IDS 2935: AI, Philosophy, and Society
Quest 1

I. General Information

Meeting Times and Location

- Monday/Wednesday 11:45 AM-12:35 PM
- LIT0101
- This course has a required discussion section (meeting times vary).

Instructor

- David Gray Grant
- Office: Griffin-Floyd Hall 301
- Office hours: Monday/Wednesday 1:30-2:30 PM
- Email: david.grant@ufl.edu

Teaching Assistant(s)

- TBA on Canvas

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.

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 Canvas.

II. Graded Work

Description of Graded Work

	Description	Points
Group exercises	Two group activities focused on experiential learning and reflection	200
Exams	Two in-class exams	400
Capstone project	Group project exploring a philosophical question raised by the course	150
Participation	Participation in lecture and section	250

Group exercise reports (2 x 100 = 200 points)

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 completed in section and graded on the basis of a brief written report. Each member of the group will submit their own version, which will be revised on the basis of discussion; the group will then select which version to submit for grading. 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 do not submit an individual version of the report will receive a zero for the assignment.

Exams (1 x 200 + 1 x 250 = 400 points)

Two in-class exams testing your understanding of course material.

Capstone project (150 points)

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 last two 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 (250 points)

A significant portion of your grade will be based on the overall **quantity** and **quality** of your participation in lecture and section. 10% of your grade will be based on your participation during the first half of the semester, and 15% will be based on your participation during the second half 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.

Grading Scale

For information on how UF assigns grade points, visit: https://catalog.ufl.edu/UGRD/academic-regulations/grades-grading-policies/

A	94 – 100%		C	74 – 76%
A-	90 – 93%		C-	70 – 73%
B+	87 – 89%		D+	67 – 69%
B	84 – 86%		D	64 – 66%
B-	80 – 83%		D-	60 – 63%
C+	77 – 79%		E	<60

Grading Rubrics

Writing Rubric

(Applies to all written work)

	Excellent	Good	Fair	Poor
Content	Fully addresses all aspects 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)	Fully addresses all aspects 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)	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)	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. (4 points)	Mostly clear, focused, and organized, but includes some confusing passages, digressions/repetitive passages, or lapses in organization. Some spelling or grammatical errors. (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. (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%)

Project Presentation Rubric

(Applies to the in-person presentation component of the capstone project)

	Excellent	Good	Fair	Poor
Content	Fully addresses all aspects of the prompt. Demonstrates excellent understanding of and critical engagement with course material. (5-6 points)	Fully addresses all aspects of the prompt. Demonstrates good understanding of course material with minor errors or omissions. Demonstrates substantial critical engagement with course material. (3-4 points)	Significant misunderstandings of course material. Key ideas are explained incorrectly or omitted. Limited critical engagement with course material. (1-2 points)	Very substantial misunderstandings. Key ideas are explained incorrectly or omitted. Very limited evidence of critical engagement with course material. (0 points)
Presentation	Ideas are presented in simple, non-technical language. Develops a small number of points thoroughly. Well-organized and easy to follow. Polished and engaging. Presentational materials (e.g., slides) are polished and effective. (4 points)	Mostly clear, focused, organized, and polished. Brief segments of the presentation are confusing, repetitive, or off-topic. Uses effective presentational materials. (3 points)	Significant lapses in clarity, focus, or organization. Presentational materials are ineffective or poorly constructed. (1-2 points)	Very difficult to understand. Presentational materials 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%)

III. Tentative schedule and readings

Please note that the following is tentative and subject to change.

Please see Canvas for an up-to-date schedule and list of readings.

Group exercise deadlines will be announced on Canvas.

Unit 1: Can machines be intelligent?

Mon 08/25 · L1 · What are the machines up to?

- No readings

Wed 08/27 · L2 · How smart are machines getting?

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

Mon 09/01 · Holiday (no class)

Wed 09/03 · L3 · How do we tell if a machine is smart?

- Turing, "Computing machinery and intelligence"

Mon 09/08 · L4 · How do we tell if a machine is smart?

- No new readings

Wed 09/10 · L5 · Are machines already smart?

- Bubeck et al., "Sparks of Artificial General Intelligence: Early experiments with GPT-4"

Mon 09/15 · L6 · Maybe machines aren't so smart?

- Marcus and Davis, "GPT-3, Bloviator: OpenAI's Language Generator Has No Idea What It's Talking About"
- Hofstadter, "The Shallowness of Google Translate"

Unit 2: Will machines replace us?

Wed 09/17 · L7 · Will machines take our jobs?

- DePillis and Lohr, "Tinkering with ChatGPT, workers wonder: Will this take my job?"
- Gheaus and Herzog, "The goods of work (other than money!)"

Mon 09/22 · L8 · If machines do it better, why bother?

- Karlan, "Human achievement and artificial intelligence"

Wed 09/24 · L9 · Can we fall in love with machines?

- HiPhi Nation, "Love in the time of Replika" (podcast)
- Frank and Nyholm, "From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible?" (pages 219–238)

Mon 09/29 · L10 · Should machines kill?

- Satariano et al., "Killer Robots Aren't Science Fiction. A Push to Ban Them Is Growing."

Wed 10/01 · L11 · Exam 1

- No new readings

Unit 3: Can machines be moral?

Mon 10/06 · L12 · Who should we blame when machines kill?

- Sparrow, "Killer robots"

Wed 10/08 · L13 · Why don't machines do what we want?

- Amodei et al, "Concrete Problems in AI Safety"

Mon 10/13 · L14 · What should we want machines to do?

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

Wed 10/15 · L15 · Are machines less biased than humans?

- Fazelpour and Danks, "Algorithmic Bias: Senses, Sources, Solutions"

Mon 10/20 · L16 · Why aren't machines less biased than humans?

- No new readings

Wed 10/22 · L17 · Can we make moral machines?

- Anderson and Anderson, "Creating an Ethical Intelligent Agent"
- Hard Fork, "Dario Amodei, C.E.O. of Anthropic, on the Parodoxes of A.I. Safety" (podcast) (Amodei interview only)

Unit 4: Will machines surpass us?

Mon 10/27 · L18 · Will machines get smarter than us?

- Chalmers, "The singularity: a philosophical analysis" (pp. 1-22)

Wed 10/29 · L19 · Could we control superintelligent machines?

- Chalmers, "The singularity: a philosophical analysis" (pp. 22-34)

Mon 11/03 · L20 · Can we become machines?

- Chalmers, "The singularity: a philosophical analysis" (34-54)

Wed 11/05 · L21 · Final project orientation

- No new readings

Mon 11/10 · L22 · Exam 2

- No new readings

Wed 11/12 · L23 · Guest lecture

- No new readings

Unit 5: Final project

Mon 11/17 · L24 · Project presentations

- No new readings

Wed 11/19 · L25 · Project presentations

- No new readings

Mon 11/24 · Holiday (no class)

Wed 11/26 · Holiday (no class)

Mon 12/01 · L26 · Project presentations

- No new readings

Wed 12/03 · L27 · Project presentations

- No new readings

IV. 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 Group Exercises, Exams, 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 Group Exercises, Exams, 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 Group Exercises, Exams, 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). Assessed by Group Exercises, Exams, 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 Group Exercises, Exams, 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 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 Group Exercises, Capstone Project, and Participation.

V. 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 nearly every lecture, and will also be incorporated into the group exercises.

VI. Required Policies

This course complies with all UF academic policies. For information on those polices and for resources for students, please see this link.