

## **AI, Science Fiction, and Human Value**

### **PHIL 4444 (EP4 Value Seminar)**

Summer 2018 Session II

MTWR 1-4 PM

Classroom: Leon Lowenstein Bldg. 518

Professor: Corey Steiner

[csteiner8@fordham.edu](mailto:csteiner8@fordham.edu)

Office Hours: MW 4:15 PM – 5:45 PM & by appointment

### **Course Description and Objectives**

This course is focused on the concept of human value. We will be asking what, if anything, makes a human being a bearer of moral value in a way that is distinct from other forms of life and, more specifically, technologically created models or imitations of life. We will survey a number of traditional and contemporary philosophical positions on the difference between humans and artificial life forms, and thus will engage with these thinkers' views on what is unique to human life. We will also use Heidegger's critique of technology to ask what value can possibly be taken away from human life in the process of technological advancement. Alongside our exploration of these theories, we will turn to several fictional portrayals of interactions between humans and artificial intelligence in order to get a better sense of what value-bearing qualities and capacities are missing on the part of technologically-developed beings. Science fiction can serve as a gateway to philosophical questions about what makes humans distinct because it offers imaginative depictions of the impact technology can have on human experience. Science fiction authors and directors encourage us to stretch our conception of what makes humans unique and what is distinctive about our relationship with technology.

### **Required Texts**

Asimov, Isaac. *I, Robot*. (New York: Spectra Publishing, 2008). ISBN: 055338256X

Dick, Philip K. *Do Androids Dream of Electric Sheep?* (New York: Ballantine Books, A Del Rey Book, 1968). ISBN: 0345404475

Lem, Stanislaw. *Solaris*, Joanna Kilmartin, trans. (Mariner, 2002). ISBN: 0156027607

Additional reading materials on Blackboard (<http://fordham.blackboard.com/>)

### **Capstone Statement**

This course provides students with a robust set of conceptual tools for discussing human values depicted in science fiction, philosophy, and appraisals of AI. It provides a capstone experience in several respects. First, it brings together in a thematic way many diverse elements of Fordham's core curriculum. Its discussion of value and morality speaks to a number of Jesuit principles, such as *cura personalis*, that is, caring for the individual. Throughout the course, students are encouraged to reflect upon the material's implications for their own lives, interactions with one another, and with the technological landscape of contemporary society. The oral and written assignments described on the syllabus are designed to challenge students to synthesize various elements of the core curriculum with an eye to deepening the ways they discuss, understand, and evaluate technology, the possibilities created by AI, and the distinctive character of human value. Through those same assignments, the course achieves the core curriculum's goal of *Eloquentia Perfecta*; students who successfully complete the course will have strengthened their capacity to understand, formulate, and evaluate arguments,

and to discuss challenging ideas both verbally and in writing. Moreover, the course's emphasis on humanity builds upon Fordham's mission of social justice. By engaging with the topic of what makes humans valuable, students are asked to think with a view towards improving the human community and respecting the dignity which is inherent in each human individual. In addition to these overarching goals, the course provides opportunity for interdisciplinary reflection on the material's implications for various fields, including computer science, information technology, neuroscience and biology. It also explores the implications of AI and technology for human value through the medium of literature and film in the works of, for example, Lem, Asimov, Dick and Kubrick. Finally, it brings together and more deeply explores themes introduced to all students at Fordham University in their core philosophy courses, namely, moral theories discussed in *Philosophical Ethics* and questions about humanity in *Philosophy of Human Nature*.

### **Grade Breakdown**

Participation: 15%

Debates: 10%

Journal: 15%

First Paper: 15%

Presentations: 20% (2 x 10%)

Final Paper (Two Drafts + Peer Review): 25%

### **Requirements and Grading**

**1. Attendance.** You are expected to come to every class session. You will be counted as present only if you are in class at the time when it is scheduled to begin, **promptly at 1:00**. If a class begins with a student presentation and you are late to class, you are expected to wait in the hallway until the presentation is completed. If this happens to you more than once over the course of the semester, you will receive an unexcused absence for each time you are late to class (not including the first one).

Legitimate reasons to seek an excused absence are serious illness, religious holidays, and outside academic obligations. An absence (or lateness) counts as excused only if I approve it *prior to class*, if you provide me with a doctor's note, or if your Dean instructs me to excuse it. An excused absence does not affect your grade, though you are responsible for learning missed material. Unexcused absences will have a dramatic effect on your overall grade. Each meeting is roughly equivalent to a week of classes during the regular academic year. Per Fordham Rose Hill policy, upon the third unexcused absence the absent student will necessarily fail the course.

**2. Participation (15%).** Classroom participation will account for half of your participation grade. Participation in classroom discussion is mandatory and will be an essential way for you to engage with course material. I expect everyone to come to class with something to say about the assigned readings. You may want to refer to questions/posts from the online forum (see below) in classroom discussion. You are also expected to listen and respond to other students' questions and contributions. If I notice that you are silent, I will warmly invite you to join the discussion by asking you to read out loud a portion of the text, to share your personal reactions towards the text, and/or to answer a simple question, none of which are meant to be intimidating. If this still may be an issue to you, please come see me and we will instead arrange regular meetings outside of class during which you will share your thoughts on the material in lieu of classroom participation.

The other half of your participation grade will be derived from your contributions to the online discussion forum posted on Blackboard. Each student is expected to post in the online discussion board **at least three times each week**. (For the purposes of this assignment, the week will begin anew on Saturdays,

so a post on July 7, for example, would count towards your posts for the following week.) Contributions may come in any of the following forms:

- (1) Responding to an open-ended discussion question, which I will post roughly once a week.
- (2) Posting a question about the assigned material. Questions should be about particularly confusing passages from the readings, concepts from the course material that you find difficult, concerns about the implications or the legitimacy of a philosopher's claim, a view discussed in class, and/or of the philosophical implications of the science fiction material we are working with. Questions should not be about factual curiosities which lie outside the text, and questions about fictional works should pertain to the themes of the course. Put differently, they should not be answerable with a simple Google search. For example, neither the question "What machines from his era would Descartes have been referring to in the *Meditations*?" nor "How did Kubrick design the sets for *2001*?" would be appropriate for the forum, though both may be valuable in their own right.
- (3) Responding to another student's question. To do this, you don't need to definitively *know* the answer, but rather can use it as an opportunity to delve further into the nuances of the issue or the implications of your peer's question.

Posts should be more than just a few lines, as they should explain the thought process which lies behind your question/response, and why you think it is important. Responses which directly engage with others' posts *and* with course material will be given extra weight. A question which prompts significant discussion will also be given extra weight. You're strongly encouraged to bring up unanswered or unresolved questions from the forum during class discussion.

**3. Debates (10%).** There will be a total of three debates. Students are expected to prepare for debates by considering various possible arguments for both their own position and the opposition, and by closely engaging with course texts in considering authors' views on the issue. **Each student is required to participate in each debate.** I will consider the quality of your contributions when assigning a grade using the following criteria: (i) relevance; (ii) responsiveness to opponents; (iii) engagement with course material.

**4. Journal (15%).** Journals are informal writing assignments in which students discuss their encounters with the topics covered in the class – their thoughts, feelings, comments, and questions about the various theories we are covering and the implications of those theories for their own lives and for the world we live in. Journal entries are not required to follow any specific format. They are graded in terms of both quantity and quality: both how much effort students put into their reflections on the course material, and the quality of that effort. Students should make at least ten journal entries on material covered in the readings, lectures, or discussions (the length of a journal entry should be approximately 1 page, though it may vary depending on your approach to the assignment). Among those entries there should be at least one devoted to each of the following topics:

- The difference between human and non-human rights (Kant) and/or the legal battle surrounding Data in Star Trek's "The Measure of a Man"
- The relationship between humanity and AI as portrayed in *Androids*, *Blade Runner*, *I, Robot*, and/or *2001*
- Heidegger's stance on the relationship between humanity and technology
- The implications of care ethics for human-AI relations

Students are encouraged to get feedback on their journals around the halfway point of the semester to check their progress, and ask for suggested improvements.

**5. First Paper (15%).** A four-page paper is due on Thursday, July 19. In this assignment, you will discuss and respond to one of the topics covered in the course readings and class discussions. A prompt will be distributed roughly one week before the assignment is due.

The assignment should be double spaced, using size twelve Times New Roman Font and standard margins. Send me an electronic copy of your paper **as a Word document** on the days noted on the syllabus. Improper submission format will result in a grade reduction. Late assignments which have not received an extension will be reduced by 1/3 of a letter grade per day. Extension requests made less than 24 hours before the due date will not be considered.

**6. Presentations (20%: 2 x 10%).** Throughout the semester, students in groups of two will give 20 minute presentations which focus on a present-day situation, discussion or fictional example which involves artificial intelligence and analyze it in the context of the philosophical positions discussed in class. Presenters should spend 8-10 minutes discussing the topic and should allot 10-12 minutes for in-class discussion. Presenters should *take a stance* on the situation (what should be done about it, whether the parties' actions are moral or immoral, etc.) and defend it before the class. Students will be evaluated on their presentations by their peers as well as myself. Each group will present twice during the semester, and each presentation will be worth 10% of your overall grade. A prompt, grading rubric, and sign-up sheet will be distributed on the second day of class.

**7. Final Paper (25%).** Students will write an eight-page final paper which engages with the various theories discussed in class and in which they will take a critical stance on the content covered in the course. The final paper is an opportunity for students to showcase the critical perspective and argumentation they have developed throughout the semester by engaging with various counter-arguments and objections. A first draft of the final paper is due on **Monday, July 30**. Students will exchange their drafts with peers and will discuss them during a peer review session on **Thursday, August 2**. I will also be available for individual meetings with students to discuss their drafts and suggestions for improvement. Final papers are due on the class' final exam date, which is scheduled for **Monday, August 6 at 1:00 pm**. There will be a symposium on this date, during which students will present and engage with one another's arguments. More information on the final exam will be given during the third week of class.

### **Academic Honesty**

Plagiarism and other forms of academic dishonesty will not be tolerated. Plagiarism, in general, consists of appropriating another person's thoughts or ideas as your own. Remember, there is nothing shameful about appropriately using the ideas or words of someone else as long as you cite the original source. When in doubt, please cite your source; it is always better to **play it safe** in borderline scenarios. If you are confused about the difference between plagiarism and proper citation, please meet with me outside of class. **All plagiarized assignments will receive a failing grade (0 out of 100) and will be reported to the university.**

### **Electronic Device Policy**

**You may not use electronic devices** in class without my permission. If I see you using an electronic device (e.g., texting, using a computer, etc.), you will lose participation credit for that class. Punishment will increase for each repeated instance.

### **Disabilities**

If you are a student with a documented disability and require academic accommodations, you need to register with the Office of Disability Services for Students (ODS) in order to request academic accommodations for your courses. Please contact the main ODS at Rose Hill at 718-817-0655 to arrange services. Staff at ODS can walk you through the process and arrange appointments depending on which campus you take courses at. Accommodations are not retroactive, so you need to register with ODS prior

to receiving your accommodations. Please see me after class or during office hours if you have questions or would like to submit your academic accommodation letter to me.

## Schedule of Readings

Students must read each reading *before* class and come prepared with questions.

DATE	READING	NOTES
Thursday, July 5	Introduction	
Monday, July 9	Descartes, Meditation II (PDF)	Theme: Human and Non-human Rights & Values
Tuesday, July 10	Kant, <i>Groundwork</i> , Section I (All) & II (Focus on 4:433 - ); Universal Declaration of Human Rights (PDF)	
Wednesday, July 11	Star Trek, “The Measure of a Man”; Contemporary Reading TBD	
Thursday, July 12	Debate I; <i>Blade Runner</i>	
Monday, July 16	Dreyfus, selections from <i>What Computers Still Can’t Do</i> ; Shapiro, <i>Embodied Cognition</i> , Chapter 2; Dennett: ‘Cognitive Wheels: The Frame Problem of AI’: <a href="https://www.youtube.com/watch?v=EVAKG6Y2uIg">https://www.youtube.com/watch?v=EVAKG6Y2uIg</a>	Theme: AI Capacities and Development
Tuesday, July 17	Haugeland, <i>Artificial Intelligence: The Very Idea</i> (selections); Kurzweil, <i>How to Create a Mind</i> (selections); Video: Meet David: <a href="https://www.youtube.com/watch?v=RJ7E7Qp-s-8">https://www.youtube.com/watch?v=RJ7E7Qp-s-8</a> Video: Are We Approaching Robotic Consciousness?: <a href="https://www.youtube.com/watch?v=J TOMNkZJRao">https://www.youtube.com/watch?v=J TOMNkZJRao</a>	
Wednesday, July 18	Asimov, <i>I, Robot</i> ; Selections from <i>Westworld</i>	
Thursday, July 19	Asimov, <i>I, Robot</i> , cont.; Debate II	<b>First Paper Due</b>
Monday, July 23	Heidegger, “The Question Concerning Technology”	Theme: Humanity’s Evolving Relationship with Technology
Tuesday, July 24	Heidegger, “The Question Concerning Technology,” cont.; SEP, “Human Enhancement”; Transhumanism Readings TBD	
Wednesday, July 25	<i>2001: A Space Odyssey</i>	
Thursday, July 26	Debate III	
Monday, July 30	Care Ethics: Readings TBD; Lem, <i>Solaris</i>	Theme: Capacity to Love <b>Final Paper Draft Due (2 copies)</b>
Tuesday, July 31	Dick, <i>Do Androids Dream of Electric Sheep?</i>	
Wednesday, August 1	Dick, <i>Do Androids Dream of Electric Sheep?</i> , cont.	
Thursday, August 2	Peer Review	<b>Completed Peer Review Packet Due (2 copies)</b>
Monday, August 6	Symposium	<b>Final Paper Due; Journal Due</b>