

# ARTIFICIAL INTELLIGENCE: BRAINS, MINDS, AND MACHINES

Summer 2020

---

<b>Instructor:</b> Shen Pan	<b>Time:</b> M–F 9:00 AM – 12:00 PM
<b>Email:</b> <a href="mailto:shenpan@umd.edu">shenpan@umd.edu</a>	<b>Place:</b> TBA

---

**Office & Office Hours:** 1110D Skinner; TBA.

**Course Description:** This course provides an introduction to the philosophy of artificial intelligence. We are on the verge of constructing artificial systems that will rival or surpass human minds in many ways. Indeed, the current world champions in chess, go, and JEOPARDY! are all AI-powered computers. Despite the importance and prominence of artificial intelligence, however, there remain controversial conceptual and foundational issues concerning the nature and creation of artificial intelligence. Just what it is for machines to be intelligent is not at all obvious. Also, when AI-powered computers find solutions to complex problems, do they engage in *thinking*? Will it be better if they do? How, if at all, are such machines different from human minds in terms of producing intelligent behavior? What methodological—in addition to technical—challenges arise when it comes to different *types* of artificial intelligence?

The ever-increasing ubiquity of artificial intelligence also raises new and urgent ethical questions. How will artificial intelligence impact our world, and what can we do to make sure the impact is beneficial? Specifically, how can we make sure that AI systems will respect our ethical principles when they make decisions at speeds mere humans can't achieve or on the basis of "reasoning" that mere human can't comprehend? Should we think of AI-powered computers merely as sophisticated machines? Or are they a new form of life? What legal and moral rights, if any, should we grant them?

The ultimate goal of this course is *not* to arrive at definitive, settled answers to these questions. Rather, it will introduce to students philosophical frameworks for thinking about them, while trying to achieve a deeper understanding of artificial intelligence. In other words, this course will teach you how to think philosophically about artificial intelligence, at a high level of abstraction and from a certain distance from practice. To this end, three modules will be covered over the span of three weeks of this course: AI & Mind, AI & Cognition, and AI Ethics & Politics.

**Learning Objectives:** I have two broad, interconnected goals for all my students: to become better thinkers, and to become better writers. Specifically for this course, if you work hard, at the end of the semester you can expect to be able to:

1. gain a foundation of knowledge in critical concepts of and theoretical approaches to artificial intelligence;
2. understand and critically review recent history of thought about artificial intelligence;
3. identify and analyze ethical, social, and political implications of both artificial intelligence and our interaction with it.

This course also has a meta-goal: to teach you to integrate philosophy and the cognitive sciences, while being critical of both.

**Learning Materials:** With the exception of the first and last days, there will be required readings (sometimes hearings or viewings) for each class. *These are supposed to be finished on your own before coming to class.* Each lecture is devoted to building upon and expanding—sometimes significantly—the content from the readings (i.e., *not* explaining them).

All learning materials will be posted on the course ELMS site. This means that there are no required purchases for this course. Nevertheless, you may wish to consult as background reading Margaret A. Boden's book, *AI: Its Nature and Future* (Oxford University Press, 2016).

I am always happy to provide you with further learning materials if you want to go more into depth with a topic. Two resources, both free and online, that you may find particularly helpful are the Stanford Encyclopedia of Philosophy (<https://plato.stanford.edu>) and the Internet Encyclopedia of Philosophy (<https://www.iep.utm.edu>).

**Communication:** I use ELMS to send out important announcements, and it is your responsibility to check them regularly. In general, regular access to the course ELMS site is required, as any additional readings, activities, assignments, etc. will be posted there as well.

The best way to contact me is via email, not the ELMS messaging service, replies to ELMS announcements, or anything else. Please feel free to email me any time. I do my best to maintain a 24-hour reply policy on weekdays, and a 48-hour reply policy on weekends. That said, in my experience and based on testimonies from previous students, in most cases the most effective way of getting a satisfying answer—not merely a reply—from me is to talk to me in person.

**Academic Integrity:** Don't cheat. Anyone who engage in academic dishonesty will be immediately reported to the relevant disciplinary authorities for further action. If you have questions about what constitutes cheating or academic dishonesty, please do not hesitate to ask me. You can also consult the Office of Student Conduct (<https://www.studentconduct.umd.edu>).

If you are having difficulty completing an assignment or are not satisfied with your performance, please see me as soon as possible so that we can discuss solutions. Do not resort to cheating in any shape or form, which can jeopardize your future at the University of Maryland.

**Accessibility and Inclusion:** *Your academic achievement is important to me.* I am committed to the principle that no qualified individual with a disability shall, on the basis of disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of the University, or be subjected to discrimination. If you have a disability that may prevent you from fully demonstrating your abilities, please contact me as soon as possible, so that we can discuss how accommodations can be implemented, to ensure full participation and to make your learning experience comfortable and effective.

The University of Maryland values the diversity of its student body. Along with the University, I am committed to providing a classroom atmosphere that encourages the equitable participation of all students regardless of age, disability, ethnicity, gender, national origin, race, religion, or sexual orientation. Potential devaluation of students in the classroom that can occur by reference to demeaning stereotypes of any group and/or overlooking the contributions of a particular group to the topic under discussion is inappropriate.

**Course Assessment:** Assessment is based on four categories of graded items, and they are worth a total of 900 points. The breakdown is as follows:

Daily Quizzes (10) .....	200 points
Short Argument Reconstruction Essays (2) .....	200 points
Critical Response Paper (1) .....	200 points
Unit Exams (3) .....	300 points

(Instructions and deadlines are individually posted for each graded item on ELMS, under “Module”.)

Your grade is determined by your performance and your performance alone (i.e., not curved or rounded up). To be fair to everyone, I have to establish clear standards and apply them consistently, so please understand that being close to a cutoff—no matter how close—is not the same thing as making the cut. For the purpose of calculating final grades, the following conversion will be used:

A+:	97%	A:	93%	A-:	90%
B+:	87%	B:	83%	B-:	80%
C+:	77%	C:	73%	C-:	70%
D+:	67%	D:	63%	D-:	60%
F:	below 60%.				

If earning a particular grade is important to you, please speak with me at the beginning of the semester so that I can offer some helpful suggestions for achieving your goal. In general, in addition to attending class regularly, you should expect to spend at least 4 hours every day on readings, quizzes, and written assignments, plus additional time on weekends. *This is in line with college courses, which typically require 2 hours outside of class for every hour in class.*

**Late Work Policy:** Because they are an integral part of each week’s learning experience, for the Daily Quizzes and Argument Reconstruction Essays, late work will *not* be accepted. Likewise, since all Unit Exams are in-class and timed, they cannot be rescheduled or made up. Late work for the Critical Response Paper will be accepted, for up to three days, with 10% of the grade deducted per day of being late.

I do realize, of course, that even the most diligent students may have to miss a deadline on occasion due to illness or some other emergency. Should such unfortunate/inconvenient situations occur, please get in touch with me as soon as possible, with relevant documentation, so that we can discuss plans for you promptly.

**Other Course Related Policies:** It is our shared responsibility to know and abide by the University of Maryland’s policies that relate to all courses. Please visit <https://go.umd.edu/ug-policy> for the Office of Undergraduate Studies’ full list of campus-wide policies, read them carefully, and follow up with me if you have any questions.

## Course Schedule

(The schedule below is subject to change, depending on the pace at which we move through material and student interests.)

### Unit 1: AI & Mind

- July 13: Discussion of syllabus; introduction to the course; kick-off discussion; *No reading*
- July 14: John Haugeland (1997), What Is Mind Design?; *Quiz 1*
- July 15: Alan Turing (1950), Computing Machinery and Intelligence; *Quiz 2*
- July 16: Hilary Putnam (1967), The Nature of Mental States; John R. Searle (1983), Can Computers Think?; *Quiz 3*
- July 17: Margaret A. Boden (1988), Escaping From The Chinese Room; *Exam 1*

### Unit 2: AI & Cognition

- July 20: John R. Searle (1990), Is The Brain a Digital Computer?; *Quiz 4*.
- July 21: Ned Block (2003), The Mind as Software in the Brain; Robert Epstein (2016), Your Brain Does not Process Information And It Is not a Computer; *Quiz 5*.
- July 22: Alan Newell and Herbert Simon (1981), Computer Science as Empirical Inquiry: Symbols and Search; *Quiz 6*
- July 23: Daniel C. Dennett (1987), Cognitive Wheels: The Frame Problem of AI; *Quiz 7*
- July 24: Yann LeCun, Yoshua Bengio, and Geoffrey Hinton (2015), Deep Learning; *Exam 2*

### Unit 3: AI Ethics & Politics

- July 27: Awad Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-Francois Bonnefon, and Iyad Rahwan (2018), The Moral Machine Experiment; *Quiz 8*
- July 28: Nassim JafariNaimi (2018), Our Bodies in the Trolley's Path, or Why Self-driving Cars Must *not* be Programmed to Kill; *Quiz 9*
- July 29: Michael LaBossiere (2017), Testing the Moral Status of Artificial Beings, or "Im Going to Ask You Some Questions"; *Quiz 10*
- July 30: Deborah G. Johnson (2006), Computer Systems: Moral Entities But Not Moral Agents; Ex Machina (2014 film); round table discussion.
- July 31: Review; *Exam 3*