

**Philosophy of Artificial Intelligence**

Code: 100315  
ECTS Credits: 6

Degree	Type	Year	Semester
2500246 Philosophy	OT	3	1
2500246 Philosophy	OT	4	1

**Contact**

Name: María Pilar Dellunde Clave  
Email: pilar.dellunde@uab.cat

**Use of Languages**

Principal working language: catalan (cat)  
Some groups entirely in English: No  
Some groups entirely in Catalan: Yes  
Some groups entirely in Spanish: No

**Teachers**

María Pilar Dellunde Clave  
Vicente Costa Bueno

**Prerequisites**

None.

**Objectives and Contextualisation**

Can a machine be creative? Can a machine think, have emotions? Artificial intelligence is present in our daily life, in science, in art, and unfortunately also in war. What are the main issues to be addressed from a philosophical point of view regarding artificial intelligence? Do we need alternative logics for the representation of these new knowledge?

In this course we will discuss the ethical limits of the application of artificial intelligences, and the possibilities of designing a general artificial intelligence. A new society where people and agents of artificial intelligence coexist creating communities with completely different norms and potentials than the ones we have lived up to now. All the intelligence of the humanities is needed to face these new challenges. Would you like to join us in the creation of these alphabets for the future?

**Competences**

- Philosophy
- Students must have and understand knowledge of an area of study built on the basis of general secondary education, and while it relies on some advanced textbooks it also includes some aspects coming from the forefront of its field of study.
- Using the symbology and procedures of the formal sciences in the analysis and building of arguments.

**Learning Outcomes**

1. Establishing relationships between science, philosophy, art, religion, politics, etc.
2. Explaining the specific notions of the History of Philosophy.
3. Formulating arguments for and against an issue, using proper vocabulary, conceptual precision and argumentative coherence.
4. Recognising and using the several forms of reasoning in the history of philosophy.
5. Regularising arguments of any source and calculating its logical correctness.
6. Rigorously building philosophical arguments.

## Content

1. What is Artificial Intelligence (AI)? Machines and algorithms. Symbolic AI vs. Subsymbolic AI.
2. Ethical, social and epistemological challenges of machine learning.
3. Gender perspective in AI.
4. Reflections on the creativity of AI.
5. Knowledge representation, natural language and explainability.
6. Robotics. Sex with robots?
7. Future perspectives. Is a general AI possible?

## Methodology

- Combination of theoretical and practical classes.
- Joint resolution of exercises.
- Use of science fiction to work on philosophical creativity.
- Self-learning activities
- Philosophical immersive role-playing games..

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

## Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Collaborative practice of Applied Philosophy	40	1.6	1, 2, 3, 4
Workshop on philosophical creativity	40	1.6	6, 2, 3, 4
Type: Supervised			
Study of the basic notions of AI and knowledge representation	30	1.2	6, 2, 5
Type: Autonomous			
Reading texts of Philosophy of the AI	24	0.96	6, 1, 4

## Assessment

The assessment is presented continuously, with three assessment activities: a synthesis test, a workshop on philosophical creativity using science fiction, philosophical texts and role-playing games, which will take place

throughout the course, and a applied philosophy activity around AI, to be chosen by students (for example, organizing a round table, writing a code of ethics, writing an academic article, etc.). Each evaluation activity will be worth one third of the final grade. At the time of each assessment activity, students will be informed (via Moodle) of the procedure and date of review of grades.

In the event of a student committing any irregularity that may lead to a significant variation in the grade awarded to an assessment activity, the student will be given a zero for this activity, regardless of any disciplinary process that may take place. In the event of several irregularities in assessment activities of the same subject, the student will be given a zero as the final grade for this subject.

In the event that tests or exams cannot be taken onsite, they will be adapted to an online format made available through the UAB's virtual tools (original weighting will be maintained). Homework, activities and class participation will be carried out through forums, wikis and/or discussion on Teams, etc. Lecturers will ensure that students are able to access these virtual tools, or will offer them feasible alternatives.

## Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Collaborative practice of Applied Philosophy	33%	4	0.16	6, 1, 3
Definitions of the main AI concepts.	33%	4	0.16	6, 5, 4
Philosophical creativity workshop	34%	8	0.32	6, 1, 2, 3, 4

## Bibliography

1. Johan van Benthem, *Modal Logic for Open Minds*, Center for the Study of Language and Information, 2000.
2. Johan van Benthem, Hans van Ditmarsch, Jan van Eijck, Jan Jaspars, *Logic in Action*, Center for the Study of Language and Information, 2016.
3. Margaret A. Boden, *AI: Its nature and future*, Oxford University Press, 2016
4. Jack Copeland, *Artificial Intelligence: A Philosophical Introduction*, Wiley-Blackwell, 1993.
5. Pedro Messeguer, Ramón López de Mantaras, *Inteligencia Artificial*, Editorial CSIC, 2017.
6. Nick Smith, *Vagueness and degrees of truth*, Oxford University Press, 2013.
7. Stuart Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd edition, Prentice Hall Press, 2009.
8. Alan Turing, *Computing Machinery and Intelligence*, *Mind*, Issue 236, pp. 433-460, 1950.
9. Mark Coeckelberg, *Ética de la inteligencia artificial*, 1a edició, Càtedra, 2018.

## Software

None