

**Filosofia de la intel·ligència artificial**

Codi: 100315

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2500246 Filosofia	OT	3	1
2500246 Filosofia	OT	4	1

**Professor/a de contacte**

Nom: María Pilar Dellunde Clave

Correu electrònic: pilar.dellunde@uab.cat

**Utilització d'idiomes a l'assignatura**

Llengua vehicular majoritària: català (cat)

Grup íntegre en anglès: No

Grup íntegre en català: Sí

Grup íntegre en espanyol: No

**Equip docent**

María Pilar Dellunde Clave

Vicente Costa Bueno

**Prerequisits**

L'assignatura no té cap prerequisit.

**Objectius**

Pot crear una màquina? Pot pensar, tenir emocions? La intel·ligència artificial és present en la nostra quotidianitat, en la ciència, en l'art, i malauradament també en la guerra. Quines preguntes i reflexions ens fem des de la filosofia? Necessitem lògiques alternatives per a la representació d'aquests nous coneixements? En aquest curs reflexionarem filosòficament sobre la intel·ligència artificial, sobre els límits ètics de la seva aplicació, les possibilitats de dissenyar una intel·ligència artificial general, i sobre les seves arquitectures cognitives. Una nova societat on persones i agents d'intel·ligència artificial convivim creant comunitats amb normes i potencialitats completament diferents de les que hem viscut fins ara, necessita tota la intel·ligència de les humanitats per fer front als nous reptes. T'apuntes a crear aquests alfabets de futur?

**Competències**

Filosofia

- Que els estudiants hagin demostrat que comprenen i tenen coneixements en una àrea d'estudi que parteix de la base de l'educació secundària general, i se sol trobar a un nivell que, si bé es basa en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda d'aquell camp d'estudi.
- Utilitzar la simbologia i els procediments de les ciències formals en l'anàlisi i la construcció d'arguments.

## Resultats d'aprenentatge

1. Construir arguments filosòfics amb rigor.
2. Establir relacions entre ciència, filosofia, art, religió, política, etc.
3. Exposar els conceptes propis de la història de la filosofia
4. Formalitzar arguments de qualsevol procedència i calcular-ne la correcció lògica.
5. Formular objeccions i contraobjeccions amb correcció lèxica, precisió conceptual i coherència argumentativa.
6. Reconèixer i utilitzar les diverses formes de raonament presents en la història de la filosofia.

## Continguts

1. Què és la Intel·ligència Artificial (IA)? Màquines i algoritmes. IA simbòlica vs. IA subsimbòlica
2. Reptes ètics, socials i epistemològics de l'aprenentatge automàtic.
3. Perspectiva de gènere en IA.
4. Reflexions sobre la creativitat de les IA.
5. Representació del coneixement, llenguatge natural i explainability.
6. Roboètica. Sexe amb robots?
7. Perspectives futures. Es possible una IA general?

## Metodologia

- Combinació de classes teòriques i pràctiques.
- Resolució conjunta d'exercicis.
- Ús de la ciència-ficció per treballar la creativitat filosòfica.
- Activitats d'autoaprenentatge.
- Jocs de rol immersius de naturalesa filosòfica

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

## Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Pràctica col·laborativa de Filosofia Aplicada	40	1,6	2, 3, 5, 6
Taller de creativitat filosòfica	40	1,6	1, 3, 5, 6
Tipus: Supervisades			
Estudi de les nocions bàsiques de IA i representació del coneixement	30	1,2	1, 3, 4
Tipus: Autònomes			
Lectura de textos de Filosofia de la IA	24	0,96	1, 2, 6

## Avaluació

L'avaluació es planteja de forma continuada, amb tres activitats avaluatives: una prova de síntesi, un taller de creativitat filosòfica utilitzant la ciència-ficció, textos filosòfics i jocs de rol, la qual tindrà lloc al llarg de tot el curs, i una activitat de filosofia aplicada al voltant de la IA, a escollir per l'alumnat (per exemple, organització d'una taula rodona, redacció d'un codi ètic, escriptura d'un article acadèmic, etc.). Cada activitat avaluativa

valdrà un terç de la nota final. En el moment de realització de cada activitat avaluativa, s'informarà l'alumnat (via Moodle) del procediment i data de revisió de les qualificacions.

En cas que l'estudiant realitzi qualsevol irregularitat que pugui conduir a una variació significativa de la qualificació d'un acte d'avaluació, es qualificarà amb 0 aquest acte d'avaluació, amb independència del procés disciplinari que s'hi pugui instruir. En cas que es produeixin diverses irregularitats en els actes d'avaluació d'una mateixa assignatura, la qualificació final d'aquesta assignatura serà 0.

En cas que les proves no es puguin fer presencialment s'adaptarà el seu format (mantenint-ne la ponderació) a les possibilitats que ofereixen les eines virtuals de la UAB. Els deures, activitats i participació a classe es realitzaran a través de fòrums, wikis i/o discussions d'exercicis a través de Teams, etc. El professor o professora vetllarà perquè l'estudiant hi pugui accedir o li oferirà mitjans alternatius, que estiguin al seu abast.

## Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Definicions dels conceptes fonamentals.	33%	4	0,16	1, 4, 6
Practica colaborativa de Filosofia Aplicada	33%	4	0,16	1, 2, 5
Taller de creativitat filosòfica	34%	8	0,32	1, 2, 3, 5, 6

## Bibliografia

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, Cátedra, 2018.

## Programari

Cap programari