

PHIL 3101: AI Ethics & Policy

Dr. Justin B. Biddle Carter School of Public Policy Office: D.M. Smith 212 Email: justin.biddle@pubpolicy.gatech.edu Phone: 404 385-2559	Course Time: TR 5:00-6:15pm Course Location: George/Scheller Tower 0268 Office Hours: TBA
--	--

Course Description:

This course will explore the ethical and philosophical dimensions of artificial intelligence (AI). AI – and related fields of data science and machine learning (ML) – are transforming the world in which we live. They have the potential to bring tremendous benefits, but they also involve risks, including risks of privacy harms; human rights violations; social injustice and inequality; alienation, and – according to some – human extinction. In this course, we will examine conceptual tools and frameworks that deepen our understanding of the ethical and philosophical issues associated with AI; we will probe these tools and frameworks in the context of current cases and challenges (including discussions of bias, human-machine interaction, surveillance, machine consciousness, impacts on democracy, environmental degradation, and others), and we will explore policies and standards that help us to realize our shared goals and values.

Texts: PDFs posted on Canvas and Perusall

Learning Objectives:

- **Ethical Awareness:** This course should help to cultivate your ethical awareness – that is, your ability to identify ethical issues in specific contexts in which they are relevant, especially in AI/ML. If you examine an AI/ML application, and someone asks you, “what are some of the ethical issues associated with this application,” you should be able to answer this question in a thoughtful and systematic manner.
- **Ethical Knowledge:** This course should help you to develop and deepen your knowledge of important ethical theories, frameworks, and concepts – including welfare, liberty, autonomy, equality, privacy, fairness, and others – and how these relate to ethical and responsible AI.
- **Ethical Reasoning:** This course should help you to identify, construct, and evaluate ethical arguments that are based on the theories, frameworks, and concepts mentioned above.
- **Communication:** This course should improve your ability to communicate, both orally and in writing, in a way that is clear, precise, and supported with reasons. You should improve your ability to do this even if the issues in question are controversial and subject to disagreement.

Course Format: The format of this course will be discussion based. We will spend most class sessions in small or large group discussions; I will lecture only as needed. For the first two months of the course, most class sessions will be structured around reading assignments, including both scholarly essays and popular media focusing on current issues. Most of these reading assignments will be done through Perusall. The final month of the course will be devoted to project work. Attendance at all sessions is required.

Course Requirements:

Perusall Reading Assignments	25%
Quizzes	30%
In-Class Assignments	10%
Final Project	35%

Grade Scale:

[90, 100]	A
[80, 90)	B
[70, 80)	C
[60, 70)	D
[0, 60)	F

This should be obvious, but I will not accept work after the semester ends. When the semester has finished, if you are unhappy with your grade, please do not contact me asking to additional work that could change your grade. The answer will be no.

For those who take the class **pass/fail**: in order to receive a grade of “satisfactory,” you must receive a grade of 70% or higher, and you must complete the short essays and final project.

Reading Assignments (Perusall): During the first part of the course, you will be required to read some assigned texts through Perusall, an AI-based learning platform that facilitates good reading practices and collaborative activity. Perusall allows you to comment on texts, pose questions that other students can read, respond to questions and comments, and more. It also assigns you a score for each reading assignment based on the extent of your comments and interactions. Part of your course grade will be calculated from these scores.

Quizzes: There will be five short quizzes given over the course of the semester. Each quiz will consist of 2-4 short-answer questions covering material from the readings, lectures, or class discussions. Quizzes will be given at the beginning of class and will last approximately 10 minutes. Please bring paper and a pen for the quizzes.

In-Class Discussion Assignments: During the first part of the course, I will regularly divide students into small groups and ask each group to discuss questions that I provide. Following these small-group discussions, I may ask students to report to the rest of the class on how they answered each question. Students will receive credit for successfully completing these assignments. I do not grade them stringently; you will receive full credit if I judge that you have taken the assignment seriously and put full effort into it.

Final Project: For the final part of the course, students will develop an ethical analysis of an AI/ML application and give a presentation on their project in class. Projects are designed to be conducted in small groups. Presentations will be 20 minutes, followed by 10 minute question and answer periods. Further details about the project will be provided during the semester.

Class Participation: It is expected that you participate regularly in class discussions and, more generally, contribute to a positive atmosphere for intellectual inquiry. This involves contributing to class discussions in both large and small groups, as well as being attentive (both to me and to your classmates), respectful, and considerate. Examples of behavior that detracts from a positive and respectful learning environment include: failure to listen to others, consistent interruption of others, disrespectful speech or incivility, failure to engage in class discussions (by,

for example, doing your thermodynamics homework in class), tardiness, falling asleep in class, etc. While you will not receive a grade for participation, you will most develop a better understanding of the course material, and will likely perform better on quizzes, if you participate actively.

Attendance: Attendance at all class sessions is required, unless it is excused – that is, cleared with me in advance or documented. If you have more than one unexcused absence, you will receive a final-grade deduction of two percentage points for each unexcused absence that follows the first. More simply: you get one freebie, and after that, it's two percentage points off for each unexcused absence. If you need to miss multiple class sessions, please reach out to the Dean of Students.

Student-Faculty Expectations: At Georgia Tech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See <http://www.catalog.gatech.edu/rules/22/> for an articulation of some basic expectations that you can have of me and that I have of you. In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class.

Academic Integrity: Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code, please visit <http://www.catalog.gatech.edu/policies/honor-code/> or <http://www.catalog.gatech.edu/rules/18/>. Any student suspected of cheating or plagiarizing on any assessed work will be reported to the Office of Student Integrity, which will investigate the incident and identify the appropriate penalty for violations.

Accommodations for Students with Disabilities: If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 or <http://disabilityservices.gatech.edu/>, as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible in order to set up a time to discuss your learning needs.