

**DATA SCIENCE 410:
ETHICAL ISSUES IN DATA SCIENCE
M/W 10:00–11:20 AM
HILL HALL 126**

Instructor: Cameron Domenico Kirk-Giannini
Office: Conklin Hall 417
Office Hours: Mondays 1:00-2:00 PM (email for appointment)
Email: camerondomenico.kirkgiannini@rutgers.edu

COURSE DESCRIPTION

In performing his or her role in the workplace, the 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 deeply 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 an 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 is no textbook for this course. All readings will be made available online.

LEARNING MATERIALS

I will upload course handouts to the course's Canvas site.

ASSIGNMENTS AND GRADING

There will be three significant course requirements. First, you will be required to attend class and participate in discussion of the course material. Second, there will be two in-class papers. Third, there will be an in-class final examination.

Grades will be determined as follows:

- Attendance and class participation: 30%
- Papers 1 and 2 (in-class): 40%
- Final exam (or debate — see explanation below): 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

Debate Option:

Students have the option of participating in a 30-minute, 4-person in-class debate instead of or in addition to taking the final exam. The debate format will be a shortened version of Policy Debate. Students may form their own teams of two and may select a pre-approved prompt or, in consultation with me and another team, design their own prompt. Students who participate in a debate will receive a grade reflecting the quality of their arguments and may choose to apply this grade to their final exam. **For example, a student who receives an A grade on the debate may skip the final exam and automatically receive an A on it.** This is optional: students who participate in a debate are still free to take the final exam, in which case they will receive whichever grade is higher.

SEMESTER OVERVIEW

Week 1: 1/17

Reading: None (Course Introduction)

PART I: ALGORITHMS AND BIAS

Week 2: 1/22 and 1/24

Readings:

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

Week 3: 1/29 and 1/31

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

Week 4: 2/5 and 2/7

Readings:

1. Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan, "Semantics Derived Automatically from Language Corpora Contain Human-like Biases."
2. Blaise Agüera y Arcas, Margaret Mitchell, and Alexander Todorov, "Physiognomy's New Clothes."

Week 5: 2/12 and 2/14

Readings:

1. Katie Crawford, *Atlas of AI*, Chapter 4, ("Classification").
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: 2/19 and 2/21

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: DATA, PRIVACY, AND CONSENT

Week 7: 2/26 and 2/28

NOTE: PAPER 1 IN CLASS ON 2/26

Readings:

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

Week 8: 3/4

NOTE: CLASS ON 3/4 BY ZOOM; NO CLASS ON 3/6

Readings:

1. Katie Crawford, *Atlas of AI*, Chapter 3, ("Data").

Week 9: 3/18 and 3/20

Readings:

1. Daniel J. Solove, "Privacy Self-Management and the Consent Dilemma."
2. Anita Allen, "Protecting One's Own Privacy in a Big Data Economy."

Week 10: 3/25 and 3/27

Readings:

1. Michal Kosinski, David Stillwell, and Thore Graepel, "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior."
2. Yabing Liu et. al., "Analyzing Facebook Privacy Settings: User Expectations vs. Reality."
3. Leslie K. John, Alessandro Acquisti, and George Loewenstein, "Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information."
4. Alessandro Acquisti, Laura Brandimarte, and George Loewenstein, "Privacy and Human Behavior in the Age of Information."

**PART III: HARM, COMPLICITY,
AND RESPONSIBILITY**

Week 11: 4/1 and 4/3

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 12: 4/10

NOTE: NO CLASS ON 4/8

Readings:

1. Katie Crawford, *Atlas of AI*, Chapter 6, ("State").

Week 13: 4/15 and 4/17

Readings:

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."

Week 14: 4/22 and 4/24

NOTE: PAPER 2 IN CLASS ON 4/24

Readings: TBD

4/22: Topic of Choice/Catchup Due to Debates

Week 15: 4/29

Readings: None

4/29: Final Exam Review Session

ATTENDANCE

During each class meeting, I will distribute a unique attendance code to each student present. You are responsible for filling out the online attendance form with your name, email, and unique attendance code. To do this, visit <http://www.cd.kg/teaching/> and click on the 'Attendance' link next to the name of this course. **Forms must be completed within 24 hours of the end of each class session for full credit.**

COURTESY

It is important that all discussion 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.

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/>

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.