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

TR 9:35 AM - 10:25 AM
Little Hall 237

This course has a required discussion section. Your discussion section time and location is determined by your registration number for the course.

TEACHING TEAM

INSTRUCTOR

David Gray Grant
Assistant Professor of Philosophy
University of Florida
david.grant@ufl.edu
Office hours: TR 10:30-11:30 AM and by appointment
Office location: TBA

TEACHING ASSISTANT

Shadi Heidarifar
Doctoral Student in Philosophy
University of Florida
sheidarifar@ufl.edu
Office hours: TBA

REQUIRED MATERIALS

All readings will be posted on Canvas (see below).

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 or your TA regarding computer issues (we are 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 your TA
Your teaching assistant is your primary point of contact for questions about assignments, grades, attendance, and administrative matters for the course. Your TA will let you know the best way to contact them with questions.

- Per the above, please do not contact your TA with questions about tech support for Canvas or other UF systems. (But please do let your TA know if technical issues are interfering with your ability to participate in the course.)

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 by far the most reliable way to get in touch with me outside of class. (Much more reliable than Canvas, for instance.)
- 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

Short reading quizzes (10%)

There will be irregular reading quizzes, designed to make sure that you are doing and understanding the readings.

Emerging technology case studies (50%)

We will have several discussions of case studies concerning emerging technologies in class or in the required discussion sections. 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 the discussion. Each group will submit the "best" written analysis at the end of the discussion. Grades for the case studies will be determined by (a) participation in the case study discussion and (b) the quality of the group's submitted written analysis.

Short Paper (20%)

You will write one short paper (1200-1500 words) in which you will critically assess a topic that we have discussed in class. I will give you a choice of two topics for this paper; I will also provide the structure that the paper must follow. One of the topics will be theoretical, addressing an argument for a particular position in ethical theory. The other topic will involve *applying* ethical concepts in order to assess a specific technology.

Attendance and participation (20%)

Philosophy is a team sport, and you will perform best in this course if you participate actively in our in-class discussions. Regular attendance and participation in class and discussion section makes up the remaining 20% of your grade.

- Attendance in lecture and section is mandatory. Students may be removed from the class roster for frequent unexcused absences.
- You are allowed two unexcused absences without penalty. If you need to miss lecture or section for any reason, please contact your TA as soon as possible so that you can discuss whether the absence will be excused and what sort of documentation will be necessary (if applicable).

- Each unexcused absence beyond the first two will result in a 10% deduction from your participation grade (and therefore 2% of your final grade in the course).
- Poor conduct in class will also result in a 10% deduction. Examples of poor conduct include disruptive behavior, falling asleep, sending text messages, checking email, etc.

> **IF YOU SUSPECT YOU ARE SICK, PLEASE DO NOT COME TO CLASS (SEE UNIVERSITY GUIDELINES BELOW)**

- If you are feeling ill, please do not come to class – even if you are vaccinated against COVID-19. The delta variant of COVID-19 can be spread by vaccinated individuals, and spreads readily to the unvaccinated. Please keep in mind that many of us have unvaccinated family members (such as small children) that are vulnerable and need to be protected.
- **Students will not be penalized for staying home due to illness.** If you do feel ill, please let me know and **I will work with you directly to help you get up to speed** when you are feeling well enough to return.

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.

Grade Scale	Grade Value
100-93=A	A=4.0
92-90=A-	A-=3.67
89-86=B+	B+=3.33
85-82=B	B=3.00
81-79=B-	B-=2.67
78-76=C+	C+=2.33
75-72=C	C=2.00
71-69=C-	C-=1.67
68-66=D+	D+=1.33
65-62=D	D=1.00
61-60=D-	D-=0.67
59-0=E	E=0.00

Grades that fall exactly on the upper threshold are awarded the higher grade. See https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx 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.
    - 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.
    - 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.
    - 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.

Academic Honesty

Please review the following guidelines on academic honesty:

1. http://www.dso.ufl.edu/studentguide/studentrights.php
2. http://www.registrar.ufl.edu/catalog/policies/students.html#honesty

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.

All sources and assistance used in preparing your papers and presentations must be precisely and explicitly acknowledged. 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.

Students with Disabilities

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure 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/."

Counseling and wellness/Emergencies

http://www.counseling.ufl.edu/cwc/Default.aspx, 392-1575;
The University Police Department: 392-1111 or 9-1-1 for emergencies.

Topics and readings

The alignment problem

T 8/24	Introduction to the course	• No readings
R 8/26	The alignment problem	• Christian, The Alignment Problem, ch 1 ("Introduction") • Hao, "What is Machine Learning?" (MIT Technology Review) • McCarthy, "What is AI?" (http://jmc.stanford.edu/artificial-intelligence/what-is-ai/)
T 8/31	Do artifacts have politics?	• Langdon Winner, "Do Artifacts Have Politics?" • Sweeney, "Saving Humanity" (keynote address) (https://www.youtube.com/watch?v=OlK_nVOM2tc) (talk starts just before the five minute mark; watch until Q&A begins)
R 9/2	AI safety	• Amodei et al, "Concrete Problems in AI Safety"

Doing good with data

| T 9/7 | AI for social good | • Chui et al, "Applying artificial intelligence for social good" (McKinsey Global Institute) (https://www.mckinsey.com/featured-insights/artificial-intelligence/applying-artificial-intelligence-for-social-good)
• Latonero, "Opinion: AI for Good is Often Bad" |

- Floridi et al., "AI4People—An Ethical Framework for a Good AI Society"

R 9/9	Interpretability	• Christian, *The Alignment Problem*, ch 4 ("Transparency")
T 9/14	Interpretability	• London, "Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability" • Holm, "In Defense of the Black Box" (*Science*)
R 9/16	Data skepticism	• Fry, "What Data Can't Do" • O'Neil, "On Being a Data Skeptic"
T 9/21	Consequentialism	• Shafer-Landau, "Consequentialism: Its Nature and Attractions"
R 9/23	Moral constraints	• Anscombe, "Mr. Truman's degree": • Nozick, "Moral Constraints and Moral Goals"
T 9/28	Problems for consequentialism	• Shafer Landau, "Consequentialism: Its Difficulties"
R 9/30	TND Network	• USC Center for Artificial Intelligence in Society, "Social Networks and Substance Abuse Prevention For Homeless Youth" (webpage with videos) (https://cais.usc.edu/projects/social-network-based-substance-abuse-prevention-for-homeless-youth)
T 10/5	Responsibility gaps	• Sparrow, "Killer Robots"
R 10/7	Responsibilities of scientists and engineers	• "J. Robert Oppenheimer, Atom Bomb Pioneer, Dies" (*NYT*) • Heaven, "OpenAI's new language generator GPT-3 is shockingly good—and completely mindless" (*MIT Technology Review*) (https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/) • Hao, "We read the paper that forced Timnit Gebru out of Google. Here's what it says." (MIT Technology Review) (https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/) • Bender et al., "On the Dangers of Stochastic Parrots"

Autonomy and privacy

T 10/12	Autonomy and the attention economy	• Tolentino, "How TikTok Holds Our Attention" (*The New Yorker*) • Kukla, "Conscientious Autonomy: Displacing Decisions in Health Care"
R 10/14	Autonomy and the attention economy	• Tolentinto, "What it Takes to Put Your Phone Away" (*The New Yorker*) • Aylsworth and Castro, "Is There a Duty to Be a Digital Minimalist?"
T 10/19	Fake news and the "infodemic"	• Frenkel et al., "Surge of Virus Misinformation Stumps Facebook and Twitter" (*The New York Times*) • David Lazer et al., "The Science of Fake News" (*Science*) • Michaelson et al., "What's New About Fake News?"
R 10/21	Fake news and free speech	• Jason Pontin, "The Case for Less Speech" *Wired* • Mill, *On Liberty*, ch 2 ("Of the Liberty of Thought and Discussion")
T 10/26	Surveillance capitalism	• Frank Pasquale, "Digital Reputation in an Era of Runaway Data" • Satariano, "What a Gambling App Knows About You" (*New York Times*) • Marlinspike, "Why 'I Have Nothing to Hide' is the Wrong Way to Think About Surveillance"
R 10/28	Why is privacy important?	• Rachels, "Why Privacy is Important"
T 11/2	Privacy and the pandemic	• Kharpal, "Use of surveillance to fight coronavirus raises concerns about government power after pandemic ends" (*CNBC*) (https://www.cnbc.com/2020/03/27/coronavirus-surveillance-used-by-governments-to-fight-pandemic-privacy-concerns.html) • Ienca and Vayena, "On the responsible use of digital data to tackle the COVID-19 pandemic" (*Nature Medicine*) (https://www.nature.com/articles/s41591-020-0832-5) • "Creating the Coronopticon" (*The Economist*) • Cozzens, "Pandemic drones to monitor, detect those with COVID-19" (*GPS World*)

(https://www.gpsworld.com/draganfly-camera-and-uav-expertise-to-help-diagnose-coronavirus/)

Fairness and bias

R 11/4	Treating people as individuals	• Eidelson, "Treating People as Individuals"
T 11/9	Treating people as individuals	• Lippert-Rasmussen, "'We are all Different': Statistical Discrimination and the Right to be Treated as an Individual
R 11/11	No class (Veteran's Day)	• No readings
T 11/16	Algorithmic bias	• Christian, *The Alignment Problem*, chs 2 ("Representation") and 3 ("Fairness")
R 11/18	Algorithmic bias	• Johnson, "Algorithmic Bias: on the Implicit Biases of Social Technology"
T 11/23	Algorithmic bias	• Fazelpour and Danks, "Algorithmic Bias: Senses, Sources, Solutions" • Obermeyer, "Dissecting Racial Bias in Health Algorithms" (video lecture) (https://www.youtube.com/watch?v=WDXaOtSQF98)
R 11/25	No class (Thanksgiving break)	• No readings
T 11/30	COMPAS	• Courtland, "Bias Detectives" (*Scientific American*) • 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*)
R 12/2	COMPAS	• Mayson, "Bias In, Bias Out"
T 12/7	Conclusion	• No readings