Ethics, Data, and Technology (PHI 3681)

COURSE DESCRIPTION

In this course, we will explore questions about how emerging technologies should be designed and regulated. What does it mean to say that an algorithm is "biased" against members of a particular social group? Should we be concerned about the fact that technology companies such as Facebook gather vast amounts of data about our online activities? What does it mean to say that a machine learning algorithm is a "black box," and is there something unfair about using such algorithms to decide how to treat people? As we investigate these and other questions about emerging technologies, we will draw on concepts and readings from a variety of different fields, including philosophy, economics, computer science, data science, and law.

MEETING TIMES AND LOCATION

T 10:40 AM–11:30 AM
R 10:40–12:35 PM
Anderson Hall 32

INSTRUCTOR

David Gray Grant, Ph.D.
Assistant Professor
Department of Philosophy
University of Florida
david.grant@ufl.edu
Office hours: T 11:30 AM-12:30 PM and 3:50-4:50 PM
Office location: 318 Griffin-Floyd Hall

REQUIRED MATERIALS

All materials will be posted on Canvas.

COURSE OBJECTIVES

1. Develop a basic vocabulary for discussing the ethical dimensions of data science and its applications.
2. Analyze issues and policies concerning emerging "big data" technologies through the application of ethical concepts.
3. Critique public policies, social practices, and social institutions that shape, and are shaped by, scientific discovery and technological design.
4. Discern the structure of arguments, represent them fairly and clearly, and evaluate them for cogency.
5. Formulate original arguments, anticipate objections, and respond in a conscientious fashion.
6. Read and discuss complex philosophical texts from both historical sources and contemporary works.
7. Speak and write clearly and persuasively about abstract and conceptually elusive matters.

COURSE WEBSITE

This course is supplemented by online content in the Canvas e-Learning environment. PDF readings, an electronic copy of the syllabus, and assignment submission portals can be found on the course website.

- To login to the e-Learning site for this course, go to https://lss.at.ufl.edu/, click the e-Learning in Canvas button, and on the next page enter your Gatorlink username and password. You can then access the course e-Learning environment by selecting PHI 3681 from the Courses pull-down menu at the top of the page.
- If you encounter any difficulties logging in or accessing any of the course content, contact the UF Computing Help Desk at (352) 392-4537 or http://helpdesk.ufl.edu.
- Please do not contact the course instructor regarding computer issues (I am unlikely to be able to help you!).

COMMUNICATION POLICY

Announcements
Course announcements will be posted on Canvas. You are responsible for checking Canvas at least once a week to make sure that you do not miss important announcements.

Contacting Dr. Grant
For questions about course content, please feel free to reach out to Dr. Grant directly by email (david.grant@ufl.edu) or come to office hours—I'm here to help!

- Email is the most reliable way to get in touch with me outside of class.
- I make effort to respond to email from students within two (2) business days. Note that emails do sometimes get lost—due to spam filtering, for instance. Please do send me another email or come up to me after class if you do not hear back within two business days. I promise I am not ignoring you intentionally!

ASSIGNMENTS

Participation (20% / 200 points)
20% of your final grade will be determined by the quality of your participation in the course. Activities that will contribute to your participation grade include:

- annotating the weekly readings in Perusall (assignments will be announced on Canvas);
- regularly attending lecture;
- participating in small-group and full-class discussions during lecture; and
- submitting occasional small ad hoc assignments (announced in class or on Canvas).

Perusall annotations and ad hoc assignments will not be accepted late.

You will receive a letter grade for participation representing the overall quality of your participation in the above activities over the course of the semester. Below is a summary of course standards for different letter grades (borderline cases will receive "+" or "−" grades as appropriate):

- A: Outstanding participation in all categories. Highly engaged during lecture, with active participation in small group discussions and frequent constructive contributions to full-class discussions. Perusall annotations, ad hoc assignments, and in-class contributions reflect excellent understanding of and critical engagement with the content of lectures and assigned readings.
- B: Good participation in all categories. Moderately engaged during lecture, with active participation in small group discussions and occasional constructive contributions to full-class discussions. Perusall annotations, ad hoc assignments, and in-class contributions reflect significant understanding of and critical engagement with the content of lectures and assigned readings.
- C: Minimally adequate participation in all categories. Minimally engaged during lecture. Perusall annotations, ad hoc assignments, and in-class contributions reflect limited understanding of and critical engagement with the content of lectures and assigned readings.
- D or F: Inadequate participation in one or more categories.

Lecture attendance is required. Repeat unexcused absences will be subject to a penalty.

- You are allowed three unexcused absences without penalty.
- Each unexcused absence beyond the first three will result in a 10% deduction from your participation grade (and therefore 2% of your final grade in the course). For example, if your participation grade based on the above standards was an 86 ('B') and you had four unexcused absences, your final participation grade in the course would be a 76 ('C').
- If you need to miss lecture for any reason, please contact me as soon as possible so that we can discuss whether the absence will be excused and what sort of documentation will be necessary (if applicable). If you do not contact me about an absence within one week following the absence, it will not be excused except under extraordinary circumstances.
- Poor conduct in class will also result in a 10% deduction. Examples of poor conduct include arriving late, disruptive behavior, falling asleep, failing to participate in small group activities, sending text messages, checking email, etc.
- Students may be removed from the class roster for frequent unexcused absences.

Note that the "Roll Call Attendance" grade that Canvas automatically generates is not directly related to your Participation grade in the class.

Emerging technology case studies (40% / 400 points)
We will have four discussions of case studies concerning emerging technologies in class. Students will be assigned to discussion groups. Each group member will provide a written analysis of the case study and will use this analysis to inform small group discussion. Groups will then work together to submit a more polished written analysis, which will inform full-class discussion of the case study. Polished group analyses will be graded.

Midterm exam (20% / 200 points)
The midterm exam will be a take-home short essay exam. The questions will concern the readings and topics we have discussed over the first half of the semester. Collaboration on the midterm exam is not permitted; your answers should represent your own understanding of the course material.

Final exam (20% / 200 points)

Like the midterm exam, the final exam will be a take-home short essay exam. The questions will concern the readings and topics we have discussed over the second half of the semester. Collaboration on the final exam is not permitted; your answers should represent your own understanding of the course material.

GRADING SCALE

The following grade scale will be used to assign final letter grades for the course. See UF grading policies for assigning grade points at:
https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx.

Points	Grade Scale	Grade Value
934–1000	93.4-100 = A	A=4.0
900–933	90.0-93.3 = A-	A-=3.67
867–899	86.7-89.9 = B+	B+=3.33
834–866	83.4-86.6 = B	B=3.00
800–833	80.0-83.3 = B-	B-=2.67
767–799	76.7-79.9 =C+	C+=2.33
734–766	73.4-76.6 =C	C=2.00
700–733	70.0-73.3 = C-	C-=1.67
667–699	66.7-69.9 = D+	D+=1.33
634–666	63.4-66.6 = D	D=1.00
600–633	60.0-63.3 = D-	D-=0.67
0–599	0-59.9 = E	E=0.00

To receive a given grade, you must meet or exceed the lower threshold specified. See https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/ for more information about UF grading policies.

ADDITIONAL INFORMATION

COVID-19

In response to COVID-19, the following practices are in place to maintain your learning environment, to enhance the safety of our in-classroom interactions, and to further the health and safety of ourselves, our neighbors, and our loved ones.

- If you are not vaccinated, get vaccinated. Vaccines are readily available at no cost and have been demonstrated to be safe and effective against the COVID-19 virus. Visit this link for details on where to get your shot, including options that do not require an appointment: https://coronavirus.ufhealth.org/vaccinations/vaccine-availability/. Students who receive the first dose of the vaccine somewhere off-campus and/or outside of Gainesville can still receive their second dose on campus.
- You are expected to wear approved face coverings at all times during class and within buildings even if you are vaccinated. Please continue to follow healthy habits, including best practices like frequent hand washing. Following these practices is our responsibility as Gators.
  - Sanitizing supplies are available in the classroom if you wish to wipe down your desks prior to sitting down and at the end of the class.

- o Hand sanitizing stations will be located in every classroom.
- If you are sick, stay home and self-quarantine. Please visit the UF Health Screen, Test & Protect website about next steps, retake the questionnaire and schedule your test for no sooner than 24 hours after your symptoms began. Please call your primary care provider if you are ill and need immediate care or the UF Student Health Care Center at 352-392-1161 (or email covid@shcc.ufl.edu) to be evaluated for testing and to receive further instructions about returning to campus. UF Health Screen, Test & Protect offers guidance when you are sick, have been exposed to someone who has tested positive or have tested positive yourself. Visit the UF Health Screen, Test & Protect website for more information.
  - o Course materials will be provided to you with an excused absence, and you will be given a reasonable amount of time to make up work.
  - o If you are withheld from campus by the Department of Health through Screen, Test & Protect you are not permitted to use any on campus facilities. Students attempting to attend campus activities when withheld from campus will be referred to the Dean of Students Office.
- Continue to regularly visit coronavirus.UFHealth.org and coronavirus.ufl.edu for up-to-date information about COVID-19 and vaccination.

University Honesty Policy
UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (https://www.dso.ufl.edu/sccr/process/student-conducthonor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

Academic Dishonesty
All sources and assistance used in preparing your papers and presentations must be precisely and explicitly acknowledged. Please review the guidelines on attribution and plagiarism found here: https://guides.uflib.ufl.edu/copyright/plagiarism. The web creates special risks here. Cutting and pasting even a few words from a web page or paraphrasing material without a reference constitutes plagiarism. If you are not sure how to refer to something you find on the internet, you can always give the URL.

You should expect the minimum penalty for academic dishonesty to be a grade of E for the class (not just the assignment). All incidents of academic dishonesty will be reported to Student Judicial Affairs. Repeat offenders may be penalized by suspension or expulsion from the university.

Students with Disabilities
Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center. Click here to get started with the Disability Resource Center. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

Classroom Conduct
Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional

courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records.

Laptop and Cellphone Policy

Laptop and cellphone use is prohibited in this class. Please turn off your phone as soon as you enter the classroom. First-time violations will receive a verbal warning. Subsequent violations will yield a loss of participation points.

Attendance, Illness, Religious Holidays, and Twelve Day Rule

Requirements for class attendance, religious holidays, and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx.

Excerpt from the statement on absences:

"Absences count from the first class meeting. In general, acceptable reasons for absence from or failure to participate in class include illness, serious family emergencies, special curricular requirements (e.g., judging trips, field trips, professional conferences), military obligation, severe weather conditions, religious holidays, and participation in official university activities such as music performances, athletic competition or debate. Absences from class for court-imposed legal obligations (e.g., jury duty or subpoena) must be excused. Other reasons also may be approved."

Course evaluation process

"Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/."

Campus Resources

Matter, We Care: If you or someone you know is in distress, please contact umatter@ufl.edu, 352-392-1575, or visit U Matter, We Care website to refer or report a concern and a team member will reach out to the student in distress.

Counseling and Wellness Center: Visit the Counseling and Wellness Center website or call 352-392-1575 for information on crisis services as well as non-crisis services.

Student Health Care Center: Call 352-392-1161 for 24/7 information to help you find the care you need, or visit the Student Health Care Center website.

University Police Department: Visit UF Police Department website or call 352-392-1111 (or 9-1-1 for emergencies).

UF Health Shands Emergency Room / Trauma Center: For immediate medical care call 352-733-0111 or go to the emergency room at 1515 SW Archer Road, Gainesville, FL 32608; Visit the UF Health Emergency Room and Trauma Center website.

GatorWell Health Promotion Services: For prevention services focused on optimal wellbeing, including Wellness Coaching for Academic Success, visit the GatorWell website or call 352-273-4450.

Academic Resources

E-learning technical support: Contact the UF Computing Help Desk at 352-392-4357 or via e-mail at helpdesk@ufl.edu.

Career Connections Center: Reitz Union Suite 1300, 352-392-1601. Career assistance and counseling services.

Library Support: Various ways to receive assistance with respect to using the libraries or finding resources.

Teaching Center: Broward Hall, 352-392-2010 or to make an appointment 352- 392-6420. General study skills and tutoring.

Writing Studio: 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers.

Student Complaints On-Campus: Visit the Student Honor Code and Student Conduct Code webpage for more information.

On-Line Students Complaints: View the Distance Learning Student Complaint Process.

Tentative schedule and readings

Please note that the following is tentative and subject to change.
Please see Canvas for an up-to-date schedule and list of readings.

Week 1 (1/6) The alignment problem

- Christian, *The Alignment Problem*, chapter 1 ("Introduction")
- Lum and Chowdhury, "What is an 'Algorithm'? It Depends on Whom You Ask"

Week 2 (1/11, 1/13) Introduction to ethics

- *Shafer-Landau, "Consequentialism: Its Nature and Attractions"
- Wells et al., "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show"

Week 3 (1/18, 1/20) AI Safety

- *Nozick, "Moral Constraints and Moral Goals"
- Amodei et al., "Concrete Problems in AI Safety" (excerpts)
- Baraniuk, "How Google's Hot Air Balloon Surprised its Creators"

Individual responses to case study discussion 1 due at 11:59 PM on 1/16

Week 4 (1/25, 1/27) Privacy and surveillance capitalism

- *Rachels, "Why Privacy is Important"
- Pasquale, "Digital Reputation in an Era of Runaway Data"

Group responses to case study discussion 1 due at 11:59 PM on 1/26

Week 5 (2/1, 2/3) Privacy and surveillance capitalism

- *Barocas and Nissenbaum, "Big Data's End Run Around Anonymity and Consent" (excerpts)
- Marlinspike, "Why 'I Have Nothing to Hide' is the Wrong Way to Think About Surveillance"

Week 6 (2/8, 2/10) Autonomy and the attention economy

- *Aylsworth and Castro, "Is There a Duty to Be a Digital Minimalist?"
- Lewis, "'Our minds can be hijacked': the tech insiders who fear a smartphone dystopia" (*The Guardian*)

Individual responses to case study discussion 2 due at 11:59 PM on 2/6

Week 7 (2/15, 2/17) Fake news and the attention economy

- *Pepp et al., "What's New About Fake News?" (excerpts)
- *Mill, *On Liberty*, chapter 2 ("Of the Liberty of Thought and Discussion") (excerpts)
- Frenkel et al., "Surge of Virus Misinformation Stumps Facebook and Twitter" (*The New York Times*)

Week 8 (2/22, 2/24) Is the attention economy noxious?

- *Castro and Pham, "Is the Attention Economy Noxious?"
- Jason Pontin, "The Case for Less Speech" *Wired*

Week 9 (3/1, 3/3) Algorithmic opacity

- *Grant, Basl, and Behrends, "What We Owe to Decision Subjects: Beyond Transparency and Explanation in Automated Decision-Making"

Spring break (3/7-3/12)

Week 10 (3/15, 3/17) Algorithmic opacity

- *Vredenburgh, "The Right to Explanation"

Week 11 (3/22, 3/24) Algorithmic bias

- *Johnson, "Algorithmic Bias: on the Implicit Biases of Social Technology"
- Latanya Sweeney, "Discrimination in Online Ad Delivery"

Week 12 (3/29, 3/31) Algorithmic bias

- *Fazelpour and Danks, "Algorithmic Bias: Senses, Sources, Solutions"
- Angwin et al., "Machine Bias" (*ProPublica*)
- Corbett-Davies et al., "A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear." (*Washington Post*)

Week 13 (4/5, 4/7) Algorithmic bias

- *Mayson, "Bias In, Bias Out" (excerpts)

Week 14 (4/12, 4/14) Responsibility gaps

- *Sparrow, "Killer Robots"

Week 15 (4/19) Conclusion

- No new readings

Final exam due by 11:59 PM on 4/25