concepts, or developing initial ideas. (For the final project, use of AI is required.) AI can be a valuable means to facilitate your learning; however, it should not be an alternative to learning. It is a violation of Wharton ethics rules to submit work that is entirely created by someone else, or by software such as AI tools. Plagiarism or other forms of cheating will be dealt with severely under relevant university procedures.

Be aware that AI tools are imperfect. (That is a central topic of this course!) They may be biased, inaccurate, or “hallucinate” false information. You are responsible for the content of your submissions, so review any AI-generated material carefully.

If you use AI tools for assignments in this course, you are expected to disclose that in your response, and identify how you employed them. You should identify prompts you used, and the relationship of what the AI created to what you eventually submitted. (This material does not count against any word limits.)

Course Requirements and Grading

Lecture Video Quizzes [10%]

There are multiple-choice questions and discussion prompted embedded in the pre-recorded lecture videos. These are designed both to gauge your comprehension and to encourage you to think deeper about the content of the videos prior to class. You receive most of the points for watching the videos and completing the assignments.

Current Events [25%]

AI and related technologies are developing quickly. And the kinds of legal and ethical controversies we cover in the class are becoming increasingly common.

There will be a “Current Events” discussion on Canvas for each week of the course. Your assignment is to find and post an online article published within the past six months that illustrates the topics of one of that week’s sessions. More-recent articles are preferable, as are news items describing real-world examples (as opposed to opinion pieces). In addition to posting the article link, you must write at least 150 words explaining how it relates to what we covered in class, and offering your own reactions or comments. In addition to your own post, you must also comment on 2 posts by other students.

Posts must be submitted by midnight Friday night of each week, and comments must be submitted by midnight the following Monday.

Note: You are only required to complete this assignment 5 times in the 7 weeks of the course. Each submission is worth 5 points. If you complete the assignment more than 4 times, you will receive your 4 highest scores.

Participation [20%]

Participation is an important aspect of this course. It is not just a matter of memorizing information; I want you to engage with these novel issues and form your own conclusions.

At the end of every class session, I will display a QR code for completion of an "exit ticket". This will be how I track attendance, and also seek your feedback and interpretation of the session. (Late arrivals will be noted manually, and if you need to leave class early, please let me know at the beginning of the session.)

Attendance is only a minority of your participation score. The majority will be based on substantive contributions to class discussions, and optionally, session reflections and comments on Canvas discussions. Substantive responses to posts by other students are encouraged, and may even earn you more participation credit than brief original posts. Participation credit will be assessed holistically: quality matters, not just quantity.

Final Project [45%]

For your final project, I am asking you to employ generative AI to examine the potential and limits of the technology. You have a choice of three options for the assignment. Each should be 1500-3000 words in length, not counting footnotes, endnotes, bibliography, or generative AI prompts and explanatory material. You are expected to reference course readings, videos, and outside materials to support your analysis; you may choose the format for those citations.

The final project is due at midnight the following Thursday after the course concludes.

Option 1. Improve on AI-generated policy

Use a generative AI tool to create a 500-word AI ethics policy for a company. (You may pick an actual company or a hypothetical one in a particular industry.) Submit the prompts you used. Then, improve the policy. Edit the AI-generated version using Microsoft Word's or Google's markup features, to show the specific changes, additions, and deletions you make. Then, write an analysis justifying the changes you made, with references to support your arguments. You should explain both what areas your policy covers, and the effectiveness of the specific guidelines or principles you select.

Option 2. Re-design the assessments of another course to be robust to generative AI

Pick a course you have taken at Penn and attach the syllabus to your answer. Write a memo that explains how a student could use generative AI to complete some or all of the course assignments. Then, offer an alternative (or alternatives) to those assignments that would be difficult for students to complete without thoughtfully engaging the material and displaying their learning. Describe why your alternative(s) will be effective ways for students to demonstrate authentic learning, given the availability of generative AI. Finally, discuss what generative AI providers, or Penn, or both, should do to address ethical concerns about use of generative AI by students.

Option 3. Black Mirror Episode

Use a generative AI to create a new episode of the Netflix show "Black Mirror" that illustrates one of the major concepts or session topics of this course. Identify which tool you used. Submit both your prompts and the episode plot generated (not included in the word limit).

Then, evaluate how effectively the episode illustrates the concepts of the course, and how useful it would be as a learning tool for students. Suggest what the episode misses or gets wrong, and how it could be improved.

Attendance Policy

Because this is an interactive, discussion-oriented course, attendance is important. As noted above, I will track attendance using in-class surveys. Attendance may be excused for the usual reasons (medical issue, personal/family emergency, religious observance) so long as you email me ahead of time with an explanation.

All class sessions will be recorded and made available to you on Canvas. If you are unable to attend a class, you may watch the video and complete the Exit Ticket within a week. Your overall attendance score will not be reduced significantly if you do so a limited number of time for excused absences.

Deadlines

I accept late submissions, up to the final class session of the course, with a graduated markdown. (With the exception of the "Current Events" assignments, because you have flexibility in which ones you complete.) Late submission of the final paper will similarly incur a graduated markdown until grades are submitted. If you encounter difficulty completing an assignment in a timely manner, please contact me before the deadline. I understand you may be facing unusual challenges in the current environment. I am always willing to listen and consider accommodations. Maximizing your learning is my primary goal.

Syllabus

Class	Readings
October 24 I, FOR ONE, WELCOME OUR NEW ROBOT OVERLORDS How might data science and AI change the relationships among firms, customers, employees, other firms, and governments? What are some of the legal or ethical concerns that may arise?	Kevin Roose, We Need to Talk About How Good A.I. Is Getting , New York Times, August 24, 2022 Andrew Paul, School District Uses ChatGPT to Help Remove Library Books , Popular Science, August 14, 2023 Graham Ruddick, Admiral to Price Car Insurance Based on Facebook Posts , The Guardian, November 1, 2016 Matthew Hutson, AI Could ID Gang Crimes, Opening an Ethical Firestorm , Science, February 28, 2018
October 26 ALGORITHMIC DECISIONMAKING All algorithmic systems, no matter how sophisticated, rely on human decisions about	R2D3, A Visual Introduction to Machine Learning

how data are collected, analyzed, and used. Failure to appreciate this can lead to problems.	Katherine J. Wu and Mike Isaac, Frontline Workers Were Left Off the Vaccine List at Stanford Medical Center in Palo Alto. They Fought Back., New York Times, December 18, 2020 Michael Luca et al, Algorithms Need Managers, Too, Harvard Business Review, January-February 2016
October 31 ACCURACY The first step to responsible use of AI and analytics is to appreciate limitations of their methods. Algorithmic decision-making is powerful, but not always effective or robust to changes.	Will Parker and Konrad Putzier, What Went Wrong With Zillow? A Real-Estate Algorithm Derailed Its Big Bet, Wall Street Journal, November 17, 2021 David Lazer et al, The Parable of Google Flu, Science, March 14, 2014 Tapani Rinta-Kahila et al, How to Avoid Algorithmic Decision-Making Mistakes: Lessons From the Robodebt Debacle, Momentum Magazine, March 16, 2022 Benjamin Weiser and Nate Schweber, The ChatGPT Lawyer Explains Himself, New York Times, June 8, 2023
November 2 TRANSPARENCY How well can we assess exactly what algorithms are doing, and why?	Jenna Burrell, How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms, Big Data & Society, January-June, 2016 Houston Federation of Teachers v. Houston Ind. School District (S.D. Texas, May 4, 2017) – <i>edited version</i> Investigation: How TikTok's Algorithm Figures Out Your Deepest Desires, Wall Street Journal, July 21, 2021 (13 minute video)
November 7 RISK AND RESPONSIBILITY Algorithmic systems may produce unintended results, which in some cases cause harm. Who should be held liable? How can organizations using analytics manage those risks?	Stefanie Dazio and tom Krisher, As a Criminal Case Against a Tesla Driver Wraps Up, Legal and Ethical Questions on Autopilot Endure, Associated Press, August 15, 2023 Karen Hao, When Algorithms Mess Up, the Nearest Human Gets the Blame, MIT Technology Review, May 28, 2019 Future of Life Institute, Pause Giant AI Experiments: An Open Letter, March 22, 2023
November 9 FAIRNESS The use of AI and analytics has the potential both to counteract and to reinforce systematic	Melissa Heikkila, AI Language Models Are Rife With Different Political Biases, MIT Technology Review, August 7, 2023 Emily Bembeneck, Rebecca Nissan, and Ziad Obermeyer, To Stop Algorithmic Bias, We First Have to Define It, Brookings, October 21, 2021

biases. But what exactly does it mean for an algorithmic system to be “fair”?	Harini Suresh and John Gutttag, A Framework for Understanding Unintended Consequences of Machine Learning (January 2019) Julia Angwin et al, Machine Bias , ProPublica, May 23, 2016
November 14 DISCRIMINATION When are differential effects of AI and analytics tantamount to illegitimate or illegal discrimination?	Ricci v. DeStefano , 557 U.S. 557 (2009) – <i>edited version</i> Texas Dept. of Housing and Community Affairs v. Inclusive Communities Project (2015) – <i>edited version</i> Katie Benner, Glenn Thrush, and Mike Isaac, Facebook Engages in Housing Discrimination With Its Ad Practices , U.S. Says, New York Times, March 28, 2019 Mike Isaac, Meta Agrees to Alter Ad Technology in Settlement With U.S. , New York Times, June 21, 2022
November 16 DATA COLLECTION Are there limits on how data should be collected, used, and shared? What are the legal or ethical concerns about the data collection and aggregation practices of AI and analytics tools?	Solon Barocas and Helen Nissenbaum, Big Data’s End Run Around Procedural Privacy Protections , Communications of the ACM (November 2014) Kashmir Hill and Surya Mattu, How a Company You’ve Never Heard of Sends You Letters about Your Medical Condition , Gizmodo, June 19, 2017 Anna Tong, OpenAI’s Sam Altman Launches Worldcoin Crypto Project , Reuters, July 24, 2023 Alex Reisner, Revealed: The Authors Whose Pirated Books Are Powering Generative AI , The Atlantic, August 19, 2023
November 21 PRIVACY What are the privacy challenges that AI and analytics create? And of sensitive attributes can be inferred from other data, does it even make sense to talk about privacy any more?	Charles Duhigg, How Companies Learn Your Secrets , N.Y. Times Magazine, Feb. 16, 2012 Kashmir Hill, The Secretive Company that Might End Privacy as We know It , N.Y. Time, January 18, 2020 Thorin Klosowski, The State of Consumer Data Privacy Laws in the US (And Why It Matters) , New York Times, September 6, 2021 Matt Burgess, ChatGPT Has a Big Privacy Problem , Wired, April 4, 2023
November 28 MANIPULATION To what extent does analysis itself influence behavior? And what are the limits on using	Vindu Goel, Facebook Tinkers With Users’ Emotions in News Feed Experiment, Stirring Outcry , New York Times, June 29, 2014 Matteo Wong, We Haven’t Seen the Worst of Fake News , The Atlantic, December 20, 2022

analytics not merely to understand and predict customer actions, but to shape them?	Chloe Xiang, AI Tasked With 'Destroying Humanity' Now 'Working on Control Over Humanity Through Manipulation' , Motherboard, April 12, 2023
November 30 SOCIETAL ISSUES AI raises many broader concerns, such as potential displacement of jobs, overwhelming government surveillance or behavior-shaping, and even the extinction of humanity.	Sarah Kessler, The A.I. Revolution Will Change Work. Nobody Agrees How, New York Times, June 10, 2023 Marc Andreessen, Why AI Will Save the World, Substack, June 6, 2023 Nicole Kobie, The Complicated Truth About China's Social Credit System, Wired, July 6, 2019 Khadeeja Safdar, On Hold for 45 Minutes? It Might Be Your Secret Customer Score, Wall St. Journal (2018)
December 5 ALGORITHMIC ACCOUNTABILITY How can governments and firms best respond to the challenges we've discussed in the course?	Future of Privacy Forum, Beyond Explainability: A Practical Guide to Managing Risk in Machine Learning Models (June 2018) Lauren Weber, New York City Starts to Regulate AI Used in Hiring Tools, Wall Street Journal, July 5, 2023 Biden Administration, Blueprint for an AI Bill of Rights, October 2022 Marietje Schaake, The European Commission's Artificial Intelligence Act, Stanford HAI, June 2021