

Titulación	Tipo	Curso
Filosofía	OT	3
Filosofía	OT	4

## Contacto

Nombre: Jordi Vallverdú Segura

Correo electrónico: jordi.vallverdu@uab.cat

## Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

## Prerrequisitos

Ninguno.

## Objetivos y contextualización

"Filosofía de la Inteligencia Artificial" es una asignatura pensada para estudiantes curiosos que quieren ir más allá de los tópicos sobre la IA. A lo largo del curso, analizaremos cómo nacieron las ideas de computación, los algoritmos y las máquinas inteligentes; cómo la lógica y los lenguajes formales han contribuido a dar forma al mundo digital; y qué significa realmente que un sistema entienda el lenguaje o simplemente lo simule. También exploraremos por qué los grandes modelos de lenguaje, como ChatGPT, plantean nuevos desafíos filosóficos, y cómo todo esto afecta a la sociedad, al conocimiento y a las decisiones éticas del presente y del futuro.

No es necesario saber programar -solo tener ganas de pensar en serio y poner la filosofía a trabajar sobre la tecnología real.

## 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

### Módulo I: Fundamentos históricos y conceptuales

1. Cálculo, censos y control: de Leibniz a Babbage
2. Lógica, máquinas y formalización (Frege, Boole, Peano, Hilbert)
3. Turing, Church y la idea de computabilidad
4. Arquitectura de Von Neumann y computación clásica
5. Computación cuántica y sus implicaciones filosóficas
6. Estadística y predicción en sistemas inteligentes

### Módulo II: Filosofía de la computación

7. ¿Qué es un algoritmo?
8. Lenguajes de programación y lógica (Lisp, Prolog)
9. Modelos computacionales de la mente (funcionalismo, Searle, Dreyfus)

### Módulo III: IA y filosofía

10. Historia de la IA: de McCarthy a AlphaGo
11. Ética de la IA: sesgos, responsabilidad, género
12. Conciencia e intencionalidad

### Módulo IV: Filosofía de los LLM

13. ¿Qué es un LLM? (Transformers, embeddings)
14. Críticas filosóficas (Stochastic Parrots, Bender, semántica)
15. Epistemología de la IA generativa

### Módulo V: Perspectivas contemporáneas

16. Debates filosóficos sobre AGI
17. IA neurodiversa e inspirada biológicamente
18. IA neuro-simbólica (sistemas híbridos)
19. Cibercultura y tecnologías emergentes
20. Conclusión: Hacia una filosofía crítica de la computación

## Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Taller de creatividad filosófica con IA (experimentación con prompting, rol de la IA)	40	1,6	1, 3, 5, 6

como autor)

Tipo: Supervisadas

Estudio de los fundamentos técnicos y filosóficos de la IA y la representación del conocimiento (conceptos clave, supervisión docente)	44	1,76	1, 3, 4
--	----	------	---------

Tipo: Autónomas

Lectura de textos de Filosofía de la IA	34	1,36	1, 2, 6
---	----	------	---------

La metodología de esta asignatura combina clases magistrales, seminarios participativos y actividades prácticas guiadas, con un fuerte énfasis en el análisis filosófico activo y el uso declarado de la inteligencia artificial como herramienta epistemológica.

Las clases magistrales introducirán los conceptos fundamentales y el contexto histórico-filosófico.

A través de lecturas comentadas semanales, los estudiantes ejercitarán la lectura crítica, el análisis conceptual y el debate colectivo.

El trabajo escrito, elaborado con ChatGPT, fomentará la experimentación filosófica con herramientas de IA, exigiendo transparencia en el prompting y reflexión sobre su aplicación.

Los exámenes evaluarán la comprensión profunda del contenido y la capacidad de aplicarlo de forma argumentada.

La asignatura promueve una participación activa, una actitud crítica y una mirada transversal sobre el papel de la computación y la inteligencia artificial en la sociedad contemporánea.

En esta asignatura se permite el uso de tecnologías de Inteligencia Artificial (IA) como parte integrante del desarrollo del trabajo, siempre que el resultado final refleje una contribución significativa del estudiante en el análisis y la reflexión personal. El estudiante tendrá que: (i) identificar qué partes han sido generadas con IA; (ii) especificar las herramientas utilizadas; y (iii) incluir una reflexión crítica sobre cómo estas han influido el proceso y el resultado final de la actividad.

La no transparencia del uso de la IA en esta actividad evaluable se considerará falta de honestidad académica y computada que la actividad se evalúe con un 0 y no se pueda recuperar, o sanciones mayores en casos de gravedad.

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.

## Evaluación

### Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Examen final: síntesis crítica	25	8	0,32	1, 2, 3, 5, 6
Examen parcial: análisis y aplicación filosófica	25	8	0,32	1, 2, 5
Lecturas comentadas y debate filosófico	25	8	0,32	1, 2, 4, 6
Trabajo escrito con IA (ChatGPT)	25	8	0,32	1, 3, 4, 5

Lecturas comentadas semanales (25%)

Los estudiantes leerán textos filosóficos clave (Turing, Searle, Bender, etc.) y deberán participar activamente en discusiones guiadas en clase.

Se valorará:

- Comprensión del texto
- Capacidad de análisis y argumentación
- Participación en el debate colectivo

2. Trabajo escrito con ChatGPT (25%)

Trabajo individual (2.500-4.000 palabras) realizado con ChatGPT.

Será necesario incluir el hilo completo como anexo.

Se valorará:

- Capacidad crítica para analizar, complementar o corregir respuestas de la IA
- Nivel, densidad y refinamiento del prompting
- Claridad filosófica y reflexión propia
- Transparencia y justificación del proceso de interacción con la IA

3. Examen parcial (25%)

Prueba escrita presencial a mitad del curso.

Preguntas abiertas, definiciones, comparaciones y análisis de casos filosóficos.

4. Examen final (25%)

Prueba final presencial con preguntas abiertas o comentarios críticos de textos.

Es posible que el Departament de Filosofia establezca (como se hará durante el primer semestre) un período de concentración de pruebas evaluativas. El profesorado indicará si existe este período o cuáles son las fechas de las pruebas al inicio de cada asignatura.

## Evaluación Única

Los estudiantes que opten por la evaluación única deberán realizar un examen final exhaustivo al final del semestre. Esta prueba tiene como objetivo evaluar la comprensión global y crítica de los contenidos del curso.

El examen incluirá:

- Varias preguntas (tanto de desarrollo como de análisis argumentativo) que abarcarán todos y cada uno de los temas del programa.
- Se prestará especial atención a:
  - Los elementos filosóficos (cuestiones ontológicas, epistemológicas y metafísicas relacionadas con la IA),
  - Las cuestiones éticas y sociales derivadas de la aplicación de sistemas inteligentes,
  - Los aspectos lógicos y computacionales, como algoritmos, arquitecturas computacionales, redes neuronales y lógica formal aplicada a la IA.

El examen será presencial, tendrá una duración de dos horas y media, y se realizará durante el período oficial de evaluaciones finales.

Esta evaluación única cubrirá la totalidad de competencias y resultados de aprendizaje establecidos en la guía docente, en condiciones equivalentes a la evaluación continua.

## Bibliografía

### Filosofía de la computación e IA

- Boden, M. A. (2016). *IA: Su naturaleza y futuro*. Oxford University Press.
- Copeland, J. (1993). *Introducción filosófica a la inteligencia artificial*. Wiley-Blackwell.
- Dennett, D. (1991). *La consciencia explicada*. Little, Brown and Company.
- Floridi, L. (2011). *La filosofía de la información*. Oxford University Press.
- Searle, J. (1980). *Mentes, cerebros y programas*. *Mind*, 89(4), 417-424.
- Smith, B. C. (1996, 2019). *Sobre el origen de los objetos / La promesa de la inteligencia artificial*. MIT Press.
- Vallverdú, J. (2024). *Causalidad para la inteligencia artificial: una perspectiva filosófica*. Springer.

### Ética y sociedad

- Bender, E. et al. (2021). *Sobre los peligros de los loros estocásticos*. FAccT.
- Coeckelbergh, M. (2018). *Ética de la inteligencia artificial*. Cátedra.
- Coeckelbergh, M. (2022). *Ética de los robots*. MIT Press.
- Crawford, K. (2021). *Atlas de la IA*. Yale University Press.

### Lógica, epistemología y lenguaje

- Smith, N. J. J. (2013). *Vaguedad y grados de verdad*. Oxford University Press.
- Van Benthem, J. (2000). *Lógica modal para mentes abiertas*. CSLI.
- Van Benthem, J., Van Ditmarsch, H., Van Eijck, J., Jaspars, J. (2016). *Lógica en acción*. CSLI.

## Software

Ningún software.

## Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

Nombre	Grupo	Idioma	Semestre	Turno
(PAUL) Prácticas de aula	1	Catalán	primer cuatrimestre	mañana-mixto
(TE) Teoría	1	Catalán	primer cuatrimestre	mañana-mixto