

COURSE SYLLABUS

YEAR COURSE OFFERED: 2025

SEMESTER COURSE OFFERED: Fall

DEPARTMENT: Philosophy

COURSE NUMBER: 3681, 3 credit hours

NAME OF COURSE: Ethics, Data, and Technology

NAME OF INSTRUCTOR: Cameron Buckner

The information contained in this class syllabus is subject to change. Students are expected to be aware of any additional course policies presented by the instructor during the course.

Time and location: T/R Period 4 10:40-11:30, LIT 0101

Office hours: T/R 9:30-10:30 FLO 330B

Instructor e-mail: cameron.buckner@ufl.edu

TAs:

Jacob Calhoun calhounc@ufl.edu Office Hours: T 11:45-2:45 FLO 316

Panagiotis Saranteas saranteas.p@ufl.edu Office Hours: T/R 9:30-10:30am
F 2:00-3:00pm, FLO 318

Philosophy office phone: (352) 392-2084

Materials and Supplies Fee: N/A

Discussion sections:

18876: F Period 5 11:45-12:35 AND 0101 Panagiotis

18877: F Period 6 12:50 - 1:40 MAT 0016 Panagiotis

18878: F Period 3 9:35-10:25 MAT 0107 Panagiotis

19740: F Period 3 9:35 - 10:25 MAT 0014 Calhoun

20492: F Period 5 11:45-12:35 MAT 0119 Calhoun

20493: F Period 6 12:50 - 1:40 MAT 0119 Calhoun

Course Description:

This course will explore philosophical issues surrounding the development and deployment of emerging technologies, focusing especially on technological advances based on “deep learning” techniques in computer science. The primary focus will be on ethical and explanatory questions surrounding the use of these systems, which in just a few years have come to have pervasive effects in our daily lives—despite the fact that our understanding of their philosophical implications remaining rudimentary. Questions we will explore are: in what senses are these systems biased, and when is their bias ethically problematic? Can we explain the workings of these vastly complex systems—containing billions of parameters and trained on Internet-scale datasets—in a way that answers to our existing scientific, legal, and ethical practices? Who is responsible when these systems

COURSE SYLLABUS

err? And finally: how can humanity adjust to the radical changes these systems are bringing to our social, political, and economic lives without losing our fundamental humanity, and can these systems be designed in a way to align with human values—as opposed to alien, machine objectives?

Please note that all lecture notes and reading responses for this course are available on the course's Canvas site.

Course objectives:

1. Understand the basic architectures and training methods used in contemporary “deep” machine learning research, the differences between this approach and earlier methods in artificial intelligence, and their current applications to software systems in daily life.
2. Develop a basic vocabulary for understanding and ethically evaluating these systems by drawing upon theories and concepts from philosophy.
3. Recognize arguments for various positions in the ethical evaluation of cutting-edge technologies, represent them fairly and clearly, and evaluate their cogency.
4. Formulate original arguments, anticipate objections, and respond to them in a conscientious manner.
5. Read and discuss complex texts from historical sources and contemporary works.
6. Speak and write persuasively on abstract and conceptually difficult issues at the intersection of philosophy and emerging technologies.

Major Assignments:

Weekly Reading Responses	10%
Mid-Term Exam	20%
Group Case Studies (5)	50%
Final Exam	20%

Note: Final grades will be calculated according to the weighting above. Please disregard any automatically calculated grade categories on the Canvas gradebook. I try to hide all these by default to avoid confusion, but settings change frequently and Canvas will sometimes include incorrectly tabulated grade averages. If you want to know what your current average is, ask me by e-mail or use an offline grade calculator with those weightings and your actual assignment grades.

Weekly Reading Responses

Weekly reading response assignments will be posted on Canvas. These will provide a “guided tour” of the readings, helping you navigate philosophical texts by recognizing key claims and arguments. You should complete the reading responses as you go through the reading, by answering true/false, multiple choice, and fill-in-the-blank style questions.

Group Case Studies

In addition to reviewing the topics covered in the main lecture component, most of the time in discussion sections will be spent working on group case studies on the topics of interest. For each case study, you will work on an activity in groups of 4-6, answering questions provided in a prompt. In most cases you will need to interact with an AI tool to complete the project. You will present your group work to the rest of the class in discussion sections in a 6-8 minute presentation and

COURSE SYLLABUS

submit a shared 2-4 page “issue brief” reviewing the work you did as a group. The in-class presentation must have slides with visual aids to present on the projector—do not just read from your phone or a paper. The issue brief should also contain appendices (not included in the page count) with full transcripts or screenshots of any interactions you had with AI models. Groups are free to divvy up the work as they see fit—some group members might do more writing, others more oral presentation, for example—but everyone should pull their weight.

In addition to your own group presentation, each group will be assigned as a “discussant” for another group. This will be like you were assigned as a commentator on their presentation at a conference. On presentation day, your whole group needs to be present in class for the presentation for which you are assigned as a discussant. All group members should attend to the presentation and be prepared to ask questions in the Q&A. You can ask clarification questions, suggest opportunities for improvement, or future directions to pursue on the same topic. Please remember to be civil and charitable in your discussant questions; your goal is not to ask “gotcha questions” or offer devastating objections, but rather help the whole class see the strengths of the presentation and critically evaluate it.

Each individual group member will turn in their own 250-500 word “discussant appraisal” that is evaluated individually by your TA, and this grade will be rolled into your overall project grade for that unit (25% of grade—you are evaluated for the discussant appraisals you write, not the ones you receive). In the discussant appraisal, you should summarize the topic and position of the group for which you were a discussant, and appraise their take on the issue. You can agree with them, disagree with them, or offer an alternative set of options on the topic. You will be individually evaluated on your ability to clearly and charitably summarize the position and work done by the target group, and the quality of the reasoning you used to justify your appraisal.

Grades will be derived from rubrics based on both the presentation and issue brief. 75% of the grade will be shared amongst the whole group for the project you present, but 25% of each case study grade will be based on your discussant appraisal of another project. Because they need to be present for their discussant appraisal, all group members are expected to attend on presentation day even if not presenting.

Students can also opt out of the group work entirely by doing individual research and writing their own longer issue brief, more like a 3-4 page individual paper. If choosing the individual option, you must also come explain your paper individually and answer questions about it with the TA in office hours. Where possible, I highly encourage all students to choose the group work option, which is likely to be more efficient for everyone and to produce more interesting research.

Exams

Class members will take two in-class essay exams: one mid-term and one final. These will be completed in Blue Books by hand in class. A list of possible topics will be handed out roughly 1 week before the exam date, from which the essay questions will be chosen.

Grading scale

Grading of each assignment will be based on a 100 point scale. Grades for group projects will be derived from the grading template rubrics, which you can see and consider ahead of time. Grade

COURSE SYLLABUS

percentages below .5 will be rounded down and .5 or above will be rounded up. For final letter grades, I reserve the discretion to further round up grades near a cutoff line in final averages for students who have reliably turned in assignments and/or improved their performance towards the end of the course.

Percentage greater than or equal to	Letter
93-100 %	A
90-92 %	A-
87-89 %	B+
83-86 %	B
80-82 %	B-
77-79 %	C+
73-76 %	C
70-72 %	C-
67-69 %	D+
63-66 %	D
60-62 %	D-
< 60 %	F

Required Reading

All readings will be posted on the Canvas site.

List of discussion/lecture topics

Students are expected to have read the chapters or articles indicated and to have completed the weekly reading response on Canvas by the assigned day.

Week	Topic	Readings
Week 1 Aug 21	Basic Machine Learning #1: AI & DCNNs	Turing 1950: Computing Machinery and Intelligence Buckner 2018: Deep Learning: A Philosophical Introduction <i>Discussion Section Group Activity #1 Assigned:</i> Case Study on the Turing Test
Week 2 Aug 26	Generative AI and Philosophy of Mind	Buckner 2018: Empiricism without Magic-Transformational Abstraction in DCNNs
Week 3 Sep 2	The Black Box Problem	Rudin 2019: Stop Explaining Black Box Machine Learning Models...

COURSE SYLLABUS

		<p>Buckner 2023: Black Boxes or Unflattering Mirrors? Comparative Bias in the Science of Machine Behaviour</p> <p><i>Discussion Section Group Activity #1</i> Presentations Sep 5 Case Study on the Turing Test</p>
<p>Week 4 Sep 9</p>	<p>Basic Machine Learning #2: Transformers</p>	<p>Millière & Buckner 2022: A Philosophical Introduction to Language Models Pt. 1 Block 1981: Psychologism and Behaviorism</p> <p><i>Discussion Section Group Activity #1</i> Group Issue Briefs and Individual Discussant Appraisals Due Wed Sep 10 11:59 PM Case Study on the Turing Test</p> <p><i>Discussion Section Group Activity #2 Assigned:</i> Case Study on Bias in DNNs</p>
<p>Week 5 Sep 16</p>	<p>Interventionist Interpretability Methods</p>	<p>Millière & Buckner 2023: Interventionist Methods for Interpreting Deep Neural Networks Vredenburg 2022 - “The Right to Explanation”</p>
<p>Week 6 Sep 23</p>	<p>Algorithmic Bias 1</p>	<p>Fazelpour and Danks 2021: Algorithmic Bias—Senses, Sources, Solutions Julia Angwin 2016 – “Machine Bias” ProPublica Corbett-Davies et al. 2016: “A computer algorithm used for bail...”</p> <p><i>Discussion Section Group Activity #2</i> Presentations Fri Sept 26 Case Study on Bias in DNNs</p>
<p>Week 7 Sep 30</p> <p>Oct 2: *Exam #1*</p>	<p>Algorithmic Bias 2</p>	<p>Johnson 2020 – Algorithmic Bias-on the implicit biases of social technology Creel and Hellman 2022 – The Algorithmic Leviathan</p> <p><i>Discussion Section Group Activity #2</i> Group Issue Briefs and Individual Discussant</p>

COURSE SYLLABUS

		<p style="text-align: center;">Appraisals Due Wed Oct 1 11:59 PM: Case Study on Bias in DNNs</p> <p style="text-align: center;"><i>Discussion Section Group Activity #3 Assigned:</i> Case Study On Responsibility for AI Accidents</p>
Week 8 Oct 7	Responsibility	<p style="text-align: center;">Mathias 2004 – The Responsibility Gap Tigard 2021 – There is no Techno-responsibility Gap</p>
Week 9 Oct 14 Homecoming Holiday Oct 17	Trust	<p style="text-align: center;">Simion & Kelp 2023 – Trustworthy Artificial Intelligence Hevelke & Nida-Rumelin 2015 – Responsibility for Crashes of Autonomous Vehicles</p>
Week 10 Oct 21	The Alignment Problem	<p style="text-align: center;">Anthropic Team 2022 – Constitutional AI Gabriel 2020 – Artificial Intelligence, Values, and Alignment</p> <p style="text-align: center;"><i>Discussion Section Group Activity #3</i> Presentations Oct 24: Case Study on Responsibility for AI Accidents</p>
Week 11 Oct 28	Algorithms, Echo Chambers, and Mental Health	<p>Nguyen 2020 – “Echo Chambers and Epistemic Bubbles”</p> <p>Munroe 2024 – “Echo Chambers, Polarization, and ‘Post-Truth’- In Search of a connection”</p> <p>Wells et al. “Facebook knows Instagram is Toxic for Teen Girls” (WSJ)</p> <p>Lewis, “Our minds can be hijacked” (The Guardian)</p> <p style="text-align: center;"><i>Discussion Section Group Activity #3</i> Group Issue Briefs and Individual Discussant Appraisals Due Wed Oct 29 11:59 PM: Case Study on Responsibility for AI accidents</p> <p style="text-align: center;"><i>Discussion Section Group Activity #4 Assigned:</i> Case study on AI Alignment</p>

COURSE SYLLABUS

Week 12 Nov 4	Generative AI, Art, and Intellectual Property	Vlaad 2024 – A Portrait of the Artist as Young Algorithm Kieval 2024 – Artificial Achievement Goetze 2024 – AI art is theft <i>Discussion Section Group Activity #4:</i> Presentations Nov 7 Case study on AI Alignment
Week 13 Nov 11 No Class Nov 11 Holiday	The Future of Work	Danaher 2017 – Will life be worth living in a world without work? Belic 2024 – Institutions, Automation, and Legitimate Expectations <i>Discussion Section Group Activity #4</i> Group Issue Briefs and Individual Discussant Appraisals Due Wed Nov 12 11:59 PM: Case Study on AI Alignment <i>Discussion Section Group Activity #5 Assigned:</i> Case Study on Generative AI
Week 14 Nov 18	Finish lecturing and review for final exam	Discussion Section Group Activity #5: Presentations on Nov 21 Case Study on Generative AI
Nov 25	Holiday	No class Nov 24-28
Dec 2 Week 15	Final Exam Dec 2 In class	Final Exam <i>Discussion Section Group Activity #5</i> Group Issue Briefs and Individual Discussant Appraisals Due Friday Dec 5 11:59 PM: Case Study on Generative AI

COURSE SYLLABUS

This course complies with all UF academic policies. For general information on those policies and resources for students, please see the “Academic Policies & Resources” page at <https://syllabus.ufl.edu/syllabus-policy/uf-syllabus-policy-links/>."

Statement on academic dishonesty:

In addition to standard policies on honor code and academic honesty linked above, clear evidence of academic dishonesty means an automatic failing grade for the entire course. Note that you are encouraged to engage with generative AI tools for many of the projects, but transcripts of all interactions with AI must be included in the appendices of group project issue briefs.

Instructional material note:

Instructional materials for this course consist of only those materials specifically reviewed, selected, and assigned by the instructor(s). The instructor(s) is only responsible for these instructional materials.