Course info

Course name:	Philosophy of AI
Course code:	WBMV05003
Credits:	7,5 ECTS
Teaching period:	2021-2022 period 3
Level (1,2,3 or M)	M
Part of curriculum:	Master Artificial Intelligence: required course Master History and Philosophy of Science: primary elective Research Master Philosophy: primary elective
Schedule:	Lectures / Lecture sessions: - Wednesdays, 9.45-10.45 Seminars (starting February, 21): - group 1, Rijk Mercuur - group 2, Miguel Segundo Ortin - group 3, Maksim Gladyshev - group 4, Janneke van Dis - group 5, Natasha Alechina - group 6, Annet Onnes
Lecturers:	Lectures are taught by: Janneke van Dis, Natasha Alechina, Rijk Mercuur, Mehdi Dastani, Sven Nyholm Seminars are taught by: Natasha Alechina, Maksim Gladyshev, Janneke van Dis, Annet Onnes, Miguel Segundo Ortin, Rijk Mercuur
Contact person for the course:	Janneke van Dis (j.h.vandis@uu.nl)

Course goals

This course will make students familiar with fundamental issues in the philosophy of AI, and will introduce them to several current discussions in the field. Students will practice their argumentation and presentation skills, both in class discussions and in writing.

Course content

The course is split up in three parts. The first part is a quick overview of the fundamental issues and core notions in philosophy of AI. It addresses topics such as the Turing Test, the Chinese Room Argument, machine intelligence, machine consciousness, weak and strong AI, and the Symbol System Hypothesis. In order to establish a shared background for all students, the material of this part will be assessed with a writing assignment already in week 3.

In the second part of the course, there will be an in-depth discussion of several current topics in the field, for example on ethics and responsibility in AI, decision making, or the relation between AI and data science. On each topic, there will be a lecture, and a seminar with class discussions and student presentations. Students prepare for those discussions by posting a thesis with one or more supporting arguments about the required reading. In the third part of the course, students will write a philosophical paper, and will provide feedback on their fellow students' draft papers.

Entry requirements

You have to be registered for one of the following degree programmes:
- Artificial Intelligence
- History and Philosophy of Science
- Philosophy (research master)

Required materials

Throughout the course, we will make use of this handbook:

- Frankish, K., & Ramsey, W. M. (Eds.). (2014). *The Cambridge handbook of artificial intelligence*. Cambridge University Press. [paperback 28 euro; freely available as e-book from the University Library]

Next to this, there is required reading for each of the seminars, and also additional reading for each of the five themes which may be used for the in-class presentations and in case you write your final paper on that theme. You can find (links to) these papers on Blackboard, under Course content.

Online and on-campus teaching

Because of the covid-19 pandemic and the current maximum group size of 75 people in higher education meetings, the lectures in this course will be online. Seminars will be on campus. We need to be flexible though, so be prepared for last-minute switches to online seminars, e.g. when a teacher needs to work from home, or when UU-regulations require us to do so. You can find the current state of University-wide rules and regulations with respect to the covid-19 situation here: https://www.uu.nl/en/information-coronavirus/principles-and-rules.

Seminars will be on campus, and in line with UU policy there won't be a full-fledged online alternative for those who can't attend because they're in quarantine. There will be *some* possibility for participating online, though. We will make sure everyone has a *buddy*, i.e. a fellow student who can give you remote access to the seminar by setting up a one-on-one videochat. It will be difficult to participate remotely in plenary discussions during the seminar, but it will be easier to have a breakout discussion group consisting solely of students with remote access. And in case you need to give a presentation but know in advance that you won't be able to attend physically, you can prepare an online presentation (either prerecorded or live).

Lectures will be online. This may cause trouble for those students who need to be on campus right after the lecture; especially, for those of you who are also taking the course Logic and Computation. For these students, there will be an opportunity to attend the online lecture together in a seminar room on campus (details via Blackboard).

We will make use of MS Teams for live meetings (team code available through Blackboard). We will make use of Blackboard for written assignments, course communication and the distribution of course materials. We will make use of the FeedBackFruits plugin in Blackboard for peerfeedback on versions of the final paper.

Lectures

Instead of regular, on-campus lectures we will have live online interactive lectures. That is, we will make use of a *flipped classroom*. Most of the content will be available in pre-recorded video lectures, which are available through Blackboard. Please make sure you have watched those videos prior to the live lecture. During the live lecture, there will be ample opportunity for discussion and questions. You can prepare for these meetings by having questions ready that you would like to ask to the lecturer.

Because most of the content is in the pre-recorded videos, lectures will take less time than the usual two hours; we will limit them to 60 minutes at most. For this reason, we will start at 9:45am (even though your schedule in MyTimetable will show 9:00am).

Live lectures will *not* be recorded. This is because we want to respect your privacy, and we want everyone to feel free to participate actively during the meeting.

Seminars

For each of the five themes of this course, there will be a lecture (on Wednesdays) followed by a seminar (on Mondays).

Preparation

You prepare for the seminars by carefully reading the assigned texts. Prior to the meeting, on Fridays at noon at the latest, you will post a *four-sentence-paper* on Blackboard related to the assigned texts. A four-sentence-paper has the following structure:

1. They say ...
2. I say ..., because ...
3. One might object that ...
4. I reply that ...

One goal of these assignments is to get a good grip on philosophical argumentation. Another goal is to get a head-start for your final paper: it would be nice if one of these five assignments is worked out later from four sentences into a 4000 word paper. (Note, though, that it is not strictly required to base your final paper on one of your assignments; you have the freedom to pick another topic.)

The seminars themselves contain three different parts:

1) Presentations

Groups of two students will present and comment on an additional paper about that week's theme. You can find suggestions for papers to present on Blackboard, under Course Content, but you are encouraged to find something to present yourself (but please check with your seminar teacher). Presentations take at most 15 minutes including discussion (and at most 10 minutes for the presentation itself). They should contain: an outline of the presentation itself, a presentation of the main standpoints and arguments from the paper, and your own critical comments on the paper. Presentations will be graded; the assessment criteria can be found on Blackboard. There will be roughly two such presentations per meeting.

2) Discussions in break-out groups

After a short break, the required reading for that particular theme will be discussed on the basis of the four-sentence papers which were posted on Blackboard. This will be done in smaller break-out groups.

3) Plenary discussion

During the final 30 minutes of the seminar, one spokesperson per breakout group will report about the discussion in their group.

Assessment

1. Writing assignment (0% - pass required to pass the course)

The material for the first two weeks of the course serves to provide all participants in the course with a shared background knowledge on philosophical issues related to AI. It will be tested by a writing assignment which can be found on Blackboard. The assignment will require you to study chapters 1-6 of *The Cambridge handbook of artificial intelligence*. The topics will also be explained in the videos and live lectures of the first two weeks.

2. Oral presentation (25% of final grade)

All participants of the course give an oral presentation as part of the seminars.

3. Final paper (50% of final grade)

All participants will write a paper of 3000-4000 words related to one of the five themes. Two earlier stages of the paper will be subjected to peer feedback: an outline of the paper and a complete first version. For the feedback, we will make use of FeedbackFruits, which is accessible through Blackboard. Only the final version will be graded.

4. Seminar participation (25% of final grade)

All participants are expected to attend all seminar meetings, to prepare for them by carefully reading the assigned texts, to post their four-sentence-paper about the assigned texts prior to the seminar meetings, to participate in the discussion, and to participate in the peer review of the draft papers.

The grade for participation will mainly be based on the quality of the four-sentence-papers, but it will also reflect the formal aspects of participation (attendance, responsiveness, being courteous during discussions) and the quality of the contributions to the class discussions.

Deadline policy

Deadlines are strict. If you miss the deadline for the final version of the paper, points will be subtracted from your grade. If, however, there are personal or covid-related circumstances for missing a deadline, you can ask for an extension by contacting your seminar teacher prior to the deadline. Extensions of more than one week will only be granted after consulting the study advisor.

Fraud and plagiarism

Please make sure that you are aware of the UU rules about these subjects: https://students.uu.nl/en/practical-information/policies-and-procedures/fraud-and-plagiarism.

The final paper will be checked for plagiarism through Urkund.

Schedule

Week	Session	Deadlines and assignments	Topic
1 - Feb 7	Mon: No class		
	Wed: Lecture		Welcome and introduction to the course
2 - Feb 14	Mon: Lecture Janneke van Dis	Read handbook Ch 1-3	Handbook Ch 1-3
	Wed: Lecture Janneke van Dis	Read handbook Ch 4-6	Handbook Ch 4-6
3 - Feb 21	Mon: Seminar	Tue Feb 22, 17:00: writing assignment	Introduction, practicalities
	Wed: Lecture Natasha Alechina	Read mandatory seminar papers Friday 12:00: Post weekly assignment on BB	Theme 1: What is a good test for intelligence?
4 - Feb 28	Mon: Seminar		Theme 1: What is a good test for intelligence?
	Wed: Lecture Sven Nyholm	Read mandatory seminar papers Friday 12:00: Post weekly assignment on BB	Theme 2: Ethics and AI
5 - Mar 7	Mon: Seminar		Theme 2: Ethics and AI
	Wed: Lecture Mehdi Dastani	Read mandatory seminar papers Friday 12:00: Post weekly assignment on BB	Theme 3: Decision making
6 - Mar 14	Mon: Seminar		Theme 3: Decision making
	Wed: Lecture Rijk Mercuur	Read mandatory seminar papers Friday 12:00: Post weekly assignment on BB	Theme 4: Simulations and models
7 - Mar 21	Mon: Seminar		Theme 4: Simulations and models
	Wed: Lecture Janneke van Dis	Read mandatory seminar papers Friday 12:00: Post weekly assignment on BB	Theme 5: Explainable AI
8 - Mar 28	Mon: Seminar	Tue Mar 29, 17:00: Final paper, research question and outline	Theme 5: Explainable AI
	individual meetings with seminar teacher on request	Wed Mar 30, 23:59: Peer feedback on research question and outline	
9 - Apr 4	Mon: Seminar		Final paper progress meeting (on campus)
		Fri Apr 8, 23:59: Final paper, first version	
10 - Apr 11	Mon: individual 10 min meetings with seminar teacher	Mon Apr 11, 23:59: Peer feedback on first version	Final paper progress meeting (individual, online)
		Fri Apr 15, 23:59: Final paper, final version	