

CSC484 Ethics In Artificial Intelligence Syllabus

Professor Clark Elliott

Spring 2022-23

Logistics:

Class meets: Asynchronous Online / informal meetings some Thursdays 8:00 PM

Professor: Dr. Clark Elliott

Class website: <http://condor.depaul.edu/~elliott/484>

email: elliott@depaul.edu.

[Include "484: " prefix in subject line and MEANINGFUL mail header!]

Course grader email: celliottgrader@gmail.com

Course Management: [D2L.depaul.edu](https://d2l.depaul.edu)

Textbook:

None. Course readings and document URLs will be provided on D2L.

Grading:

Assignments	61.5%
Quizzes	23.5%
Exams	15%

Grading Scale:

95%	A
90%	A-
86, 83, 80	B+, B, B-
78, 74, 70	C+, C, C-
65, 60	D+, D

I reserve the right to raise the grade of a student that has demonstrated exceptional contributions in some particular portion of the class. I reserve the right to make changes in the points distribution of the three categories as necessary during the quarter, with notification to the students.

Topics:

In this graduate seminar course we will cover first a quick introduction in to traditional ethics reasoning, introducing Ethical Reasoning Frameworks such as Social Contract Theory, Kant and utilitarianism, along with several others. Following the introductory material we will introduce, and extensively discuss the many ethical challenges facing us in the modern world as AI becomes increasingly integrated into most aspects of society. Sample topics might include:

- ChatGPT and the future of education
- When it is UN-ethical NOT to use AI
- Massive AI-generated fake news and social media posts
- Lethal military AI & autonomous killing machines
- Implicit bias in capturing “the way it has been done”
- Artificial intimacy
- AI lie detection
- AI hiring and job performance reviews
- Enhanced cyborg humans
- AI Art & AI Art theft
- New forms and applications of AI
- AI-generated fake news
- Societal benefit of AI systems vs. protection for individuals
- AI Psychotherapists and psychiatrists
- AI Lawyers and Judges, including sentencing
- The need for explanations for AI-generated decisions
- Artificial sentience / life
- When LLMs are attacked through massive LLM posted fakes
- When the bulk of LLM input content becomes previous LLM output

CONTENT WARNING:

This is a graduate course at a major university, with a truly international student body. We will necessarily be discussing a wide range of topics in this course. Some can be psychologically and emotionally challenging. THIS IS A WARNING.

If you are uncomfortable with any of the topics we are discussing, or with what others are saying as part of the discussion, you must **immediately leave the session--both live and via asynchronous discussion, or stop the lecture video**. Then, send me email and we'll work it out. While it is important for us, in our mission as a major university, to cover many challenging subjects as part of our curriculum, **I don't want my students to be uncomfortable**. It is unlikely to be necessary for you to participate in any particular discussion. There are many things we can talk about in this ethics course, and we can find something else for you to do during a particular module.

I can't know what might be a trigger for you, but, for example we may very well cover subjects like the following:

- Ethical frameworks that are not consistent with particular religious teachings.
- Gender issues, such as may, e.g., come with artificial personalities, artificial companions, AI personality reasoning, and etc.

- Implicit bias built into AI systems regarding race, ethnic stereotyping, residency status, gender, age and many other aspects of being an individual in society who is also part of sub-groups.
- Sometimes frightening possible futures because of the rise of AI, as existing systems (banking, news, social media, military, political, stock markets and etc.) undergo major changes, or even become obsolete.
- The implications of overwhelming floods of fake news, fake faces, fake videos, fake voices, fake personalities hitting the Internet.
- Artificially intimate companions that may include aspects of human sexuality.
- Concerns about artificial child pornography and the resultant breakdown of legal systems based on de facto harm to children, which no longer apply because there are no longer real, de facto, children involved.
- Artificially intelligent autonomous killing machines for the military.
- Varied cultural norms concerning the acceptance of cheating in academia.
- Many other subjects.

Always let your conscience be your guide. NEVER put yourself, including your psychological or emotional well-being, at risk. It is better to be cautious first, then adjust as necessary later.

Class structure:

This is a reading, lecture, and discussion seminar-style class. Scholarly discussion of the ideas is the focus of the course. Discussion will take place online on the D2L forums. Students are expected to keep up on the reading, lectures and quizzes each week, and to actively participate in the discussions.

Each collection of readings / lecture topics will have a (generally weekly) closing date, and class discussion of the topic will officially end on that date. However, this is a *soft* close, and if you can inspire others to continue the dialogue with you, you are free to continue to post in older forums throughout the quarter.

Learning Goals:

At the end of class you will:

- Be able to give an outline of how basic, traditional, ethical reasoning works.
- Be able to give some details of four commonly respected frameworks for reasoning about ethical scenarios.
- Be testably well-studied in many of the most pressing new ethical issues arising in the modern world because of AI technology
- Be able to intelligently and productively converse with others on hot-button topics regarding the ethics of AI in the modern world.

Office hours for the course are available from my faculty link at cdm.depaul.edu

All assignments, the assignment schedule, and the course materials, are available online at d2l.depaul.edu.

Submission File Formats:

All submissions to D2L MUST BE IN THE SPECIFIED FORMAT or they will not be accepted for credit. There are very strict rules on file formats for this class and thus no other (alternate) formats will be graded. Follow the submission instructions for each assignment.

Sometimes this means in STANDARD ZIP FORMAT, including submissions of a single file. No 7zip files, no rar files, etc. No exceptions. In some cases zip files are prohibited.

Sometimes this means Microsoft Word format (.docx), or in plain HTML (.html), or Text (.txt) including text files contained within a ZIP archive. (Free programs are available to produce each of these formats.) NO PDF FILES.

Follow the instructions for each assignment.

Students are responsible for downloading their assignments after uploading, to make sure that files have not been corrupted. Corrupted files will not be graded.

NO LATE ASSIGNMENTS will be accepted for credit, unless otherwise noted. If you miss a deadline, time to move on!

Academic Integrity:

Students may be required to view the short lecture on Academic Integrity and to pass the Academic Integrity Quiz with a 100% score before other coursework will be accepted for credit.

Cheating, plagiarism, and unethical conduct are not allowed, and will be sanctioned, including referral to the dean's office, and failure in the class (default) or (occasionally) failure on the assignment and an additional 20-40% reduction in the course grade. Please refer to the academic handbook by which rules you are expected to abide. This course will be subject to the university's academic integrity policy. More information can be found at <http://academicintegrity.depaul.edu/>. If you have any questions be sure to consult with the professor.

Violations include, but are not limited to: making false claims on any checklist for work that has not been done; including ANY un-cited work of others in any documents you turn in; turning in work, including any program, that has been authored by someone other than yourself and in some cases including *any* work of others, whether cited or not—see the rules for each assignment; using artificial intelligence tools to author text of any assignments.

Publicly sharing or posting online any prior or current materials from this course (including exam questions and/or answers) is considered to be providing unauthorized assistance prohibited by the policy. Both students who share/post and students who access or use such materials are considered to be cheating under the Policy and will be subject to sanctions for violations of

Academic Integrity.

ChatGPT and other LLMs

Sadly, because it can be time-consuming and challenging to reliably detect the use of AI LLMs to produce written assignments, I have retired most such assignments until such time as I can determine that they can be fairly graded.

You are free to use ChatGPT (or other LLMs) for *background research* as you would Wikipedia and Google Search. However, you are NEVER allowed to paste output from an LLM directly into the discussion forums (or assignments). EVER! The forums are to be used strictly for your ideas, from your head, as *integral parts of the discussions*. I reserve the right to remove postings, or refuse assignments, I believe are the work of ChatGPT, and I am the sole judge. Officially, posting directly from an LLM is an academic integrity violation. Using unattributed verbatim text in assignments from an LLM is an academic integrity violation.

Checklists

Checklists for this course, when we have them, are a **contract** between the student and the professor. The default for all checklist items is *No*. If you change an item to *Yes* you are claiming, absolutely, to have done the work indicated. An inaccurate checklist that claims work not actually done results in zero grade, a significant points penalty in the course, and, possibly a failing course grade, and sanction through the dean's office. We will *assume* that you are attempting to cheat your peers by claiming work that was not done. If you are really in doubt, change *No* to *Maybe* or *Probably* and give a short explanation at the bottom of the checklist. Examples would be, "I wrote the code and it usually works, but sporadically fails every tenth time or so" and "I wasn't sure if you meant three list items or three nodes; I did implement three nodes, but not three list items" and so on.

Example: Sam spends 35 hours on her JokeServer assignment, and does a very good, but not quite complete job on it. She turns it in before completing randomizing of the jokes (a minor part of the assignment worth only a few points). She checks *Yes* next to "randomizes jokes" and turns in her checklist with her assignment. Ordinarily she would have received 95 of 100 points for her work. **Result:** Zero points on the assignment. Additionally a 20% course penalty resulting in a D in the class. Referral to the dean's office. No remedy.

Citations

There may be written assignments, and programs, turned in for this class. The default *assumption* is that any work you turn in is your own. If you submit ANY work of others that is not clearly cited as being the work of others, you will be sanctioned: zero grade on the assignment, outright failure in the class, referral to the dean's office for possible expulsion from the university.

Example: Louis spends 100 hours over the course of the quarter, completing a thirty-page ethics paper. He includes a really interesting pair of paragraphs that he found on the web, but fails to cite the original work and author. He submits his paper for grading. Ordinarily he would have received full credit for his fine paper, and possibly even extra points for exceptional work in this area. **Result:** Zero grade for his paper. Failure in the class. Referral to the dean's office. No remedy.

When in doubt, cite! (But note that you still might be guilty of plagiarism if you've included too much of the work of others.) Change the font of included text, and possibly use quotation marks, to make absolutely clear that it is the work of others. Make absolute certain that you are allowed to include the cited work of others in a particular assignment. (E.g., not allowed in the study logs!)

Gray Areas

Sometimes there are gray areas. When you believe there is one, **ask!** There is no assumption of innocence because you forgot to ask, when it comes to academic integrity.

"Minor points" notation:

From time to time I use the point box as a communication vehicle in two specific ways, and I reserve the right to add minor points for this purpose:

- One point extra: I am tipping my hat to you for particularly fine work. That is, if you get 101 points on a 100 point assignment, I may be saying, "Hey, I noticed the five extra modules you wrote. Good job!"
- Two points extra: If you receive two extra points, I am acknowledging an *exceptional* contribution beyond expectations, so 102 points on a 100 point assignment is something to feel really good about, and is a rare compliment.
- Grade of "1": used as a placeholder to let a student know that I have reviewed an assignment, and am waiting for further information or work as per correspondence. A "1" will *always* be resolved to a different grade.
- Grade of "2": a serious warning that you need to communicate with me about possible plagiarism or some other irregularity that is being investigated.

More Policies

Changes to Syllabus

This syllabus is subject to change as necessary during the quarter. If a change occurs, it will be addressed during class, posted in D2L, and sent via email.

Online Course Evaluations

Evaluations are a way for students to provide valuable feedback regarding their instructor and the course. Detailed feedback will enable the instructor to continuously tailor teaching methods and course content to meet the learning goals of the course and the academic needs of the students. They are a requirement of the course and are key to continue to provide you with the highest quality of teaching. The evaluations are anonymous; the instructor and administration do not track who entered what responses. A program is used to check if the student completed the evaluations, but the evaluation is completely separate from the student's identity. Since 100% participation is

our goal, students are sent periodic reminders over three weeks. Students do not receive reminders once they complete the evaluation. Students complete the evaluation online in [CampusConnect](#).

Students will be given formal time to complete course evaluations.

Academic Integrity and Plagiarism

This course will be subject to the university's academic integrity policy. More information can be found at <http://academicintegrity.depaul.edu/>. If you have any questions be sure to consult with the professor.

Publicly sharing or posting online any prior or current materials from this course (including exam questions and/or answers) is considered to be providing unauthorized assistance prohibited by the policy. Both students who share/post and students who access or use such materials are considered to be cheating under the Policy and will be subject to sanctions for violations of Academic Integrity.

There is NO CHEATING OF ANY KIND in this class!

Academic Policies

All students are required to manage their class schedules each term in accordance with the deadlines for enrolling and withdrawing as indicated in the [University Academic Calendar](#). Information on enrollment, withdrawal, grading and incompletes can be found at: cdm.depaul.edu/enrollment.

Students with Disabilities

Students who feel they may need an accommodation based on the impact of a disability should have CSD contact the instructor privately to discuss their specific needs. All discussions will remain confidential. To ensure that you receive the most appropriate accommodation based on your needs, contact the instructor as early as possible in the quarter (preferably within the first week of class), and make sure that you have contacted the Center for Students with Disabilities (CSD) at: csd@depaul.edu.

Lewis Center 1420, 25 East Jackson Blvd.

Phone number: (312)362-8002

Fax: (312)362-6544

TTY: (773)325.7296

Other Course Policies

Attendance: Students are expected to attend each class when there is an on-campus section (unless otherwise notified), or view the class online. Attendance will not be formally taken, except as required for on-campus class sections. Unless otherwise noted ALL the course material presented in the lectures is suitable for exams.

Class Discussion: Student participation in class discussions is expected, and this will take place in class for local students, and online for all students. But see the following:

Uncomfortable with lecture or discussion topics: This is a university course. DePaul is a major urban university with many students from all over the world. If you are uncomfortable with any of the lectures, and/or discussions, for religious, gender, social or any other reasons, then stop the video, leave the room and leave the discussion immediately. Make contact with the professor in a timely way to discuss your concerns and work towards a resolution. We cover challenging topics in this university environment, but we are also sensitive to our students' needs.

Attitude: A professional and academic attitude is expected throughout this course. Measurable examples of non-academic or unprofessional attitude include but are not limited to: talking to others when the instructor is speaking, mocking another's opinion, cell phones ringing, emailing, texting or using the internet whether on a phone or computer. If any issues arise a student may be asked to leave the classroom.

[Cell phones / laptops in class]: If you need to use your cell phone for any reason, or your laptop for any reason other than following the class slides, and taking notes, *leave the room*. You may quietly leave and re-enter as often as necessary unless I note otherwise. Your peers devote many hours out of their busy lives, and thousands of dollars, to come to class. They deserve a vibrant, focused, environment. If you have a special case, discuss it with the instructor ahead of time. NO TEXTING, EMAIL, FACEBOOK, etc. in the classroom.]

Civil Discourse: DePaul University is a community that thrives on open discourse that challenges students, both intellectually and personally, to be [Socially Responsible Leaders](#). It is the expectation that all dialogue in this course is civil and respectful of the dignity of each student. Any instances of disrespect or hostility can jeopardize a student's ability to be successful in the course.