

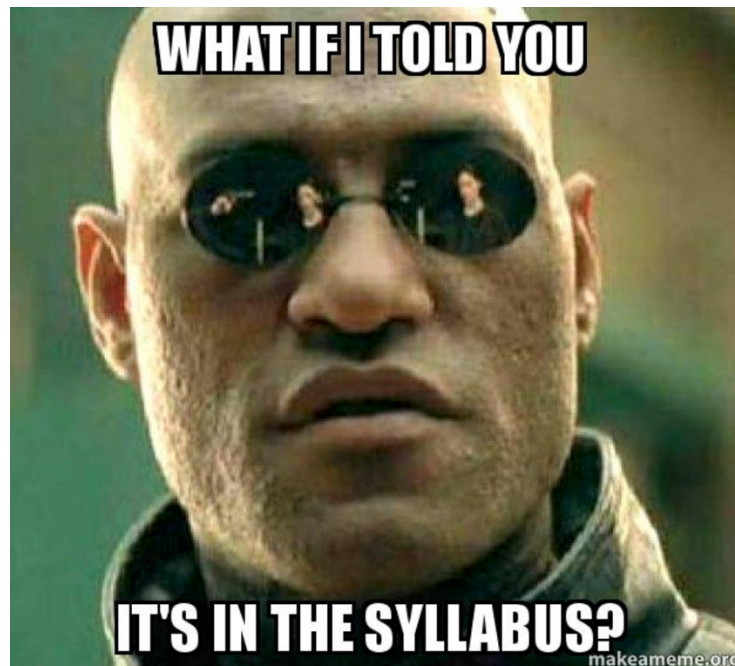
PHI 1680: AI, Philosophy, and Society

Quest 1

Read the syllabus!

Please read this syllabus carefully. Look, I understand that syllabi are tedious to read. Nobody *wants* to read a syllabus (except maybe the stuff about topics, readings, and assignments). However, you really do need to read this document carefully, as you need to know what it says in order to (a) benefit as much as possible from the course and (b) get a good grade. I promise reading it will be well worth your time!

Besides, if you don't read this syllabus carefully, and you email me to ask a question that is answered explicitly in it, I am going to annoy you by responding with this (admittedly pretty dated) meme:



I. General Information

Meeting Times and Location

Lecture

Monday/Wednesday 12:50–13:40 PM in TURL005

Section

This course has a required discussion section (meeting times and locations vary).

Instructor

David Gray Grant, Ph.D.

Office: Griffin-Floyd Hall 301

Office hours: Mondays 2:00–4:00 PM

Email: david.grant@ufl.edu

Teaching Assistant(s)

Nathan Thomas

Email: nathan.thomas1@ufl.edu

Office hours: Wednesdays 9:30 AM–12:30 PM

Office hours location: Griffin-Floyd Hall 316

Joshua Hasker

Office hours: Tuesdays and Thursdays 12:30–2:00 PM

Office hours location: Griffin-Floyd Hall 316

Email: jhasker@ufl.edu

Course Description

In the past few years, the capabilities of AI-based systems have grown explosively due to the development of a new technology, large language models. These systems, known informally as “chatbots,” are trained on a significant portion of the text and images that humanity has collectively produced over centuries. As a result, they have developed the ability to perform tasks that we normally associate with human-level intelligence, such as writing essays, writing computer programs, and passing graduate-level exams. In this course, we will explore the philosophical and social implications of this powerful new technology. Are chatbots intelligent in the same sense that we are intelligent? Will they take jobs previously held by highly skilled human workers, such as lawyers, doctors, and software engineers? Could we fall in love with chatbots, and could they fall in love with us? Why are so many of the experts developing chatbots concerned that they might destroy humanity? What can we do to stop them from destroying humanity? Will we one day be able to “upload” our minds to computer servers, in effect becoming chatbots ourselves?

As we explore these questions, we will engage with research from several academic disciplines, including computer science, psychology, philosophy, and economics. Assignments will focus on original research into existing AI-based technologies as well as critical reflection on how we want AI to shape society going forward.

This course provides both Quest 1 and General Education Humanities credit. As a Quest 1 course, it focuses on multidisciplinary exploration of an essential question about the human condition (how AI is changing society, and how should we, as citizens, respond?) and emphasizes both experiential learning and self-reflection. As a General Education Humanities course, it familiarizes students with analytical tools from several humanities disciplines in order to equip them with the ability to approach this question in a rigorous way and from multiple theoretical perspectives.

Credit hours: 3

Quest and General Education Credit

- Quest 1
- Humanities

This course accomplishes the [Quest](#) and [General Education](#) objectives of the subject areas listed above. A minimum grade of C is required for Quest and General Education credit. Courses intended to satisfy Quest and General Education requirements cannot be taken S-U.

Required Readings and Works

No required materials or fees. All readings and other course materials will be posted to or linked from Canvas.

II. Communication policy

Announcements

Course announcements will be posted on Canvas. You are responsible for checking Canvas and your UF email at least once a week (and ideally every day that UF is in session) to make sure that you do not miss important announcements.

Contacting your TA

Your teaching assistant is your primary point of contact for questions about assignments, grades, attendance, extensions, and other administrative matters for the course. Your TA will let you know the best way to contact them.

Please contact your TA first if you have questions about one of these administrative matters. If you contact me first about assignments, grades, attendance, or other administrative matters, I will almost certainly have to refer you back to your TA, as they:

1. maintain our attendance records;
2. handle excused absences and assignment extensions;
3. grade your assignments and maintain our grading records; and
4. address clarificatory questions about assignments and course policies.

If you are unable to resolve the issue after speaking with your TA, or if your TA does not respond within two business days, or if the matter is genuinely urgent and you are not sure that your TA can help, please feel welcome to reach out to me directly and I will be happy to assist you.

Contacting your instructor

For questions about course content, or for other issues that you are unable to resolve by communicating with your TA, please feel welcome to reach out to me (Dr. Grant) directly by email or by coming to office hours. I'm here to help!

- Email (david.grant@ufl.edu) and office hours are the most reliable way to get in touch with me outside of class.
- **I make every effort to respond to email from students within two (2) business days.** Please do not contact me again if two business days have not passed unless the matter is genuinely urgent.
- Note that emails do sometimes get lost (due to spam filtering, for instance). **Please feel welcome to send me another email or come up to me after class if you do not hear back within two business days.** I promise I am not ignoring you intentionally!

When you email me, please (always) specify both (1) the title or course number of the class your question is about and (2) your TA's name. I regularly teach two entirely different classes with hundreds of students and several TAs, so it will be much easier for me to help you if you provide this information.

Tech support

We are unable to provide tech support for software that you will (or may) be using in this class, such as Canvas, UF email, Perusall, iClicker, Google Drive, etc., as we lack the requisite back-end access to these systems, and are in any case not experts. However, **we do need to know if you are experiencing technical issues** that are interfering with your ability to fully participate in the course.

So, if you are experiencing technical issues, please:

1. Contact the UF Help Desk (<https://it.ufl.edu/helpdesk/>) by phone, chat, or email. (They are available 24/7.)
2. Let your TA know if the issue is interfering with your ability to participate in the class, complete assignments, etc.

III. Graded Work

Description of Graded Work

	Description	Points
Perusall	Participation on Perusall	150
Group exercises	Two group activities focused on experiential learning and reflection	200
Short essays	Two short analytical essays	350
Capstone project	Group presentation exploring a philosophical question raised by the course	150
Participation	Participation in lectures and discussion sections	150

Perusall

Collaborative annotation of weekly readings on the Perusall online collaborative annotation platform. You will be asked to annotate at most one reading per week. Your grade will be determined by the quantity and quality of your participation in these collaborative discussions over the course of the semester.

Perusall annotations are required and will not be accepted late.

- You can skip two Perusall annotation assignments without penalty.
- Each subsequent Perusall annotation assignment that you fail to submit will result in a 10% deduction from your Perusall grade.

You are responsible for doing all assigned readings, regardless of whether there is an associated Perusall assignment.

- Note that while many of the readings for this course are short, they are also very dense—you will need to read them *carefully* and *multiple times* in order to fully absorb the material.

Group exercises

Two small group exercises focused on experimenting with contemporary AI systems or reflecting on the role AI plays (or should play) in our society. The exercises will be graded on the basis of a brief written report. Each member of the group will submit their own version, which will be revised by the group on the basis of discussion.

- Group members will normally receive the same grade, based on the overall quality of the report submitted by the group (see the rubric for written assignments below). However, students that fail to contribute adequately (based on peer evaluations) may be required to re-do the assignment for an individualized grade. Students that repeatedly fail to adequately contribute to group assignments may also lose participation points.
- You will be given time to work on small group exercises during section or lecture. If you are unable to attend the relevant section or lecture, you may be asked to complete the assignment individually.
- Students that do not submit an individual version of the report will receive a zero for the assignment.

Each group exercise will be worth 100 points.

Short essays

Two short, thesis-driven analytical essays (1000-1250 words each) on assigned topics relating to the material read and discussed in class. Each essay will ask you to develop and defend your own position on a question posed by the readings and/or lectures for the course. The first short essay will be worth 150 points; the second will be worth 200.

Capstone project presentation

Students will form small groups and select a philosophical question raised by the course to explore in greater detail. Groups will work together to formulate an answer to their chosen philosophical question, and develop a short presentation articulating and defending that answer. Groups will then deliver their presentations in lecture during the final weeks of class. Groups will normally receive a single

grade for the presentation, except in cases where one or more group members fails to adequately contribute.

Participation

You will receive a letter grade for participation representing the overall **quantity** and **quality** of your participation over the course of the semester. Below is a summary of course standards for different letter grades (borderline cases will receive “+” or “-” grades as appropriate):

- A (100%): Outstanding participation. Highly engaged during lecture and section, with active participation in small group discussions and frequent constructive contributions to full-class discussions. In-class contributions reflect excellent understanding of and critical engagement with the content of lectures and assigned readings.
- B (86%): Good participation. Moderately engaged during lecture and section, with active participation in small group discussions and occasional constructive contributions to full-class discussions. In-class contributions reflect significant understanding of and critical engagement with the content of lectures and assigned readings.
- C (76%): Minimally adequate participation. Minimally engaged during lecture and section. In-class contributions are infrequent and reflect limited understanding of and critical engagement with the content of lectures and assigned readings.
- D (66%): Poor participation. Attentive during lecture and section, but participating only very infrequently in small-group and full-class discussions. In-class contributions do not reflect understanding of or critical engagement with the content of lectures and assigned readings.
- F (0%): Inadequate participation. Not attentive during lecture and section and/or not participating in either small-group or full-class discussions.

Attendance requirement

Lecture attendance

Attendance in lecture is not formally required, and we will not be taking attendance. **However, you will not do well in this course if you do not regularly attend lecture.** There are two reasons for this:

1. Participation in lecture is part of your grade in the course (see “Participation”). Pretty obviously, you cannot participate in lecture (and your TA cannot observe your participating in lecture) if you are not present.
2. The lectures will include crucial information that you will be expected to know when completing course assignments.

Section attendance

Attendance in your weekly discussion section is formally required, and excessive absences will negatively affect your participation grade.

- **You are allowed three (3) unexcused absences from section without penalty.** We do this in part to allow for students to miss section for reasons that, while perfectly legitimate, would not normally qualify for an excused absence (such as travel undertaken for pleasure, rather than to address emergencies or participate in official UF activities).
- **Each unexcused absence beyond the first three will result in a 10% deduction from your total participation grade for the course.**
- Unexcused absences will not be penalized during the first two weeks of section, as we know that enrollment will be in flux.
- If you need to miss your weekly discussion section for any reason, please contact your TA as soon as possible so that you can discuss whether the absence will be excused and what sort of documentation will be necessary (if applicable).
- **If you do not contact your TA about an absence within one week following the absence, it will not be excused except under exceptional circumstances.**
- Poor conduct in lecture or discussion section will also result in a 10% penalty to your participation grade. Examples of poor conduct include arriving late, disruptive behavior, falling asleep, failing to participate in small group activities, sending text messages, checking email, etc. See “Other course policies” below for examples of conduct that may result in a penalty.

Section attendance will be taken using the Canvas “Roll Call” attendance tool. Some notes on how this will work:

- Your TA will use the “Roll Call” tool in Canvas to keep track of *all* of your absences, both excused and unexcused.

- Your TA will maintain a separate list of your excused absences, so that we can calculate your total number of unexcused absences from lecture and section at the end of the semester.
- **The “Attendance” grade that Canvas automatically generates is not directly related to your participation grade in the class**, as discussed below (see “The ‘Attendance’ grade in Canvas”).

Deadlines and extensions

All assignment deadlines will be posted on Canvas.

If you think you may need an extension, please reach out to your TA to explain the situation and see if you qualify.

- We are understanding people, so please reach out to your TA if you think you *might* qualify for an extension, even if you are not sure.
- Note that rates of mental illness among college-age adults are at historically high levels (as we will discuss in the course in connection with our discussion of social media algorithms). In general, mental health issues *will* qualify you for an excused absence, even if you do not have an official diagnosis (though we may require you to reach out to the UF Counseling and Wellness Center for assistance and a doctor’s note).

Verbal Preparation Rule

Whenever you submit work for this course, you must be prepared to present it orally, and to answer questions about it. This will help ensure that you do not submit work that you do not fully understand (either because it was prepared too hastily, or because it was prepared with unauthorized assistance, such as unauthorized generative AI use).

To ensure that you are prepared to orally present each assignment that you submit, we will randomly select several students to orally defend their submissions during their required discussion sections. Oral defenses will have two parts:

1. **Brief presentation.** The student will give a brief presentation outlining the main ideas contained in their submission.

2. **Brief Q&A.** The student will be asked questions about those ideas by their classmates and TA.

During your presentation, you are welcome to have your paper in front of you as a memory aid. You are also welcome to refer to notes as you speak. However, you cannot simply read out a presentation that you prepared beforehand; you should explain your paper informally and largely from memory.

Oral defenses will be graded pass/fail. A passing performance simply means demonstrating that you understand what you wrote—that you can explain your argument in your own words and engage thoughtfully with questions. I expect the only failing performances will be cases where the presenter is unable to explain the ideas expressed in their own paper (which should be straightforward if you did the work yourself).

- Since being prepared for a possible oral defense is part of the assignment itself, and since the defense is graded pass/fail, students that pass will not receive extra credit (though they will benefit from the additional feedback that they receive during the defense).
- Students that fail the defense may be required to complete a new, substitute assignment in a controlled environment and without external aids.

If you suspect that defending your submissions in front of your fellow students would cause you significant anxiety, please let us know at the beginning of the semester, and we will allow you to conduct your defenses (if any) during office hours.

Final grades

Grading scale

The following grade scale will be used to assign final letter grades for the course.

Points	Grade
940 – 1000	A
900 – < 940	A-
870 – < 900	B+

840 – < 870	B
800 – < 840	B-
770 – < 800	C+
740 – < 770	C
700 – < 740	C-
670 – < 700	D+
640 – < 670	D
600 – < 640	D-
0 – < 600	E

For information on how UF assigns grade points, visit:

<https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/>

No rounding

Final grades will not be rounded (so please don't ask). This means that, to receive a given grade, you must meet or exceed the lower threshold specified; for instance, to get an 'A' you must score at least 940 points. Neither points nor grades will be rounded.

The Canvas Gradebook

We will report grades and other grade-related information to you on Canvas in two ways:

1. assignment grades will be posted on Canvas and viewable in the Canvas gradebook for our course; and
2. attendance records for the required discussion sections will be maintained using the Canvas Roll Call tool.

Note that Canvas automatically reports some misleading information about grades, as discussed below.

The “Total” column in Canvas

Ignore the “Total” grade calculated by Canvas. During the semester, the “Total” column in Canvas is not a reliable guide to how well you are doing in the course, as it does not take into account your performance on assignments that have not yet been graded. Since many students misinterpret the Total column, I

have hidden it on our Canvas gradebook. If you would like to determine how your final grade will depend on your performance on future assignments, you can determine that using a grade calculator or (even better) basic algebra. I encourage you to learn how to do this sort of calculation in an Excel spreadsheet, as you will find that to be a useful skill going forward.

The “Attendance” grade in Canvas

The “Attendance” grade calculated by Canvas is not directly related to your grade in the course. Your TA will use the Canvas Attendance tool to track your attendance in our required discussion sections. (This will also help your TA get to know your name.) Canvas automatically calculates an “Attendance” grade on the basis of the resulting attendance records. However, that grade is not used to calculate your final grade in the course, and is not directly related to your participation grade. Instead, your TA will use your total number of unexcused absences as just one data point in determining how actively (and how successfully) you participated in the course.

You should view the Attendance grade reported by Canvas as a rough (and not necessarily reliable) summary of how many total absences you have accrued so far—both excused and unexcused.

- Please see the “Participation” subsection above for more information about how attendance and participation will affect your final grade in the course.
- Please feel welcome to reach out to your TA if you have any questions about how many unexcused absences you have accrued, their present assessment of the quality and quantity of your participation in the course so far, etc.

Grading Rubrics

Writing Rubric (applies to all written work)

	Excellent	Good	Fair	Poor
Content	Fully addresses all elements of the prompt. Demonstrates excellent understanding of and critical engagement with course material. Uses appropriate details and citations from the text(s). (5-6 points)	Addresses all elements of the prompt. Demonstrates good understanding; minor points may be misstated or omitted. Demonstrates good critical engagement with course material. Uses appropriate details and citations from the text(s). (3-4 points)	Fails to address all elements of the prompt. Significant misunderstandings of course material. Key ideas are explained incorrectly or omitted. Limited critical engagement with course material. Little textual evidence to back up interpretation of the readings. (1-2 points)	Fails to address multiple elements of the prompt. Very substantial misunderstandings. Key ideas are explained incorrectly or omitted. Very limited evidence of critical engagement with course material. Textual evidence is missing or not relevant. (0 points)
Writing	Written in simple, non-technical language. Develops a small number of points thoroughly. Well-organized and easy to follow. Few, if any, spelling or grammatical errors that interfere with understanding. (4 points)	Mostly clear, focused, and organized, but includes some confusing passages, digressions/repetitive passages, or lapses in organization. Some spelling or grammatical errors that interfere with understanding. (3 points)	Significant lapses in clarity, focus, or organization. Significant spelling or grammatical errors that interfere with understanding. (1-2 points)	Very difficult to understand; many spelling or grammatical errors that interfere substantially with understanding. (0 points)

10 points	A+ (100%)
9 points	A (96%)
8 points	A- (93%)
7 points	B+ (89%)
6 points	B (86%)
5 points	B- (83%)
4 points	C+ (79%)
3 points	C (76%)
2 points	C- (73%)
1 points	D (66%)
0 points	F (60%)

Presentation Rubric (applies to the capstone project presentation)

	Excellent	Good	Fair	Poor
Content	Fully addresses all elements of the prompt. Demonstrates excellent understanding of and critical engagement with course material (i.e., ideas from assigned readings that were discussed in lecture). Uses appropriate details and citations from course readings. (5-6 points)	Addresses all elements of the prompt. Demonstrates good understanding of course material with minor errors or omissions. Demonstrates substantial critical engagement with course material. Uses appropriate details and citations from course readings. Properly cites all sources. (3-4 points)	Fails to address all elements of the prompt. Significant misunderstandings of course material. Key ideas are explained incorrectly or omitted. Limited critical engagement with course material. Little textual evidence to back up interpretation of course readings. (1-2 points)	Fails to address multiple elements of the prompt. Very substantial misunderstandings. Key ideas are explained incorrectly or omitted. Very limited evidence of critical engagement with course material. Textual evidence is missing or not relevant. Fails to properly cite sources. (0 points)
Presentation	Ideas are presented in simple, non-technical language. Develops a small number of points thoroughly. Well-organized and easy to follow. Slides and presentation are polished, engaging, and effective. (4 points)	Mostly clear, focused, organized, and polished. Brief segments of the presentation are confusing, repetitive, or off-topic. Slides and presentation are engaging and effective. (3 points)	Significant lapses in clarity, focus, or organization. Slides or presentation are poorly constructed or otherwise ineffective. (1-2 points)	Very difficult to understand. Slides or presentation are absent or incomplete. (0 points)

10 points	A+ (100%)
9 points	A (96%)
8 points	A- (93%)
7 points	B+ (89%)
6 points	B (86%)
5 points	B- (83%)
4 points	C+ (79%)
3 points	C (76%)
2 points	C- (73%)
1 points	D (66%)
0 points	F (60%)

IV. Other course policies

Classroom conduct

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to loss of participation points or other penalties (as appropriate).

In order to avoid unnecessary disruptions to the learning environment, you are expected to conform to the following policies (in addition to other commonsense and generally accepted norms of classroom conduct that are not explicitly identified below). First-time violations of these policies will receive a verbal warning. Subsequent violations will yield a loss of participation points.

Side conversations, etc.

Side conversations (including whispered ones) and other forms of disruptive behavior that make it difficult for others to concentrate are not permitted.

- While lecture is in progress, you are expected to remain silent unless you have been called on to speak by the lecturer or are participating in a small group activity.
- Note that this is the expectation in most formal settings with large groups (for obvious reasons).

Timely arrival and departure

You are expected to:

1. **Arrive on time.** Please be seated and ready for lecture to begin by the official start time specified above. This includes checking yourself in with iClicker (if we are using it) and putting your devices away *before* the official start time for the class.
2. **Remain seated and engaged until class officially ends.** Please remain seated and engaged with lecture until the official end time specified above. Please do feel free (of course) to leave the room to go to the bathroom, address genuine emergencies, etc.

3. **Refrain from packing up early.** Avoiding disruptions includes refraining from packing up your belongings until the official end time specified above or when the instructor announces that the lecture is over (whichever comes first). If you need to leave class early, please let us know (in advance if possible).

We understand that students sometimes need to arrive late or leave early for legitimate reasons. If you need to leave class early, please contact your TA as soon as possible to explain.

Laptop and Cellphone Policy

Laptop and cellphone use is prohibited in this class, in part because it has been shown to negatively affect learning outcomes for *other* students, and not just the student using the device.

- You may use a tablet if it is lying flat and opened to a note-taking app or course reading.
- Please let us know as soon as possible if you have (or think you may need) an accommodation letter from the University permitting you to use a laptop to take notes.

Accommodations for students with disabilities

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center. [Click here to get started with the Disability Resource Center.](#) **It is important for students with accommodations to share their accommodation letter with their instructor and discuss their access needs as early as possible in the semester.**

Plagiarism

Acknowledgement of sources

Per the UF Honor Code and Student Conduct Code, “[s]ubmitting materials from any source without proper attribution” constitutes plagiarism and carries potentially serious penalties. **To avoid plagiarism, all sources and assistance used in**

preparing your assignments must be precisely and explicitly acknowledged.

- The web creates special risks here. Cutting and pasting even a few words from a web page or paraphrasing material without a reference constitutes plagiarism. If you are not sure how to refer to something you find on the internet, you can always give the URL.
- You can find a copy of the UF Honor Code and Student Conduct Code here: <https://policy.ufl.edu/regulation/4-040/>.

Unauthorized generative AI use

Students are not permitted to use large language models (such as ChatGPT) or other AI-based text generation tools to complete written assignments for this course unless specifically authorized to do so by the relevant assignment prompt. Submitting work produced using one of these tools without prior written authorization (such as explicit authorization from the assignment prompt) constitutes plagiarism.

I'm sure that none of you would do such a thing intentionally, in part because you want to learn the material we will cover, and in part because you want to develop the (useful and marketable) skills that you will develop by doing the work yourself. However, I have heard that some students believe that unauthorized generative AI use is not detectable, and so will not result in adverse consequences.

While it is of course difficult to prove that a student has used generative AI in an unauthorized way, students that do so are *routinely* found guilty of violating other provisions of the honor code by the Student Conduct and Conflict Resolution Committee. The reason for this is that chatbots have a well-known tendency to both:

1. plagiarize other authors without attribution; and
2. hallucinate sources and citations.

These are much easier to prove, and either is sufficient for a guilty finding for plagiarism as defined by UF. So, please don't do it—for your sake, for our sake, and for the sake of your fellow students! If you have any questions about what constitute unauthorized use of sources, plagiarism, etc., please let us know and we will be happy to discuss it. (You should of course review

Procedure for suspected cases

Students that the teaching team suspects have engaged in plagiarism may be required to undergo an oral defense of their submission, and may be reported to the Student Conduct and Conflict Resolution Committee (SCCR).

In the event that the SCCR finds that you are guilty of plagiarism or other forms of academic misconduct prohibited by the UF Honor Code and Student Conduct Code, you should expect the minimum penalty to be a grade of E for the class (not just the assignment).

Perusall privacy and accessibility policies

You can find Perusall's privacy policy [here](#) and their accessibility statement [here](#).

V. Student Learning Outcomes (SLOs)

At the end of this course, students will be expected to have achieved the [Quest](#) and [General Education](#) learning outcomes as follows:

Content: Students demonstrate competence in the terminology, concepts, theories and methodologies used within the discipline(s).

- Identify, describe, and explain how the resources available in the humanities and sciences can help us understand artificial intelligence's present and future effects on society (Content SLOs for Gen Ed Humanities and Q1). Assessed by Perusall, Group Exercises, Short Essays, Capstone Project, and Participation.
- Identify, describe, and explain how the resources available in the humanities and sciences can help us reflect on how we *should* want artificial intelligence to affect society (Content SLOs for Gen Ed Humanities and Q1). Assessed by Perusall, Group Exercises, Short Essays, Capstone Project, and Participation.

Critical Thinking: Students carefully and logically analyze information from multiple perspectives and develop reasoned solutions to problems within the discipline(s).

- Identify, analyze, and evaluate themes in public and scholarly discussions of artificial intelligence and its effects on society (Critical Thinking SLOs for Gen Ed Humanities and Q1). Assessed by Perusall, Group Exercises, Short Essays, Capstone Project, and Participation.
- Analyze and evaluate the particular questions about artificial intelligence and its effects on society that we discuss in the course, such as questions about how artificial intelligence will affect our work and romantic lives (Critical Thinking SLOs for Gen Ed Humanities and Q1). Perusall, Group Exercises, Short Essays, Capstone Project, and Participation.
- **Communication:** Students communicate knowledge, ideas and reasoning clearly and effectively in written and oral forms appropriate to the discipline(s). Develop and present clear and effective responses to essential questions about how artificial intelligence is, will, and should affect society in oral and written forms appropriate to the relevant humanities disciplines incorporated into the course (Communication SLOs for Gen Ed Humanities and Q1). Assessed by Perusall, Group Exercises, Short Essays, Capstone Project, and Participation.

Connection: Students connect course content with meaningful critical reflection on their intellectual, personal, and professional development at UF and beyond.

- Reflect on how artificial intelligence is affecting us and our communities (Connection SLO for Q1). Assessed by Perusall, Group Exercises, Capstone Project, Participation.
- Reflect on what we value, and how those values should shape the way we interact with AI-based technologies and the way we respond to AI's social effects (Connection SLO for Q1). Assessed by Perusal, Group Exercises, Capstone Project, and Participation.

VI. Quest Learning Experiences

1. Details of Experiential Learning Component

Experiential learning is an important part of this course and will be incorporated into several assignments, including the group exercises, short analytical essays, and capstone project. These assignments will involve experimenting with large language models and conducting original philosophical research.

2. Details of Self-Reflection Component

Throughout the course, you will have the opportunity to reflect on how AI-based technologies have affected your own life and the lives of those around you, whether you believe those effects are good or bad (and why), and how we should address the social changes that AI is already bringing about and may bring about in the future. Time for self-reflection will be integrated into the lectures and group exercises.

VII. University policies

This course complies with all UF academic policies. For information on those policies and for resources for students, please see [this link](#).

VIII. Schedule and readings

The following schedule is tentative and subject to change. Please see Canvas for an up-to-date schedule and list of readings.

Unit 1: The present

Mon 01/12 · L1 · What is artificial intelligence?

- The Ezra Klein Show, "A.I. could solve some of humanity's hardest problems. It already has" (podcast)

Wed 01/14 · L2 · AI safety

- OpenAI, "Detecting and reducing scheming in AI models" (<https://openai.com/index/detecting-and-reducing-scheming-in-ai-models/>)
- Amodei et al, "Concrete Problems in AI Safety," sections 1-4 (optional; skim the technical bits and just focus on the more general descriptions of the problems the authors are characterizing)

Mon 01/19 · No class (holiday)

- No readings

Wed 01/21 · L3 · Are chatbots intelligent? (part 1)

- Turing, "Computing machinery and intelligence," sections 1-3 (read carefully), section 4 (just skim this one to get the gist), section 6

Mon 01/26 · L4 · Are chatbots intelligent? (part 2)

- Bubeck et al., "Sparks of Artificial General Intelligence: Early experiments with GPT-4" (sections 1 and 10, plus whatever strikes your interest)
- Marcus and Davis, "GPT-3, Bloviator: OpenAI's Language Generator Has No Idea What It's Talking About"

Wed 01/28 · L5 · Group exercise workshop

- No new readings

Mon 02/02 · L6 · AI and decision-making: privacy

- King and Meinhardt, "Rethinking Privacy in the AI Era," ch. 3 ("Provocations and Predictions")

Wed 02/04 · L7 · AI and decision-making: transparency

- Christian, *The Alignment Problem*, ch. 3 ("Transparency")

Mon 02/09 · L8 · AI and decision-making: fairness

- Barocas, Hardt, and Narayanan, *Fairness and machine learning: Limitations and Opportunities*, ch. 1 ("Introduction")
(<https://fairmlbook.org/introduction.html>)

Wed 02/11 · L9 · Paper writing workshop

- Pryor, "Guidelines on Writing a Philosophy Paper"
- Group exercise 1 due Friday

Mon 02/16 · L10 · Unit review

- No new readings

Unit 2: The near future

Wed 02/18 · L11 · Autonomous vehicles

- Hao, "Should a Self-Driving Car Kill the Baby or the Grandma? Depends on Where You're From"
- Jaques, "Why the Moral Machine is a Monster" (pages 1–9)

Mon 02/23 · L12 · Lethal autonomous weapons (part 1)

- Chivers, "The Dawn of the A.I. Drone" (NYT)

Wed 02/25 · L13 · Lethal autonomous weapons (part 2)

- Sparrow, "Killer robots"

Mon 03/02 · L14 · Can we fall in love with chatbots?

- Kraft, "They Fell in Love With A.I. Chatbots — and Found Something Real" (NYT Magazine)
- Frank and Nyholm, "From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible?" (pages 219–238)

Wed 03/04 · L15 · AI and the future of work

- Brynjolfsson et al., "Canaries in the Coal Mine: Six Facts about the Recent Employment Effects of Artificial Intelligence" (skim to get a sense of the main results)
- Gheaus and Herzog, "The goods of work (other than money!)"
- Short essay 1 due Friday

Mon 03/09 · L16 · Unit review

- No new readings

Unit 3: The far future

Wed 03/11 · L17 · The singularity hypothesis

- Chalmers, "The singularity: a philosophical analysis," sections 1 and 2
- Group exercise 2 due Friday

Mon 03/16 · No class (spring break)

- No readings

Wed 03/18 · No class (spring break)

- No readings

Mon 03/23 · L18 · AI and human achievement

- Karlan, "Human achievement and artificial intelligence"

Wed 03/25 · L19 · The control problem

- Chalmers, "The singularity: a philosophical analysis," sections 5–7

Mon 03/30 · L20 · Uploading

- Chalmers, "The singularity: a philosophical analysis," section 10

Unit 4: Final project

Wed 04/01 · L21 · Final project workshop

- No new readings

Mon 04/06 · L22 · Project presentations

- No new readings

Wed 04/08 · L23 · Project presentations

- No new readings

Mon 04/13 · L24 · Project presentations

- No new readings

Wed 04/15 · L25 · Project presentations

- No new readings

Mon 04/20 · L26 · Project presentations

- No new readings

Wed 04/22 · L27 · Conclusion

- No new readings
- Short essay 2 due