

**Ethics, Data, and Technology (PHI 3681)**  
*Fall 2024*

**Lecture:** Tuesday and Thursday | Period 3 (9:35 am – 10:25 am) in MCCA G186  
Discussion 28503 (section 3630): Friday Period 5 (11:45 am – 12:35 pm) in MAT 0005  
Discussion 28504 (section 3631): Friday Period 6 (12:50 pm – 1:40 pm) in WEIM 1070  
Discussion 28499 (section 3626): Friday Period 3 (9:35 am – 10:25 am) in MAT 0016  
Discussion 28500 (section 3627): Friday Period 5 (11:45 am – 12: 35 pm) in TUR 2354  
Discussion 28501 (section 3628): Friday Period 6 (12:50 pm – 1:40 pm) in FAC 0120  
Discussion 28502 (section 3629): Friday Period 3 (9:35 am – 10:25 am) in MAT 0005

**Instructor:** Duncan Purves

Contact: [dpurves@ufl.edu](mailto:dpurves@ufl.edu)  
Office hours: Thursday, 11:30 am – 1:00 pm in Griffin-Floyd Hall, rm. 332 and by appointment

**Teaching Assistants:**

Julianna Costanzo

Contact: [jsc1998@ufl.edu](mailto:jsc1998@ufl.edu)  
Office hours: 11 am – 2 pm in Griffin-Floyd Hall, rm. 316.

Jack Madock

Contact: [johnmadock@ufl.edu](mailto:johnmadock@ufl.edu)  
Office hours: TBD in Griffin-Floyd Hall, rm. 316.

---

**COURSE DESCRIPTION**

This course examines the ethical dimensions of contemporary machine learning and its applications. Here “ethics” refers to an academic discipline that investigates questions like: What actions are right or wrong; What values should guide my actions; What things are good and bad; What kind of person should I be? Machine learning is a method by which computer and data scientists train computer systems to perform tasks such as recognizing objects and patterns, reasoning, strategizing, and physically navigating the world. In this course, students will come to understand how ethics and machine learning collide. Through real-world examples and case studies, students will learn how to explain how four features of machine learning applications give rise to ethical questions: bias, scale, opacity, and autonomy. Applications of machine learning discussed in the course include criminal risk assessment, targeted advertising, lethal autonomous weapon systems, and generative artificial intelligence. Students will leave the class with a more holistic, ethically sensitive, approach to evaluating novel machine learning technologies.

---

**COURSE OBJECTIVES**

1. Identify and explain the philosophical dimensions of foundational concepts in data science and machine learning including bias, fairness, opacity, privacy, and consent.
  2. Develop a basic vocabulary for discussing the ethical dimensions of data science and big data technologies.
  3. Evaluate contemporary uses of data-driven systems through the application of ethical concepts.
  4. Discern the structure of arguments, representing them fairly and clearly, and evaluating them for cogency.
  6. Formulate original arguments, anticipate objections, and respond in a conscientious fashion.
  7. Read and discuss complex texts from contemporary work in philosophy, law, and information and data science.
  8. Speak and write clearly and persuasively about abstract and conceptually elusive matters.
- 

### GRADING POLICIES

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/UGRD/academic-regulations/grades-grading-policies/>

Grade Scale	Grade Value
100-92.5=A	A=4.0
92.5-89.5=A-	A-=3.67
89.5-85.5=B+	B+=3.33
85.5-81.5=B	B=3.00
81.5-78.5=B-	B-=2.67
78.5-75.5=C+	C+=2.33
75.5-71.5=C	C=2.00
71.5-68.5=C-	C-=1.67
68.5-65.5=D+	D+=1.33
65.5-61.5=D	D=1.00
61.5-59.5=D-	D-=0.67
59.5-0=E	E=0.00

Grades that fall exactly on the threshold are awarded the higher grade.

---

## COURSE WEBSITE AND TECHNICAL ISSUES

This course is supplemented by online content in the e-Learning environment known as "Canvas." 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>. Do not contact the course instructor regarding computer issues (I am unlikely to be of any help!).

---

## REQUIRED READINGS

The primary reading for this course is a book manuscript by David Boonin titled *Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords*. It will be supplemented by a variety of academic articles as well as news articles, documentaries, and other media.

This manuscript and all other readings and media can be found on Canvas either in the 'Files' tab or linked from the Course Schedule below.

---

## COURSE REQUIREMENTS (see due dates on Canvas and the course schedule below)

**Weekly discussion comments (15%)**: Each week, students will submit substantive critical comments and questions on the readings or lectures through Canvas. You can also engage with your peers' comments to increase your depth of engagement with course material. Credit for these comments is all-or-nothing. To receive credit the comment must be "substantive," as characterized below:

- A substantive comment is one that provides something in addition to the initial post or previous comments.
- Substantive comments can be critical of the points made in the reading or they can be questions of clarification, either about the reading or about a comment made by another student on the discussion thread.
- Posting a question that prompts clarification or leads the discussion deeper into the topic would be considered substantive.
- A post that relates a concept from the reading to one's personal experience in an illuminating way would be considered substantive.
- Posts that only say "I agree" or "this is an interesting point" are not substantive.

**Two Short Papers (40%)**: You will write two short papers (900 – 1200 words each) each worth 20% of your final grade. This short paper involves a critical analysis of a position or argument concerning a topic that we have discussed in class. I will give you a choice of topics for each paper, I will provide the structure that the paper must follow, and I will provide you with a detailed grading rubric.

These papers assess whether students have acquired:

- skill in discerning the structure of arguments, representing them fairly and clearly, and evaluating them for cogency.
- skill in formulating original arguments, anticipating objections, and responding in a conscientious fashion.
- skill in reading and discussing complex philosophical texts from both historical sources and contemporary works
- and skill in speaking and writing clearly and persuasively about abstract and conceptually elusive matters.

**Case studies (30%):** Students will complete two case studies. These case studies ask students to synthesize and apply ethical concepts from the unit to a novel application of machine learning or AI. See the assignment description on Canvas for more details.

**Curating your social media diet: a blog (15%):** Students will select one aspect of their social media consumption that they feel needs adjustment or improvement. This could include reducing overall time spent on platforms, curating a more positive feed, limiting usage during specific hours, or even increasing your social media consumption in some way, if that's something you've been wanting to do. Throughout the duration of the course unit on autonomy, Students will document their experience in a weekly blog post that explores the reasons behind the chosen change. They will reflect on the challenges they faced, insights gained, and the impact of this adjustment on their daily life and well-being. More details and assignment rubric are available on Canvas.

---

## ADDITIONAL INFORMATION

### *Academic Honesty*

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 Conduct Code specifies a number of behaviors that are in violation of this code and the possible sanctions. [Click here to read the Conduct Code](#). If you have any questions or concerns, please consult with the instructor or TAs in this class.

### *AI Use Policy*

This policy covers any generative AI tool, such as ChatGtP, Elicit, etc. This includes text and artwork/graphics/video/audio. **1.** Students are not permitted AI tools UNLESS under direct instruction from the instructor to do so. Please contact your instructor if you are unsure or have questions BEFORE using AI for any assignment. **2.** If AI is permitted to be used, students must indicate what part of the assignment was written by AI and what was written by the student. **3.** The first assignment in the course is to sign the AI contract stating that you understand and agree

to these policies. Violations of this contract will be understood as violations of the pledge described in the above section titled “Academic Honesty.”

### *Students with Disabilities*

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, <https://disability.ufl.edu/get-started/>). 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.

### *Late work, 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>.

\*\*If you think you have an acceptable reason to be excused for late work, you must contact the [Dean of Students Office](#) to receive an excuse letter. This prevents me from having to make difficult discriminations between excuses, and it provides you with something you can use for all of your missed classes.

### *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/>.

### *Email*

The best way to reach me or your TAs is by email (see page 1 of the syllabus). Although we typically check email numerous times per day, you shouldn’t expect us to reply to an email before the next business day. Please plan ahead if you have questions before major assignments are due. We are happy to answer whatever questions you have over email, so long as they are the kinds of questions that can be answered in just a couple of sentences.

We prefer that you ask more involved questions—especially conceptual questions about the lectures or readings—during office hours.

*\*I will occasionally email the class list with important information, like changes in the reading, discussion questions, or modifications for assignments. Therefore, please check your email regularly.*

### Campus Resources:

#### Health and Wellness

**U Matter, We Care:** If you or someone you know is in distress, please contact [umatter@ufl.edu](mailto: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](mailto: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.

## COURSE SCHEDULE

<i>Date</i>	<i>Topic</i>	<i>Reading</i>
<b>Week 1</b> Thurs., 8/22		Introductions/Syllabus/Discussion of ProPublica’s Machine Bias
<b>Week 2</b> Tues., 8/27	Technology and Society: The Ethics of Artificial Intelligence and Machine Learning	Langdon Winner, “Technologies as Forms of Life”
Thurs., 8/29	Ethics, Machine Learning, and AI	David Boonin (Forthcoming). <i>Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords</i> , Chapter 1: The Power of the State Meets the Power of Machine Learning
<b>Unit 1: Bias and Machine Learning</b>		
<b>Week 3</b> Tues., 9/3	Bias in Machine Learning: Sources and Measures	Julia Angwin et al, <a href="#">“Machine Bias,” ProPublica</a>  Sam Corbett-Davies et al (2016). <a href="#">“A computer algorithm used for bail and sentencing decisions was labeled biased against blacks. It’s actually not that clear,”</a> Corbett-Davies et al, Washington Post
Thurs., 9/5	Bias in Machine Learning: Sources and Measures	David Boonin (Forthcoming). <i>Artificial Intelligence, Ethics and the State</i> , Chapter 2: Predictive Policing and the Problem of Bias (Section 2.0 - 2.2.3)
<b>Week 4</b> Tues., 9/10	Is Bias in Machine Learning a Form of Racial Discrimination?	David Boonin (Forthcoming). <i>Artificial Intelligence, Ethics and the State</i> , Chapter 2: Predictive Policing and the Problem of Bias (Section 2.0 - 2.3.2.1)  Altman, “Discrimination,” sections 1–2 (Stanford Encyclopedia of Philosophy)
Thurs., 9/12	What’s Wrong with Bias in Machine Learning?	Selections from Clinton Castro (2019). “What’s wrong with machine bias?” <i>Ergo</i> 6(15), pp. 415 – 424.
<b>Week 5</b> Tues., 9/17	“Using Race” to Improve Fairness in Machine Learning?	Hellman, “Measuring Algorithmic Fairness” <i>Virginia Law Review</i> 106(4) (846 – 864)
<b>Unit 2: Machine Learning and Scale</b>		
Thurs., 9/19	Privacy in the Age of Big Data and Machine Learning	Amnesty International, “Surveillance Giants,” pages 1–26: <a href="https://www.amnesty.org/en/documents/pol30/1404/2019/en/">https://www.amnesty.org/en/documents/pol30/1404/2019/en/</a>



		James Rachels (1975). "Why Privacy is Important" <i>Philosophy &amp; Public Affairs</i> , 4(4), 323–333. <a href="http://www.jstor.org/stable/2265077">http://www.jstor.org/stable/2265077</a>
<b>Week 6</b> Tues., 9/24	<b>In-class Case Study Activity:</b> <b>Machine Bias in Predictive Policing</b>	
Thurs., 9/26	Why Does Privacy Matter?	Moxie Marlinspike, "Why 'I Have Nothing to Hide' is the Wrong Way to Think About Surveillance" <i>Wired</i> <a href="https://www.wired.com/2013/06/why-i-have-nothing-to-hide-is-the-wrong-way-to-think-about-surveillance/">https://www.wired.com/2013/06/why-i-have-nothing-to-hide-is-the-wrong-way-to-think-about-surveillance/</a>
<b>Week 7</b> Tues., 10/1	Against Privacy	Tony Doyle, "Privacy and Perfect Voyeurism" <b>Guest Lecture from Teaching Assistant</b>
Thurs., 10/3	Protecting Privacy in the Age of Big Data	Barocas and Nissenbaum, "Big Data's End Run Around Anonymity and Consent," in <i>Privacy, Big Data, and the Public Good: Frameworks for Engagement</i> , Edited by Julia Lane, Victoria Stodden, Stefan Bender, Helen Nissenbaum. Book DOI: <a href="http://dx.doi.org/10.1017/CBO9781107590205">http://dx.doi.org/10.1017/CBO9781107590205</a> <a href="https://nissenbaum.tech.cornell.edu/papers/BigDataEndRun.pdf">https://nissenbaum.tech.cornell.edu/papers/BigDataEndRun.pdf</a>
<b>Week 8</b> Tues., 10/8	<b>First Case Study Due</b> Protecting Privacy: From Anonymity to Noise	Kearns and Roth (2020). "Algorithmic Privacy: From Anonymity to Noise," chapter 2 of <i>The Ethical Algorithm</i> . Oxford University Press. pp. 22 - 56.
<b>Unit 3: Opacity</b>		
Thurs., 10/10	Machine Learning, Transparency, and the Right to an Explanation	David Boonin (Forthcoming). <i>Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords</i> , Chapter 4: Risk Assessment Tools and the Problem of Opacity (Section 4.0 - 4.2.3)  Jenna Burrell. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. <i>Big Data &amp; Society</i> , 3(1). <a href="https://doi.org/10.1177/2053951715622512">https://doi.org/10.1177/2053951715622512</a>
<b>Week 9</b> Tues., 10/15	Opaque Juries and the Right to an Explanation	David Boonin (Forthcoming). <i>Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords</i> , Chapter 4: Risk Assessment Tools and the Problem of Opacity (Section 4.3 - 4.3.3)
Thurs., 10/17	Explanation Opaque Juries and the Right to an Explanation cont'd	No New Reading

		<b>Unit 4: Autonomy</b>
<b>Week 10</b> Tues., 10/22	Social Media, Misinformation, and the Ethics of Censorship  Begin Social Media Diet Blog	<p>Sheera Frenkel, Davey Alba and Raymond Zhong (2020). "Surge of Virus Misinformation Stumps Facebook and Twitter" (The New York Times) <a href="https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html">https://www.nytimes.com/2020/03/08/technology/coronavirus-misinformation-social-media.html</a></p> <p>Jason Pontin, (2018) "The Case for Less Speech" (Wired) <a href="https://www.wired.com/story/ideas-jason-pontin-less-speech/">https://www.wired.com/story/ideas-jason-pontin-less-speech/</a></p> <p>John Stuart Mill, On Liberty, ch. 2 ("Of the Liberty of Thought and Discussion") (excerpts)</p>
Thurs., 10/24	Social Media, Misinformation, and the Ethics of Censorship	<p>Georgia Wells, Jeff Horwitz, and Deepa Seetharaman (2021). "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show" WSJ (<a href="https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739">https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739</a>)</p> <p>Castro, Clinton &amp; Pham, Adam (2020). "Is the Attention Economy Noxious?" Philosophers' Imprint 20 (17):1-13. <a href="https://quod.lib.umich.edu/cgi/p/pod/dod-idx/is-the-attention-economy-noxious.pdf?c=phimp;idno=3521354.0020.017;format=pdf">https://quod.lib.umich.edu/cgi/p/pod/dod-idx/is-the-attention-economy-noxious.pdf?c=phimp;idno=3521354.0020.017;format=pdf</a></p>
<b>Week 11</b> Tues., 10/29	<b>Paper #1 Due at 11:59 pm</b>	
	Catch-up Day	No new reading
Thurs., 10/31	<b>In-class Case study Activity: Censorship, the Attention Economy, and Rational Autonomy</b>	
<b>Week 12</b> Tues., 11/5	Lethal Autonomous Weapon Systems- Arguments for and Against	"David Boonin (Forthcoming). Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords, Chapter 5: Autonomous Weapon Systems and the Problem of Responsibility (Section 5.0 - 5.1.3)

		Jonah M. Kessel, Natalie Reneau and Melissa Chan (2019) "Killing in the Age of Algorithms" New York Times <a href="https://www.nytimes.com/video/technology/100000006082083/lethal-autonomous-weapons.html">https://www.nytimes.com/video/technology/100000006082083/lethal-autonomous-weapons.html</a>
Thurs., 11/7	Lethal Autonomous Weapon Systems and Moral Responsibility	"David Boonin (Forthcoming). Artificial Intelligence, Ethics, and the State: Two Cheers for Our New Robot Overlords, Chapter 5: Autonomous Weapon Systems and the Problem of Responsibility (Section 5.2 - 5.3.2.1)
<b>Week 13</b> Tues., 11/12	<b>Second Case Study Due at 11:59 pm</b>  Lethal Autonomous Weapon Systems and Moral Responsibility Cont'd	No New Reading
Thurs., 11/14	Generative AI and Intellectual Property	Trystan S. Goetze (2024). "AI Art is Stealing from Artists: Labour, Extraction, and Exploitation Or, On the Dangers of Stochastic Pollocks" <a href="https://arxiv.org/abs/2401.06178">https://arxiv.org/abs/2401.06178</a>  Kyle Chayka, "Is AI Art Stealing from Artists?" <i>The New Yorker</i> <a href="https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists">https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists</a>  Yiwen Lu (2024). Digital Media Outlets Sue OpenAI for Copyright Infringement. NYT. <a href="https://www.nytimes.com/2024/02/28/technology/openai-copyright-suit-media.html?campaign_id=4">https://www.nytimes.com/2024/02/28/technology/openai-copyright-suit-media.html?campaign_id=4</a>
<b>Week 14</b> Tues., 11/19	Generative AI and the Future of Work	John Danaher (2017). "Will Life Be Worth Living in a World Without Work? Technological Unemployment and the Meaning of Life." <i>Sci Eng Ethics</i> 23, 41–64.  New York Times, "China is Testing More Driverless Cars than Any Other Country" <a href="https://www.nytimes.com/2024/06/13/business/china-driverless-cars.html">https://www.nytimes.com/2024/06/13/business/china-driverless-cars.html</a>
Thurs., 11/21	TBD <b>Social Media Diet Blog Due at 11:59 pm</b>	
<b>Week 15</b> Tues., 11/26		<b>Thanksgiving No Class</b>

<i>Thurs., 11/28</i>		Thanksgiving No Class
<b>Week 16</b> <i>Tues., 12/3</i>	TBD Paper #2 Due at 11:59 pm	