

INF 385T.12: Ethics of AI

Instructor: Ken Fleischmann

Pronouns: he/him/his

Email: kfleisch@ischool.utexas.edu

TA: Jaxsen Day (Primary Course Contact)

Pronouns: he/him/his

Email: jaxsen@utexas.edu

Course Description

COURSE DESCRIPTION

Artificial intelligence (AI) is both a product of and a major influence on society. As AI plays an increasingly important role in society, it is critical to understand both the ethical factors that influence the design of AI and the ethical dimensions of the impacts of AI in society. The goal of this course is to prepare information professionals for the important ethical responsibilities that come with developing systems with important, perhaps even life-and-death, consequences.

Students first learn about both the history of ethics and the history of AI, to understand the basis for contemporary, global ethical perspectives (including non-Western and feminist perspectives) and the factors that have influenced the design, development, and deployment of AI-based systems. Students then explore the societal dimensions of the ethics and values of AI. Finally, students explore the technical dimensions of the ethics and values of AI, including design considerations such as fairness, accountability, transparency, power, and agency.

Students should take this course to prepare them for the ethical challenges that they will face throughout their careers, and to carry out the important responsibilities that come with being an information professional. The ethical dimensions of AI may have important implications for information professionals and their employers. For example, the release of unsafe or biased AI-based systems may cause liability issues and reputational damage. This course will help students to identify design decisions with ethical implications, and to consider the perspectives of users and other stakeholders when making these ethically significant design decisions.

Students who perform well in this class will be positioned to take on a leadership role within their organizations and will be able to help guide and steer the design, development, and deployment of AI-based systems in ways that benefit users, other stakeholders, their organizations, and society. The knowledge and skill gained through this course will benefit students throughout their careers, and society as a whole will benefit from ensuring that information professionals are prepared to consider the important ethical dimensions of AI.

LEARNING OUTCOMES

1. You will learn about the history of AI and the ethical challenges that arise from recent advances in AI
2. You will gain an awareness of a wide range of ethical theories and learn to apply them to the ethics of AI
3. You will learn about efforts to develop principles for the design of ethical AI and develop your own view

How Will You Learn?

STATEMENT OF LEARNING SUCCESS

The ethics of AI is a timely, important, and expansive topic that defies the boundaries of space and time. Although ethics of AI was considered a fairly narrow, minor, and esoteric topic when I began pursuing research on this topic over 20 years ago, today it is understood as a critical issue across academia, industry, government, and nonprofits. I hope that this course will provide a launchpad for everyone to begin to explore the ethics of AI, and I will be happy to support your learning during and beyond the semester in any ways that I can. I hope the flexibility of the videos and online discussions will give you space and time to grapple with these important issues and help fit the course into your busy schedules.

TEACHING MODALITY INFORMATION

This course is fully asynchronous, and work can be done at any point during the week, although it is important to note that topics are covered on a week-by-week basis, and there are multiple consistent deadlines during each week. The deadlines are Wednesdays for discussion questions (which can be submitted before or after completing the readings and watching the videos) and Mondays for quizzes and discussion answers. During each week, you should first post your discussion question, then complete the readings, then watch the videos (including completing the embedded quizzes), and then post your answers. The goal is to ensure that you are actively engaged in learning each week, as this engagement will assist you in retaining this knowledge in the future. The quizzes and other elements of the course are intentionally designed to be as straightforward as possible – the goal of the course is to assist and support you in your learning journey, and to build a learning community that can facilitate everyone’s learning experience.

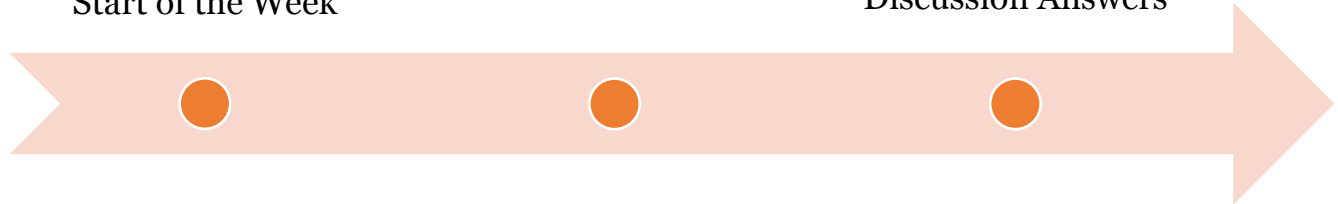
WEEKLY ACTIVITIES



WEEKLY SCHEDULE

Tuesdays 12:00 am
Start of the Week

Mondays 11:59 pm
Readings
Videos/Quizzes
Discussion Answers



Wednesdays 11:59 pm
Discussion Question

COMMUNICATION

The course Canvas site can be found at utexas.instructure.com. You are responsible for ensuring that the primary email address you have recorded with the university is the one you will check regularly for course communications because that is the email address that Canvas uses. Due to the asynchronous, online nature of the course, it is critical that you regularly check Canvas and your Canvas-linked email account.

ASKING FOR HELP

Whenever you have any questions or need help, please contact the course TA, Jaxsen Day, at jaxsen@utexas.edu

UNIVERSITY POLICIES AND RESOURCES

For a list of important university policies and helpful resources that you may need as you engage with and navigate your courses and the university, see the [University Policies and Resources Students Canvas](#) page. The page includes the language of the University Honor Code, Title IX legal requirements for Texas employees, and information about how to receive support through the office of Disability & Access.

REQUIRED MATERIALS AND DEVICES

All required materials are available via the course Canvas site. To be successful in this course, you will need a device capable of accessing the course Canvas site, as well as reliable internet access to use Canvas and email.

Course Requirements and Grading

CLASS EXPECTATIONS

Class participation Since this is an online, asynchronous course, class participation for this course will involve completing all weekly readings, watching all weekly videos and completing the embedded quizzes, and participating in all weekly discussion boards (discussion questions and answers board and take-aways board). Timely completion of all assignments is essential for you to stay on track and be successful in the class.

Professional standards You are individually responsible for completing all assignments on time and independently – everyone will be graded on their own individual work. We are collectively responsible for maintaining a respectful environment in the discussion boards at all times, and treating peers with respect. Discrimination, personal attacks, or hate speech of any kind will not be tolerated, and will have consequences.

ASSIGNMENTS

The following table represents how you will demonstrate and how we will assess your learning:

Assignments	Percent of Total Grade
Quizzes	20%
Discussion Questions	10%
Discussion Answers	20%
Midterm Assignment	25%
Final Assignment	25%

ASSIGNMENT DEADLINES AND LATE WORK PENALTIES

The Canvas gradebook is set up to automatically deduct 10% of the total assignment grade per day, on top of any deductions based on the assignment rubric. Please make sure to turn in all assignments on time. You are responsible for your own time management and setting aside the time to complete class assignments – think of the videos as a way to more flexibly adjust the learning that you would usually do during a set weekly three-hour block to potentially multiple times throughout the week, and just as the readings and assignments for most classes would be in addition to this class contact time, similarly these activities will be in addition to the videos. Please manage your time accordingly, and make sure that you will be able to set aside enough time each week to complete all of the required activities and assignments. You have some flexibility during the week, and you can always choose to turn in assignments earlier during the week, but setting aside time during each week to devote to completing the weekly class activities and assignments is essential. As such, we will not accept excuses such as having a busy week or not getting around to the work in time – in those cases, the late penalty will stand as is. For emergency situations, please work with Student Emergency Services so that they can notify us, allowing us to extend deadlines as needed to accommodate the circumstances.

QUIZZES (20%)

To ensure that you are completing, paying attention to, and understanding the readings and the videos, the videos will include embedded quizzes. These quizzes will cover content from the readings and/or videos. Please answer the quizzes thoughtfully, as you will only be able to complete them once.

DISCUSSION QUESTIONS (10%)

At the start of the week, post a discussion question about any aspect of any topic covered during the week. Asking the question does not require prior preparation; your question does not need to demonstrate a full comprehension of the week's topic(s), just that you are intellectually curious to learn more about the topic(s).

DISCUSSION ANSWERS (20%)

Choose any **two classmates' questions** and write one paragraph answers to each. Please make sure to support your answers with evidence from the readings and videos, citing at least **two sources per answer**. In each answer, at least one of your sources must be from the **current week's** course materials. Citations can be informal and can include the video (for example, you can just say: "as mentioned in the lecture about Deep Blue"). Please note that citing the same source twice will count as one source, even if you reference different parts of it. For full credit on this assignment, post your answers by the following Monday.

MIDTERM ASSIGNMENT (25%)

We will use generative AI tools to speculate about how prominent ethical theorists might react to contemporary AI ethics dilemmas for this assignment. By asking a tool like ChatGPT about how one of the theorists we have studied might react to an ethical dilemma, we have the opportunity to reflect how can we apply ethical theories to AI ethics dilemma and the capabilities and limitations of generative AI tools.

FINAL ASSIGNMENT (25%)

We will use generative AI tools to develop strategies for auditing a generative AI tool in terms of a specific value (e.g., fairness, accountability, transparency, power, agency) and then reflect on what you learned about performing value audits for generative AI as well as about the strengths and limitations of generative AI as a potential collaborator.

GRADE CUTOFFS



Course Outline

All instructions, assignments, readings, rubrics and essential information will be on the Canvas website at utexas.instructure.com. Check Canvas regularly. **Changes** to the schedule may be made at my discretion if circumstances require. I will announce any such changes in class and will also communicate them via a Canvas announcement. It is your responsibility to note these changes when announced, and I will do my best to ensure that you are notified of changes with as much advance notice as possible.

Week	Class Topic(s)	Readings	Assignments
1	Introduction	Class Syllabus	Videos/Quizzes
Units 1a and 1b: History of Ethics and History of AI			
2	Indian Ethics	Bilimoria	Videos/Quizzes Discussion Qs/As
	Classical Chinese Ethics	Hansen	
	Babbage's Engines	Schaffer	
3	Buddhist Ethics	DaSilva	Videos/Quizzes Discussion Qs/As
	Islamic Ethics	Nanji	
	Dartmouth Conference on AI	Kline	
4	Kantian Ethics	ONeill	Videos/Quizzes Discussion Qs/As
	Consequentialism	Hooker	
	Deep Blue	Hankey	
5	Distributive Justice	Murray	Videos/Quizzes Discussion Qs/As
	Virtue Ethics	Slote	
	Watson	Chandrasekar	
6	Ethics of Care	Held	Videos/Quizzes Discussion Qs/As
	Ubuntu	Ujomudike	
	Autonomous Cars	Oliveira et al.	

Unit 2: Ethics and Values of AI: Societal Dimensions			
7	Human Values and Value-Sensitive Design	Cieciuch et al. Friedman et al. Cheng & Fleischmann	Videos/Quizzes Discussion Qs/As
8	Codes of Ethics	ACM Code Of Ethics ALA Code Of Ethics ALA Core Values Fleischmann	Videos/Quizzes Discussion Qs/As Midterm Due
9	AI Ethics Guidelines	Jobin et al. Hagendorff et al. White House	Videos/Quizzes Discussion Qs/As
Unit 3: Ethics and Values of AI: Technical Dimensions			
10	Fairness	Buolamwini & Gebru Guo et al. Lee	Videos/Quizzes Discussion Qs/As
11	Accountability	Shah Raji et al. Slota et al.	Videos/Quizzes Discussion Qs/As
12	Transparency	Fleischmann & Wallace Larsson & Heintz Miller	Videos/Quizzes Discussion Qs/As
13	Power	Kalluri Campolo & Crawford Birhane et al.	Videos/Quizzes Discussion Qs/As
14	Agency	Fleischmann Johnson & Verdicchio Neff & Nagy	Videos/Quizzes Discussion Qs/As Final Due