Computer Science 330:
Ethical Issues in Data Science
Fall 2020

Instructor: Cameron Domenico Kirk-Giannini
Format: Asynchronous Remote
Office Hours: Tuesdays 2–3 PM
Email: camerondomenico.kirkgiannini@rutgers.edu

COURSE DESCRIPTION

In performing his or her role in the workplace, the practicing computer or data scientist confronts a number of important ethical questions: questions about the possibility of bias in automated decision-making systems, about what constitutes appropriate collection, aggregation, and use of personal information about the users of technology services, and about collective and individual responsibility for the social impacts of newly developed technologies. The purpose of this course is to consider these questions directly in order to prepare students to think critically about the ethical questions which will arise in connection with their work later in their lives.

In our first unit, we will discuss the problem of bias in machine learning, focusing especially on the controversy surrounding whether COMPAS, an algorithm used in the US judicial system to make bail and sentencing decisions, is biased against Black defendants. In our second unit, we will discuss the conflict between consumers who wish to make use of technological services while maintaining their privacy and corporations who wish to obtain as much information as possible in exchange for providing their services. In particular, we will consider a number of theories of what privacy is and why it is valuable, survey some of the ways privacy can be compromised by new technologies, and then engage with some suggestions for how to manage privacy concerns in light of these technologies. In our third unit, we will consider the question of corporate and individual responsibility for the harms caused by new technologies. We will focus in this unit on the case study of China's growing surveillance infrastructure, asking to what extent the corporations and individual engineers working on the technologies which make surveillance possible are collectively or individually responsible for the harms it causes.

LEARNING OUTCOMES

- By critically engaging with the course material, students will gain a detailed understanding of some of the most important ethical issues relevant to the field of data science.

- The topics covered in the course will also serve as convenient introductions to some major concepts in value theory such as: justice, fairness, bias, privacy, consent, moral responsibility, collective agency, and complicity.

- Through class discussion and structured writing exercises, students will develop crucial philosophical abilities like reconstructing and evaluating arguments, articulating ideas in conversation, and writing clearly and cogently.

TEXTBOOK

There will be no textbook for this course. All readings will be made available online.

LEARNING MATERIALS

I will upload lecture videos, slides, course handouts, and review materials to the course's Blackboard site.

ASSIGNMENTS AND GRADING

There will be three significant course requirements. First, students will be required to work together in small groups to produce weekly responses to the course readings. Second, there will be one 1200-1500 word paper written in response to a prompt related to the themes of the course. Third, there will be a take-home final examination.

Grades will be determined as follows:

- Weekly Assignments: 40%

- Paper: 30%

- Final Exam: 30%

Grading Scale:

- A = 89.5-100

- B+ = 84.5-89.49

- B = 79.5-84.49

- C+ = 74.5-79.49

- C = 69.5-74.49

- D = 59.5-69.49

- F = 0-59.49

SEMESTER OVERVIEW

Week 1 (Week of Tuesday 9/1):
Reading: None (Course Introduction)

PART I: ALGORITHMS AND BIAS

Week 2 (Week of Monday 9/7):
Readings:

1. Batya Friedman and Helen Nissenbaum, "Bias in Computer Systems."

Week 3 (Week of Monday 9/14):
Readings:

1. Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification"

2. Jeffrey Dastin, "Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women"

Week 4 (Week of Monday 9/21):
Readings:

1. Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan, "Semantics Derived Automatically from Language Corpora Contain Human-like Biases."

2. Safiya Umoja Noble, "Missed Connections: What Search Engines Say About Women."

Week 5 (Week of Monday 9/28):

Readings:

1. Blaise Agüera y Arcas, Margaret Mitchell, and Alexander Todorov, "Physiognomy's New Clothes."

2. Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks."

Week 6 (Week of Monday 10/5):

Readings:

1. Sam Corbett-Davies, Emma Pierson, Avi Feller, and Sharad Goel, "A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear."

2. Anthony W. Flores, Kristin Bechtel, and Christopher T Lowenkamp, "False Positives, False Negatives, and False Analyses: A Rejoinder to 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks.'"

PART II: PRIVACY AND CONSENT

Week 7 (Week of Monday 10/12):

1. Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Chapter 4 ("Locating the Value in Privacy")

2. Herman T. Tavani, "Informational Privacy: Concepts, Theories, and Controversies."

Week 8 (Week of Monday 10/19):

Readings:

1. Anita Allen, "Protecting One's Own Privacy in a Big Data Economy."

2. Daniel J. Solove, "Privacy Self-Management and the Consent Dilemma."

Week 9 (Week of Monday 10/26):

Readings:

1. Yabing Liu et. al., "Analyzing Facebook Privacy Settings: User Expectations vs. Reality."

2. Michal Kosinski, David Stillwell, and Thore Graepel, "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior."

Week 10 (Week of Monday 11/2):

Readings:

1. Leslie K. John, Alessandro Acquisti,and George Loewenstein, "Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information."

2. Alessandro Acquisti, Laura Brandimarte, and George Loewenstein, "Privacy and Human Behavior in the Age of Information."

PART III: HARM, COMPLICITY, AND RESPONSIBILITY

Week 11 (Week of Monday 11/9):

Readings:

1. Phillip Pettit, "Responsibility Incorporated."

Week 12 (Week of Monday 11/16):

Readings:

1. Christopher Kutz, *Complicity: Ethics and Law for a Collective Age*, Chapter 5 ("Complicitous Accountability")

2. Andrew Clapham and Scott Jerbi, "Categories of Corporate Complicity in Human Rights Abuses."

Week 13 (Week of Monday 11/23):

(SHORT WEEK: THANKSGIVING)

Readings:

1. Kristina Wilson, Yaneli Ramos, Daniel Harvey, and Wayne Norman, "Google in China: 'The Great Firewall'."

2. Simon Denyer, "China's scary lesson to the world: Censoring the Internet works."

Week 14 (Week of Monday 11/30):

1. Simon Denyer, "Beijing Bets on Facial Recognition in a Big Drive for Total Surveillance."

2. Paul Mozur and Aaron Krolik, "A Surveillance Net Blankets China's Cities, Giving Police Vast Powers."

3. Paul Mozur and Nicole Perlroth, "China's Software Stalked Uighurs Earlier and More Widely, Researchers Learn."

4. Ryan Gallagher, "How U.S. Tech Giants Are Helping to Build China's Surveillance State."

PAPER DUE FRIDAY 12/4 5:00 PM EST

Week 15 (Week of Monday 12/7):

Readings: None (Review and Take-Home Final)

FINAL EXAM TUESDAY 12/8 5:00 PM EST – FRIDAY 12/10 5:00 PM EST (72 HOURS)

WEEKLY ASSIGNMENTS

Each week, you will be responsible for crafting a written response to the course readings together with the other members of your small group. Groups will consist of approximately five students. You will be expected to work collaboratively on the Slack platform, and in addition to grading the responses you submit, I will be monitoring your Slack groups to ensure that everyone is participating in the writing process. I will shuffle the groups around a number of times in the course of the semester — my hope is that this will enable you to get to know at least some of the other students in the course!

COURTESY

It is important that all discussion on Slack be conducted calmly and respectfully. Professional courtesy and consideration for

our classroom community are especially important with respect to topics dealing with differences such as race, color, gender and gender identity/expression, sexual orientation, national origin, religion, disability, age, and veteran status.

Meaningful and constructive dialogue requires mutual respect, a willingness to listen, and tolerance of opposing points of view. Respect for individual differences and alternative viewpoints will be maintained at all times in this class. Our choices of words and use of language are critical components of respectful discourse as we work together to achieve the full benefits of creating a classroom in which all people can feel comfortable expressing themselves.

LATENESS POLICY

Assignments may be turned in late, but will be subject to a one-letter-grade penalty per day late (a one-letter-grade drop is the difference between an A and a B and between a B+ and a C+).

If you would like an extension on an assignment, you must notify me at least 24 hours in advance of the deadline for that assignment. Extensions will be granted on a discretionary basis.

ACADEMIC INTEGRITY

As an academic community dedicated to the creation, dissemination, and application of knowledge, Rutgers University is committed to fostering an intellectual and ethical environment based on the principles of academic integrity. Academic integrity is essential to the success of the University's educational and research missions, and violations of academic integrity constitute serious offenses against the entire academic community.

Academic Integrity Policy: http://academicintegrity.rutgers.edu/academicintegrity-policy/

COLLABORATION

There are two kinds of collaborators. *Reviewers* offer comments on a pre-existing piece of writing. They are not responsible for the writing or major aspects of the argument of the pieces on which they comment. You are encouraged to seek reviewers for your pieces of writing. Please acknowledge your reviewers by name in a footnote at the end of your essay.

Co-authors enter into the creative process with the author. They are responsible for significant parts of the writing and/or argument of the pieces which they co-author.

Your group reading responses are expected to be co-authored pieces. However, you are not permitted to submit co-authored essays or examinations for credit in this course.

CORRUPT FILES

It is your responsibility to ensure that you provide your written work to me in a format I can access. If you submit a corrupt file, I will do my best to notify you as soon as possible. It is then your responsibility to send me another copy of the file in a timely manner. This means it is very important that you open and read emails from me. If I request a second copy of a file and you do not reply in a timely manner, I will regard the assignment as not having been submitted.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES

Every effort will be made to accommodate students who present a valid Letter of Accommodations. For more information, see: https://ods.rutgers.edu/my-accommodations/letter-of-accommodations

RELIGIOUS OBSERVANCE

I am happy to accommodate special needs related to students' religious practices. However, I require that you notify me in writing within the first two weeks of class if you will need such accommodation at any point during the semester.

COUNSELING SERVICES

Counseling services are available at the Counseling Center, Room 101, Blumenthal Hall. For more information, call (973) 353-5805 or visit http://counseling.newark.rutgers.edu/. Please note that I am required to report certain sensitive information you might relate to me to the University.

NO EXTRA CREDIT

There will be no opportunities to earn extra credit in this course.