

INF 385T/350G, Ethics of AI: Theorizing Good Systems
27284/27134
Fall 2019

Class Meets: *Thursdays, 12:00-2:45 pm, UTA 1.504*

Instructor: Dr. Kenneth R. Fleischmann

Pronouns: he, him, his

Email: kfleisch@ischool.utexas.edu

Office: UTA 5.534

Office Hours:

Thursdays, 11:30 am – noon, 2:45-3 pm,
via e-mail, and by appointment

Course Description

Course Description: Artificial intelligence is becoming an increasingly important fixture in our everyday lives. This course will help you to understand the ethical challenges related to AI and how to address them. To do this, the course will provide you with a broad range of ethical theories, including non-Western and feminist theories, and teach you to apply these theories to contemporary ethical challenges resulting from AI.

What will I learn?

1. You will gain an awareness of a wide range of ethical theories and learn to apply them to the ethics of AI
2. You will learn about efforts to develop principles for the design of ethical AI
3. You will improve your ability to articulate your own values and to understand and appreciate others' values

How will I learn?

1. You will learn from the readings, which will cover ethical approaches and their application to AI
2. You will learn from the class discussions, which will cover reading topics as well as writing a research paper
3. You will learn by writing a research paper exploring an ethical dilemma in AI and approaches for addressing it

How to succeed in this course

1. Please make sure to complete all readings prior to class and all assignments by their due dates and times
2. Please make sure to come into the class with an open mind, and to develop your own perspectives
3. Please make sure to be respectful during all class discussions, and to appreciate others' perspectives

Course Requirements

Required Materials All course readings will be available via the course Canvas site.

Required Devices There are no required devices, but please bring the readings for the week, in print or electronically.

Classroom expectations Please make sure to be engaged and respectful at times, and to contribute to discussions.

Assignments

1. **Assignment 1: Attendance and Participation**

- (a) Because the vast majority of the learning in this class will occur within the classroom, you are required to attend class regularly. Attendance will be taken during each class period. Absences will only be excused in situations following university policy (illness, religious holy days, participation in University activities at the request of university authorities, and compelling absences beyond your control) with proper documentation and timely notification (prior to class for non-emergencies). Excessive tardiness may be considered as an unexcused absence except in situations following university policy.
- (b) Class participation is a critical element of this course. The effectiveness of the course will be significantly impacted by the quality of your participation. Class participation is not merely attendance, but rather factors in your overall contributions to the collaborative learning environment, based on both the quantity and quality of your interactions in all aspects of the course. Discussion of class participation with the instructor is encouraged in order to ensure that you are making the most of the classroom experience and the accompanying opportunities for learning. You are expected to participate in all aspects of class discussion. You should come to class prepared

to discuss the required readings, as well as your perspectives on these readings. You should strive for balance in your contributions, and your participation will not be based on who speaks the loudest or the longest, but on consistent participation of significant quantity and, most importantly, quality.

- (c) Your attendance and class participation grade will be calculated by multiplying the numerical assessment of your class participation by the percentage of classes that you attend (with exceptions made for documented, university-recognized absences as noted above). Please note that regular attendance and active participation in each class session are critical for receiving a good grade in this course. For example, by actively participating in each class, you will receive a full letter grade higher than if you were to skip half of the classes or to be half-awake for all of the classes.

2. **Assignment 2: Discussion Questions**

At 9 am on each class day (except the first week and last three weeks), you will submit three discussion questions via Canvas. Each question should be one sentence in length. Please be sure to directly refer to and synthesize all three of the readings for the week.

3. **Assignment 3: Research Paper**

Throughout the semester, you will develop a research paper that explores a specific topic within ethics of AI. You are required to apply readings from the course as well as readings from outside of the course, including incorporating any three ethical theories as part of your analysis. We will spend a significant amount of time early in the semester to helping you select your topic, and we will continue to devote approximately half of the class time each week to collectively brainstorming how to develop and improve everyone's research papers. You are ultimately responsible for your own paper, but there will be many opportunities to get formal and informal feedback on your ideas and writing.

- (a) Paper Proposal: Your paper proposal must include the proposed title of your research paper as well as a one-page description of the goals of your research paper. At this stage, you are not required to have settled on your ethical theories yet. Please make sure to include your strategy for finding appropriate outside readings. Research paper proposals will be graded according to the following criteria: suitability of the topic, creativity of the topic, and appropriateness of the plan for finding outside readings.
- (b) Paper Outline: Your paper outline will typically include bulleted lists, filling in as much detail as you have ready at this point, and incorporating feedback from the paper proposal. Your paper needs to be broken into sections, and you need to include your three planned ethical theories as part of your paper outline. You also need to include a list of at least 10 references of relevance to the paper, including at least 5 references that are peer-reviewed journal articles, refereed conference proceedings papers, or academic books or book chapters from beyond the course syllabus. Citations must follow APA Style (there are many useful online resources available; I recommend: <https://owl.english.purdue.edu/owl/resource/560/01/>). Research paper outlines will be graded according to the following criteria: clarity of the plan for the paper; appropriateness of the plan for the paper; and appropriateness of the references.
- (c) Paper Rough Draft: Your rough draft must be at least a half-complete version of your final paper and must incorporate feedback from the proposal and outline. You may choose to what extent and in what ways you develop the paper; for example, you can provide the first half of the paper, or half of each section of the paper. The more complete your rough draft is, the more useful the feedback you receive from the instructor and your peers will be. The rough draft should be 1,500-5,000 words. You must cite at least 20 sources, including at least 10 outside scholarly readings, and all citations must follow APA Style. Research paper rough drafts will be graded according to the following criteria: suitability of the writing, creativity of the writing, and overall coherence and clarity of writing.

- (d) **Reviews of Rough Drafts:** You will be selected to review two papers written by peers. Please prepare one-page reviews of each paper, including a brief summary of the purpose and content of the paper as you understood it, the strengths of the paper, and constructive feedback on how to improve it. Reviews will be graded according to the following criteria: clarity, insightfulness, and helpfulness.
- (e) **Final Paper:** Your final paper must incorporate feedback from all previous stages of the paper development. Your paper must be complete, coherent, and easy to read. Please make sure to proofread your paper thoroughly prior to submission. The final paper should be 3,000-5,000 words. You must cite at least 30 sources, including at least 15 outside scholarly readings, and all citations must follow APA Style. Research paper rough drafts will be graded according to the following criteria: suitability of the writing, creativity of the writing, and overall coherence and clarity of writing.
- (f) **Final Presentation:** During one of the three final weeks of the semester, you will present your research paper to the class in the form of an audiovisual presentation. The presentation week and length will be announced well in advance of the date. You may give your presentation using a visual aid such as PowerPoint, Prezi, or Keynote; create and demonstrate a website or other technology; or create a video. Research paper presentations will be graded according to the following criteria: quality of research, quality of visual presentation, and quality of oral presentation.

Grading for this Course

The following table represents how you will demonstrate your learning and how I will assess the degree to which you have done so.

Assignments	Percent of Total Grade
1. Attendance and Participation	20%
2. Discussion Questions	20%
3. Research Paper	60%
3a. Paper Proposal	5%
3b. Paper Outline	5%
3c. Paper Rough Draft	10%
3d. Peer Reviews	10%
3e. Final Paper	20%
3f. Final Presentation	10%

Course Outline

All instructions, assignments, readings, rubrics and essential information will be on the Canvas website at <https://utexas.instructure.com>. Check this site regularly and use it to ask questions about the course schedule.

Changes to the schedule may be made at my discretion and if circumstances require. It is your responsibility to note these changes when announced (although I will do my best to ensure that you receive the changes with as much advanced notice as possible).

Week	Date	Class Topic	Class Readings	Assignments
1	8/29	Introduction to the Ethics of AI		
Unit 1: Ethical Foundations				
2	9/5	Hindu Ethics, Classical Chinese Ethics, and Virtue Theory	Bilimoria Hansen Pence	
3	9/12	Buddhist Ethics, Islamic Ethics, and Kantian Ethics	De Silva Nanji O'Neill	
4	9/19	Utilitarianism, Social Contract Theory, and Rational Egoism	Goodin Kymlicka Smith, T.	
5	9/26	Ethics of Care, Ubuntu, and Situated Knowledges	Held Prinsloo Haraway	
6	10/3	Philosophy of Technology	Tiles & Oberdiek Smith, M.R. Winner	
Unit 2: Applying Ethical Theories to AI				
7	10/10	AI and Work	Daugherty & Wilson Grossman Mateescu & Elish	
8	10/17	AI and Health Care	Jiang et al. Vayena et al. Wang & Siau	
9	10/24	AI and Transportation	Faulhaber et al. Morton et al. Dietrich & Weisswange	
10	10/31	AI and Warfare	Coker Jensen et al. Umbrello et al.	
11	11/7	Establishing Principles for Ethical AI	Hagendorff Zeng et al. Torresen	
Unit 3: Final Presentations				
12	11/14	Final Presentations		
13	11/21	Final Presentations		
14	12/5	Final Presentations		

Policies

Classroom Policies

Statement on Learning Success

Your success in this class is important to me. We will all need accommodations because we all learn differently. If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we'll develop strategies to meet both your needs and the requirements of the course. I also encourage you to reach out to the student resources available through UT. Many are listed on this syllabus, but I am happy to connect you with a person or Center if you would like.

Absences

Excused Absence: The only absences that will be considered excused are for religious holidays or extenuating circumstances due to an emergency. If you plan to miss class due to observance of a religious holiday, please let me know at least two weeks in advance. You will not be penalized for this absence, although you will still be responsible for any work you will miss on that day if applicable. Check with us for details or arrangements.

If you have to be absent, use your resources wisely. Ask your team and other classmates to get a run-down and notes on any lessons you miss. If you find there are topics that we covered while you were gone that raise questions, you may come by during office hours or schedule a meeting to discuss. Email specific questions you have in advance so that we can make the most of our time. "What did I miss?" is not specific enough.

Student Rights & Responsibilities

- You have a right to a learning environment that supports mental and physical wellness.
- You have a right to respect.
- You have a right to be assessed and graded fairly.
- You have a right to freedom of opinion and expression.
- You have a right to privacy and confidentiality.
- You have a right to meaningful and equal participation, to self-organize groups to improve your learning environment.
- You have a right to learn in an environment that is welcoming to all people. No student shall be isolated, excluded or diminished in any way.

With these rights come responsibilities:

- You are responsible for taking care of yourself, managing your time, and communicating with the teaching team and with others if things start to feel out of control or overwhelming.
- You are responsible for acting in a way that is worthy of respect and always respectful of others. • Your experience with this course is directly related to the quality of the energy that you bring to it, and your energy shapes the quality of your peers' experiences.
- You are responsible for creating an inclusive environment and for speaking up when someone is excluded.
- You are responsible for holding yourself accountable to these standards, holding each other to these standards, and holding the teaching team accountable as well.

Personal Pronoun Use (She / He / They / Ze / Etc)

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, unless they have added a "preferred name" with the Gender and Sexuality Center (<http://diversity.utexas.edu/genderandsexuality/publications-and-resources/>). I will gladly honor your request to address you by a name that is different from what appears on the official roster, and by the

[gender pronouns you use \(she/he/they/ze, etc\)](#). Please advise me of any changes early in the semester so that I may make appropriate updates to my records.

University Resources for Students

Services for Students with Disabilities

This class respects and welcomes students of all backgrounds, identities, and abilities. If there are circumstances that make our learning environment and activities difficult, if you have medical information that you need to share with me, or if you need specific arrangements in case the building needs to be evacuated, please let me know. I am committed to creating an effective learning environment for all students, but I can only do so if you discuss your needs with me as early as possible. I promise to maintain the confidentiality of these discussions. Any student with a documented disability who requires academic accommodations should contact Services for Students with Disabilities at 471-6259 (voice) or 512-410-6644 (Video Phone) as soon as possible to request an official letter outlining authorized accommodations. For more information, visit <http://ddce.utexas.edu/disability/about/>.

Counseling and Mental Health Center

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. <http://www.cmhc.utexas.edu/individualcounseling.html>

The Sanger Learning Center

Did you know that more than one-third of UT undergraduate students use the Sanger Learning Center each year to improve their academic performance? All students are welcome to take advantage of Sanger Center's classes and workshops, private learning specialist appointments, peer academic coaching, and tutoring for more than 70 courses in 15 different subject areas. For more information, please visit <http://www.utexas.edu/ugs/slc> or call 512-471-3614 (JES A332).

Undergraduate Writing Center: <http://uwc.utexas.edu/>

Libraries: <http://www.lib.utexas.edu/>

ITS: <http://www.utexas.edu/its/>

Student Emergency Services: <http://deanofstudents.utexas.edu/emergency/>

BeVocal

BeVocal is a university-wide initiative to promote the idea that individual Longhorns have the power to prevent high-risk behavior and harm. At UT Austin all Longhorns have the power to intervene and reduce harm. To learn more about BeVocal and how you can help to build a culture of care on campus, go to: wellnessnetwork.utexas.edu/BeVocal.

Important Safety Information:

If you have concerns about the safety or behavior of fellow students, TAs or Professors, call BCAL (the Behavior Concerns Advice Line): 512-232-5050. Your call can be anonymous. If something doesn't feel right – it probably isn't. Trust your instincts and share your concerns.

The following recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767, <http://www.utexas.edu/safety/>

Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.

- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.
- In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
- Link to information regarding emergency evacuation routes and emergency procedures can be found at: www.utexas.edu/emergency

Title IX Reporting

Title IX is a federal law that protects against sex and gender based discrimination, sexual harassment, sexual assault, sexual misconduct, dating/domestic violence and stalking at federally funded educational institutions. UT Austin is committed to fostering a learning and working environment free from discrimination in all its forms. When sexual misconduct occurs in our community, the university can:

1. Intervene to prevent harmful behavior from continuing or escalating.
2. Provide support and remedies to students and employees who have experienced harm or have become involved in a Title IX investigation.
3. Investigate and discipline violations of the university's [relevant policies](#).

Faculty members and certain staff members are considered “Responsible Employees” or “Mandatory Reporters,” which means that they are required to report violations of Title IX to the Title IX Coordinator. **I am a Responsible Employee and must report any Title IX related incidents** that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee. If you want to speak with someone for support or remedies without making an official report to the university, email advocate@austin.utexas.edu For more information about reporting options and resources, visit titleix.utexas.edu or contact the Title IX Office at titleix@austin.utexas.edu.

The following recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767, <http://www.utexas.edu/safety/>

University Policies

Academic Integrity

Each student in the course is expected to abide by the University of Texas Honor Code: “As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity.” **Plagiarism is taken very seriously at UT.** Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Otherwise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT’s Academic Honesty and the University Honor Code which can be found at the following web address: http://deanofstudents.utexas.edu/sjs/acint_student.php

Q Drop Policy

If you want to drop a class after the 12th class day, you’ll need to execute a Q drop before the Q-drop deadline, which typically occurs near the middle of the semester. Under Texas law, you are only allowed six Q drops while you are in college at any public Texas institution. For more information, see: <http://www.utexas.edu/ugs/csacc/academic/adddrop/qdrop>