

# Ethics and Algorithms

## (Abridged Syllabus)

*Note to the reader: this syllabus is for a small seminar of 15–25 advanced undergraduate students who have taken at least one course in moral or political philosophy.*

### Course Description

Computer algorithms play an increasingly central role in our lives. They help decide what movies we watch, what news stories we read, what jobs we get, and how much we pay for health insurance. In this seminar, we will consider how to make decisions with algorithms in a morally responsible way. Should social media companies delete “fake news” from their platforms? Should we permit internet service providers to collect our web browsing data and sell it to advertisers? Should public health organizations use algorithms to decide which people to help? Do algorithms used by the criminal justice system wrongfully discriminate against African Americans? In exploring these issues, we will draw on the work of philosophers, social scientists, journalists, legal scholars, and computer scientists. Students will also conduct three “algorithm audits” of real-world automated decision-making systems.

### Assignments and Grading

You must complete *all* required assignments in order to pass this course. Your grade will be determined as follows:

- 20%: Participation
- 15%: Short paper (4–5 pages)
- 25%: Group project (algorithm audit)
- 40%: Final paper (10–15 pages)

### Schedule and Readings

#### *Unit 1: Introduction*

#### 1. Wed 9/2 Introduction to the course

- James Moor, “What is Computer Ethics?”
- Zeynep Tufekci, “Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency”

#### 2. Mon 9/7 No class (Labor Day)

#### 3. Wed 9/9 Can ethics be computed?

- Susan Leigh Anderson and Michael Anderson, “Machine Ethics: Creating an Ethical Intelligent Agent” (*AI Magazine*)
- Boer Deng, “The Robot’s Dilemma” (*Nature*)

4. **Mon 9/14 Can ethics be computed soon?**

- Michael Anderson, Susan Leigh Anderson, and Chris Armen, [“Towards Machine Ethics”](#)
- Cathy O’Neill, [On Being a Data Skeptic](#) (essay)

5. **Wed 9/16 What are we evaluating?**

- Langdon Winner, [“Do Artifacts Have Politics?”](#)
- William Langewiesche, [“What Really Brought Down the Boeing 737 Max?”](#) (*The New York Times Magazine*)

*Unit 2: Cost-benefit Analysis and Its Limits*

6. **Mon 9/21 Effective Altruism I**

- William MacAskill, [Doing Good Better](#), chapters II-IV

7. **Wed 9/23 Effective Altruism II**

- William MacAskill, [Doing Good Better](#), chapters V and VI

8. **Mon 9/28 Thinking about risk I**

- Nick Bostrom, [“Existential Risk Prevention as Global Priority”](#)
- Allan Dafoe and Stuart Russell, [“Yes, We Are Worried About the Existential Risk of Artificial Intelligence”](#) (*MIT Technology Review*)

9. **Wed 9/30 Thinking about risk II**

- Amia Srinivasan, [“Stop the Robot Apocalypse”](#) (*The London Review of Books*)
- Cass Sunstein, [“Cost-Benefit Analysis and the Environment”](#)

10. **Mon 10/5 Limits of cost-benefit analysis**

*Short paper due*

- Steven Kelman, [“Cost-benefit Analysis: An Ethical Critique”](#)
- David Schmidtz, [“A Place for Cost-Benefit Analysis”](#)

*Unit 3: Auditing Algorithms*

11. **Wed 10/7 Auditing algorithms**

- Dillon Reisman et al., [“Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability”](#) (AI Now Institute report)
- William Kymlicka, [“Moral Philosophy and Public Policy: The Case of NRTs”](#)

12. **Mon 10/12 No class (Indigenous People’s Day)**

13. **Wed 10/14 Algorithm audit I: TND Network**

- USC Center for Artificial Intelligence in Society, [“Social Networks and Substance Abuse Prevention For Homeless Youth”](#) (webpage with videos)
- Thomas Valente et al., [“Peer Acceleration: Effects of a Social Network Tailored Substance Abuse Prevention Program Among High Risk Adolescents”](#) (*Addiction*)
- The National Association of Social Workers’ [Code of Ethics](#)

**14. Mon 10/19 Algorithm audit I debrief**

- John Harris, “QALYfying the Value of Life”

*Unit 4: Autonomy and Privacy*

**15. Wed 10/21 Fake news and free speech**

- John Stuart Mill, *On Liberty*, chapter II (“Of the Liberty of Thought and Discussion”)
- David Lazer et al., “The Science of Fake News” (*Science*)
- The New York Times Editorial Board, “Facebook and the Digital Virus Called Fake News” (*The New York Times*)

**16. Mon 10/26 Autonomy and manipulation**

- Jia Tolentino, “How TikTok Holds Our Attention” (*The New Yorker*)
- John Stuart Mill, *On Liberty*, chapter III (“On Individuality As One of the Elements of Well-being”)

**17. Wed 10/28 Why should we care about privacy?**

- James Rachels, “Why Privacy is Important”
- Moxie Marlinspike, “Why ‘I Have Nothing to Hide’ Is the Wrong Way to Think About Surveillance” (*Wired*)
- Fahrad Manjoo, “I Visited 47 Sites. Hundreds of Trackers Followed Me” (*The New York Times*)

*Unit 5: Fairness and Discrimination*

**18. Mon 11/2 What is fairness? What is discrimination?**

- Deborah Hellman, *When Is Discrimination Wrong?*, chapter I (“The Basic Idea”)
- Batya Friedman and Helen Nissenbaum, “Bias in Computer Systems”

**19. Wed 11/4 Big data’s disparate impact**

*Group projects due*

- Latanya Sweeney, “Discrimination in Online Ad Delivery”
- Solon Barocas and Andrew Selbst, “Big Data’s Disparate Impact,” pages 671–693
- Marten Wattenberg et al., “Attacking Discrimination with Smarter Machine Learning” (interactive)

**20. Mon 11/9 Group project presentations**

- No readings.

**21. Wed 11/11 The COMPAS algorithm I**

- Julia Angwin et al., “Machine Bias” (*ProPublica*)
- Karen Hao et al., “Can You Make AI Fairer Than a Judge? Play Our Courtroom Algorithm Game” (*MIT Technology Review*) (interactive)

**22. Mon 11/16 The COMPAS algorithm II**

- Robert Long, “Fairness in Machine Learning: Against False Positive Rate Equality as a Measure of Fairness” (manuscript), sections 1-3
- Equivant’s [response to ProPublica](#)

**23. Wed 11/18 The COMPAS algorithm III**

- Robert Long, “Fairness in Machine Learning: Against False Positive Rate Equality as a Measure of Fairness” (manuscript), sections 4 and 5
- Deborah Hellman, “Measures of Algorithmic Fairness” (excerpts)

*Unit 6: Conclusion*

**24. Mon 11/23 Algorithm audit II: The Allegheny Family Screening Tool**

- Rhema Vaithianathan et al., “[Developing Predictive Risk Models to Support Child Maltreatment Hotline Screening Decisions: Allegheny County Methodology and Implementation](#)” (first report only)

**25. Wed 11/25 Algorithm audit II debrief**

- Virginia Eubanks, *Automating Inequality*, chapter IV (“The Allegheny Algorithm”)
- Tim Dare and Eileen Gambrell, “[Ethical Analysis: Predictive Risk Models at Call Screening for Allegheny County](#)” (second report in the document)
- The Department of Human Services’ [official response to Automating Inequality](#)

**26. Mon 11/30 No class (Thanksgiving Break)**

**27. Wed 12/2 Conclusion**

*Final paper due*