

Filosofía de la inteligencia artificial

Código: 100315
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
2500246 Filosofía	OT	3	1
2500246 Filosofía	OT	4	1

Contacto

Nombre: María Pilar Dellunde Clave
Correo electrónico: pilar.dellunde@uab.cat

Uso de idiomas

Lengua vehicular mayoritaria: catalán (cat)
Algún grupo íntegramente en inglés: No
Algún grupo íntegramente en catalán: Sí
Algún grupo íntegramente en español: No

Equipo docente

María Pilar Dellunde Clave
Vicente Costa Bueno

Prerequisitos

Ninguno.

Objetivos y contextualización

¿Puede crear una máquina? ¿Puede pensar, tener emociones? La inteligencia artificial está presente en nuestra cotidianidad, en la ciencia, en el arte, y desgraciadamente también en la guerra. ¿Qué preguntas y reflexiones nos hacemos desde la filosofía? ¿Necesitamos lógicas alternativas para la representación de estos nuevos conocimientos?

En este curso reflexionaremos filosóficamente sobre la inteligencia artificial, sobre los límites éticos de su aplicación, las posibilidades de diseñar una inteligencia artificial general, y sobre sus arquitecturas cognitivas. Una nueva sociedad donde personas y agentes de inteligencia artificial convivimos creando comunidades con normas y potencialidades completamente diferentes de las que hemos vivido hasta ahora, necesita toda la inteligencia de las humanidades para hacer frente a los nuevos retos ¿Te apuntas a crear estos alfabetos de futuro?

Competencias

- Filosofía
- Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio.

- Utilizar la simbología y procedimientos de las ciencias formales en el análisis y construcción de argumentos.

Resultados de aprendizaje

1. Construir argumentos filosóficos con rigor.
2. Establecer relaciones entre ciencia, filosofía, arte, religión, política, etc.
3. Exponer los conceptos propios de la historia de la filosofía
4. Formalizar argumentos de cualquier procedencia y calcular su corrección lógica.
5. Formular objeciones y contraobjeciones con corrección léxica, precisión conceptual y coherencia argumentativa.
6. Reconocer y utilizar las diversas formas de razonamiento presentes en la historia de la filosofía.

Contenido

1. ¿Qué es la Inteligencia Artificial (IA)? Máquinas y algoritmos. IA simbólica vs. IA subsimbólica
2. Retos éticos, sociales y epistemológicos del aprendizaje automático.
3. Perspectiva de género en IA.
4. Reflexiones sobre la creatividad de las IA.
5. Representación del conocimiento, lenguaje natural y explainability.
6. Roboética. ¿Sexo con robots?
7. Perspectivas futuras. ¿Es posible una IA general?

Metodología

- Combinación de clases teóricas y prácticas.
- Resolución conjunta de ejercicios.
- Uso de la ciencia ficción para trabajar la creatividad filosófica.
- Actividades de autoaprendizaje.
- Juegos de rol inmersivos de naturaleza filosófica.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Práctica colaborativa de Filosofía Aplicada	40	1,6	2, 3, 5, 6
Taller de creatividad filosófica	40	1,6	1, 3, 5, 6
Tipo: Supervisadas			
Estudio de las nociones básicas de IA y representación del conocimiento	30	1,2	1, 3, 4

Evaluación

La evaluación se plantea de forma continuada, con tres actividades evaluativas: una prueba de síntesis, un taller de creatividad filosófica utilizando la ciencia ficción, textos filosóficos y juegos de rol, que tendrá lugar a lo largo de todo el curso, y una actividad de filosofía aplicada en torno a la IA, a escoger por el alumnado (por ejemplo, organización de una mesa redonda, redacción de un código ético, escritura de un artículo académico, etc.). Cada actividad evaluativa valdrá un tercio de la nota final. En el momento de realización de cada actividad evaluativa, se informará al alumnado (vía Moodle) del procedimiento y fecha de revisión de las calificaciones.

En caso de que el estudiante lleve a cabo cualquier tipo de irregularidad que pueda conducir a una variación significativa de la calificación de un determinado acto de evaluación, este será calificado con 0, independientemente del proceso disciplinario que pueda derivarse de ello. En caso de que se verifiquen varias irregularidades en los actos de evaluación de una misma asignatura, la calificación final de esta asignatura será 0.

En caso de que las pruebas no se puedan hacer presencialmente, se adaptará su formato (sin alterar su ponderación) a las posibilidades que ofrecen las herramientas virtuales de la UAB. Los deberes, actividades y participación en clase se realizarán a través de foros, wikis y / o discusiones de ejercicios a través de Teams, etc. El profesor o profesora velará para asegurarse el acceso del estudiantado a tales recursos o le ofrecerá otros alternativos.

Actividades de evaluación

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Definiciones de los conceptos fundamentales en IA.	33%	4	0,16	1, 4, 6
Práctica colaborativa de Filosofía Aplicada	33%	4	0,16	1, 2, 5
Taller de creatividad filosófica	34%	8	0,32	1, 2, 3, 5, 6

Bibliografía

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 edición, Cátedra, 2018.

Software

Ningún software