

PHILOSOPHY 290 (A01): Selected Topics
Ethics of Artificial Intelligence
CRN: 22467
Course Outline
Spring 2025

Instructor: Dr. C. Klatt

Email: cklatt@uvic.ca

Office Hours: Mondays and Thursdays, 1 -2pm, CLE B311

Text: No text to buy for this course. Readings will be provided online.

Course Objectives:

The development of technology that uses some form of Artificial Intelligence (AI) is increasing at a rapid pace. AI systems determine which news stories you view in social media, determine which applicants make it to the top of the list for job interviews, and interpret medical scans in order to diagnose any disease you may have. As the abilities of AI develop, we will increasingly see these algorithms making decisions on our behalf. But should we allow AI to take over roles that have so far only be entrusted to a rational human mind? Can we expect that AI will be able to make ethical decisions that reflect our values and desires? In this course we will examine three fundamental questions. First, should we consider AI to be capable of rational thought and ethical decision making? Second, what are the specific ethical issues that we need to consider when applying AI in various everyday uses? And third, what moral status do we give to an AI that has greater mental capacity than human beings?

Grading:

5 Online Quizzes @ 5% each	= 25%
Best 5 of 6 In-Class Written Assgs @ 6% each	= 30%
Mid-term Test @ 20% (Feb 24 th)	= 20%
End of term Test @ 15% (April 3 rd)	= 15%
Participation	= 10%

No Final Exam

The final grade will be recorded as a percentage.

A+ = 90 – 100; A = 85 – 89; A- = 80 – 84; B+ = 77 – 79; B = 73 – 76; B- = 70 – 72;
C+ = 65 – 69; C = 60 – 64; D = 50 – 59; F = 0 – 49.

Quizzes:

Quizzes will be completed online through BrightSpaces. I will spend some time during the first week of class explaining how to write the quizzes online. If you have any questions, please email me. The quizzes may be written any time **before 11:30 pm (Victoria time)** on the day it is due and they will be available to write on BrightSpace no fewer than two days before they are due. For exact due dates see the *Topic List* below. If

you miss the quiz you will be able to view the answers online after the due date as long as you have opened the quiz before the due time. Late quizzes will not be graded.

In-Class Written Assignments and Participation:

Although I will be presenting some material in each lecture, I intend to run the class as a group discussion of the issues. As such, you will be graded on your level of engagement through a participation grade and short written assignments that will be completed during class time.

Course Withdrawal Deadlines:

- Jan 19th: Withdrawal with 100% reduction of tuition fees
- Feb 9th: Withdrawal with 50% reduction of tuition fees
- Feb 28th: Last day for withdrawal (no fees returned)

Course Material on BrightSpaces:

I will put my lecture overheads, quizzes and any required readings or links to viewing material on BrightSpaces. I will also use our class BrightSpaces site to post grades.

When you need help:

If you have any questions or concerns about the course material or any other issue, please feel free to contact me at my email cklatt@uvic.ca. Alternatively, you may come and talk to me in person. I will have office hours in CLE B311 after our classes on Mondays and Thursdays from 1 – 2pm.

When you miss a test or assignment:

If you miss a test, email me as soon as possible. We will make arrangements for you to write the test when you are able.

If you are ill and think you will be unable to complete a quiz in time, contact me BEFORE the quiz due date and I will extend your time.

I will be counting the best 5 of your 6 written assignments. If you miss one in-class written assignment, it will count as your dropped grade. If you miss more than one written assignment and have a legitimate excuse (e.g. illness, serious family issue, athletic participation etc.), contact me and we will make arrangements for you to make up the missed material.

Academic Misconduct:

See the section *Policy on Academic Integrity* in the UVic calendar for information on cheating and its consequences.

https://www.uvic.ca/calendar/undergrad/index.php#/policy/Sk_0xsM_V?bc=true&bcCurrent=08%20-%20Policy%20on%20Academic%20Integrity&bcGroup=Undergraduate%20Academic%20Regulations&bcltemType=policies.

In particular, note that:

“Cheating includes, but is not limited to:

- copying the answers or other work of another person
- sharing information or answers when doing take-home assignments, tests or examinations except where the instructor has authorized collaborative work
- having in an examination or test any materials or equipment other than those authorized by the examiners
- accessing unauthorized information when doing take-home assignments, tests or examinations
- impersonating a student on an examination or test, or being assigned the results of such impersonation
- accessing or attempting to access examinations or tests before it is permitted to do so

Students found communicating with one another in any way or having unauthorized books, papers, notes or electronic devices in their possession during a test or examination will be considered to be in violation of this policy.

Aiding Others to Cheat

It is a violation to help others or attempt to help others to engage in any of the conduct described above.”

Cheating will not be tolerated and will result in a zero on the test or assignment or failure in the course.

Use of Generative Artificial Intelligence Technology

The use of generative AI technologies, like ChatGPT, is NOT allowed on quizzes and written assignments in this course. Such use will be considered Academic Dishonesty and will result in disciplinary action. You may use generative AI technologies for non-testing activities such as creating summary notes etc.

Intellectual property of materials on the LMS website

Please note that all assignments for this course and all materials posted to the LMS website are the intellectual property of myself and the University of Victoria. Do not circulate this material or post it to note-sharing sites without my permission. Posting course materials to note-sharing sites or otherwise circulating course materials without the permission of your instructor violates the Policy on Academic Integrity. (<http://web.uvic.ca/calendar/FACS/UnIn/UARe/PoAcI.html>). Any evidence you are circulating materials without permission will be referred to the Chair of the Philosophy Department for investigation.

Territory Acknowledgement

The University of Victoria acknowledges with respect the Lekwungen-speaking peoples on whose traditional territory the University of Victoria stands, and the Songhees, Esquimalt and WSÁNEĆ peoples whose historical relationships with the land continue to this day.

Equity

UVic is committed to promoting, providing and protecting a positive, supportive and safe learning and working environment for all its members. Information about UVic policies on human rights, equity, discrimination and harassment are available at <https://www.uvic.ca/calendar/future/undergrad/index.php#/policy/HkQ0pzdAN>
If you have any particular concerns about equity in our course please don't hesitate to contact me.

Religious Accommodation

Information regarding accommodation of religious observance can be located in the UVIC Calendar at <https://www.uvic.ca/calendar/future/undergrad/index.php#/policy/r1q0gofdN?bc=true&bcCurrent=10%20-%20Accommodation%20of%20Religious%20Observance&bcGroup=Undergraduate%20Academic%20Regulations&bcItemType=policies>

“By far, the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.” —Eliezer Yudkowsky

Topic List

	Monday	Thursday
Week 1: Jan 6 th – 10 th	Introduction to the Course, Issues in the Ethics of Artificial Intelligence	Introduction to Ethical Theories (Consequentialism and Deontology)
Week 2: Jan 13 th – 17 th	Can Computers Think? Turing and Searle	Can Computers Think? Moural In-Class Written Assg #1
Week 3: Jan 20 th – 24 th	Robot Rules Asimov Quiz #1 Due	Robot Rules Anderson
Week 4: Jan 27 th – 31 st	Movie: I, Robot and Discussion	Movie: I, Robot and Discussion In-Class Written Assg #2
Week 5: Feb 3 rd – 7 th	Robot Rules Unesco Report Quiz #2 Due	Robot as Moral Agent Sullins
Week 6: Feb 10 th – 14 th	Data and Privacy In-Class Written Assg #3	Guest Speaker Quiz #3 Due
Reading Break		
Week 7: Feb 24 th – 28 th	Test #1	Bias in AI
Week 8: Mar 3 rd – 7 th	AI and Warfare Movie: Killer Robots	AI and Warfare In-Class Written Assg #4
Week 9: Mar 10 th – 14 th	AI and Social Media Movie: Social Dilemma	AI and Social Media In-Class Written Assg #5
Week 10: Mar 17 th – 21 st	Autonomous Vehicles Quiz #4 Due	AI and Healthcare In-Class Written Assg #6
Week 11: Mar 24 th – 28 th	Singularity Chalmers	Autonomy Wheeler
Week 12: Mar 31 st – Apr 4 th	Guest Speaker Quiz #5 Due	Test #2

Links for readings and viewings are available on BrightSpace.

This syllabus is tentative and subject to change.