

Titulació	Tipus	Curs
Intel·ligència Artificial / Bachelor in Artificial Intelligence	OB	3

Professor/a de contacte

Nom : Roger Deulofeu Batllori

Correu electrònic : roger.deulofeu@uab.cat

Prerequisites

No hi ha prerequisites

Idiomes dels grups

Podeu consultar aquesta informació al [final](#) del document.

Objectius

L'objectiu de l'assignatura és introduir l'estudiant en els principals problemes ètics vinculats al desenvolupament i l'ús de la intel·ligència artificial. A partir de conceptes bàsics de teoria ètica com el conseqüencialisme, la deontologia, l'autonomia, la responsabilitat o la justícia, el curs analitza casos concrets relacionats amb el biaix algorítmic, la privacitat, l'explicabilitat i l'opacitat, l'autonomia, la presa de decisions automatitzada i la governança de la IA. L'assignatura vol proporcionar eines conceptuais i argumentatives perquè l'estudiant pugui identificar dilemes ètics, avaluar críticament diferents posicions i justificar de manera raonada les seves pròpies decisions professionals. Aquesta dimensió és especialment rellevant en la formació d'enginyers i enginyeres, ja que les eines que dissenyen, desenvolupen i implementen poden tenir conseqüències d'alt impacte ètic i social, i és essencial que siguin capaços d'anticipar-les, avaluar-les i assumir-ne responsablement les implicacions.

Resultats d'aprenentatge

- KM10 (Identificar els mecanismes de protecció de la privacitat i la propietat intel·lectual en el marc dels sistemes d'intel·ligència artificial.) Identificar els mecanismes de protecció de la privacitat i la propietat intel·lectual en el marc dels sistemes d'intel·ligència artificial.
- KM11 (Descriure la normativa i la regulació de la intel·ligència artificial en els àmbits nacional, europeu i internacional.) Descriure la normativa i la regulació de la intel·ligència artificial en els àmbits nacional, europeu i internacional.
- KM12 (Descriure els supòsits de responsabilitat civil en l'ús de la intel·ligència artificial.) Descriure els supòsits de responsabilitat civil en l'ús de la intel·ligència artificial.
- SM10 (Analitzar l'impacte en la privacitat i el dret a la intimitat de les persones d'aplicacions d'intel·ligència artificial.) Analitzar l'impacte en la privacitat i el dret a la intimitat de les persones d'aplicacions d'intel·ligència artificial.
- SM12 (Determinar la normativa i la regulació aplicables a un cas d'ús de la intel·ligència artificial.) Determinar la normativa i la regulació aplicables a un cas d'ús de la intel·ligència artificial.

Continguts

1. Introducció a l'ètica de la IA: per què és important?

La rellevància social, política i professional de la reflexió ètica sobre la intel·ligència artificial.

2. Marcs ètics normatius

Principals enfocaments per avaluar accions, decisions i tecnologies: conseqüencialisme, deontologia i ètica de la virtut.

3. Llibertat, agència i autonomia

Com els sistemes d'IA poden afavorir, limitar o transformar la presa de decisions humanes i l'autonomia individual.

4. Responsabilitat, rendició de comptes i justícia

La distribució de responsabilitats entre dissenyadors, desenvolupadors, institucions i usuaris, i el problema de l'equitat en els sistemes d'IA.

5. Opacitat, transparència i confiança

Els reptes que plantegen els sistemes algorítmics opacs i el paper de la transparència en la construcció d'una IA fiable.

6. Explicabilitat i interpretabilitat

La importància de comprendre, justificar i comunicar com els sistemes d'IA produeixen els seus resultats.

7. Privacitat, protecció de dades i consentiment

Qüestions ètiques relacionades amb les dades personals, la vigilància, el consentiment i la governança de la informació.

Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Participació activa a classe	15	0,6	KM10
Debats a classe	20	0,8	SM10
Pràctiques i exercicis a classe	20	0,8	KM11, SM12

DE FORMA EXTRAORDINÀRIA, DURANT EL CURS 2026-2027 NO S'IMPARTIRÀ DOCÈNCIA EN AQUESTA ASSIGNATURA A CAUSA DE LA MODIFICACIÓ DEL GRAU EN INTEL·LIGÈNCIA ARTIFICIAL.

EL PROFESSOR OFERIRÀ ALS ESTUDIANTS MATRICULATS, LA POSSIBILITAT DE REUNIR-SE EN TUTORIES PER A DISCUTIR SOBRE EL PROCÉS D'APRENTATGE I AVALUACIÓ.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, perquè l'alumnat completi les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura.

Avaluació

Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Assaig filosòfic	20%	10	0,4	KM10, SM12
Examen	80%	10	0,4	KM11, KM12, SM10

De forma extraordinària, l'avaluació del curs 2026-27 es farà a partir d'un examen (80%), i d'un assaig (prèvia discussió del tema amb el professor).

L'examen s'haurà de preparar a partir de l'estudi i anàlisi crítica d'una selecció de lectures de la bibliografia.

Consultar el professor per a més informació.

Bibliografia

Binns, R. (2018, January). Fairness in machine learning: Lessons from political philosophy. In *Conference on fairness, accountability and transparency* (pp. 149-159). PMLR.

Boddington, P. (2023). *AI ethics*. Singapur: Springer International Publishing, 48.

Boge, F.J. (2022). Two Dimensions of Opacity and the Deep Learning Predicament. *Minds & Machines* 32, 43-75 . <https://doi.org/10.1007/s11023-021-09569-4>

Burell, J. (2016). *How the machine? thinks': Understanding opacity in machine learning algorithms*. *Big Data & Society*.

Creel, K.A. (2020). Transparency in complex computational systems. *Philosophy of Science* [Vol. 87, No. 4 \(October 2020\)](#), pp. 568-589 (22 pages)

Coeckelbergh, M. (2020). *AI ethics*. MIT press.

Cooper, A. F., Moss, E., Laufer, B., & Nissenbaum, H. (2022, June). Accountability in an algorithmic society: relationality, responsibility, and robustness in machine learning. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 864-876).

Diakopoulos, Nicholas. (2020). 'Transparency'. In Markus D. Dubber, Frank Pasquale, and Sunit Das (eds),

The Oxford Handbook of Ethics of AI (2020; online edn, Oxford Academic, 9 July 2020),

Dignum, V. (2020). Responsibility and artificial intelligence. *The oxford handbook of ethics of AI*, 4698, 215.

Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and machines*, 14(3), 349-379.

Lara, F., & Deckers, J. (Eds.). (2024). *Ethics of artificial intelligence*. Springer Nature.

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and information technology*, 6(3), 175-183.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big data & society*, 3(2), 2053951716679679.

Nissenbaum, H. (1996). Accountability in a computerized society. *Science and engineering ethics*, 2(1), 25-42.

Prunkl, C. (2024). Human Autonomy at Risk? An Analysis of the Challenges from AI: C. Prunkl. *Minds and Machines*, 34(3), 26.

Sandel, M. J. (2010). *Justice: What's the right thing to do?*. Macmillan.

Suresh, H., & Gutttag, J. (2021, October). A framework for understanding sources of harm throughout the machine learning life cycle. In *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1-9).

Tigard, D. W. (2021). Artificial moral responsibility: How we can and cannot hold machines responsible. *Cambridge Quarterly of Healthcare Ethics*, 30(3), 435-447.

Programari

No es fa servir cap programari per aquesta assignatura.

Grups i idiomes de l'assignatura

La informació proporcionada és provisional fins al 30 de novembre.
A partir d'aquesta data, podreu consultar l'idioma de cada grup a través d'aquest [enllaç](#). Per accedir a la informació, caldrà introduir el CODI de l'assignatura