

Guia docent

270425 - AESIA - Aspectes Ètics i Socials de la Intel·ligència Artificial

Última modificació: 03/02/2025

Unitat responsable: Facultat d'Informàtica de Barcelona
Unitat que imparteix: 270 - FIB - Facultat d'Informàtica de Barcelona.
Titulació: GRAU EN INTEL·LIGÈNCIA ARTIFICIAL (Pla 2021). (Assignatura obligatòria).
Curs: 2024 **Crèdits ECTS:** 6.0 **Idiomes:** Castellà

PROFESSORAT

Professorat responsable: CLAUDIO ULISES CORTÉS GARCÍA
Altres: Segon quadrimestre:
CLAUDIO ULISES CORTÉS GARCÍA - 10

CAPACITATS PRÈVIES

El curs no té requisits previs, sinó elements sòlids d'IA i lògica.

La lectura de l'ètica nicomaquea d'Aristòtil és recomanada

COMPETÈNCIES DE LA TITULACIÓ A LES QUALS CONTRIBUEIX L'ASSIGNATURA

Específiques:

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CE17. Desenvolupar i avaluar sistemes interactius i de presentació d'informació complexa i la seva aplicació a la resolució de problemes de disseny d'interacció persona-ordinador i persona-robot.

Genèriques:

CG5. Treballar en equips i projectes multidisciplinaris relacionats amb la intel·ligència artificial i la robòtica, interactuant fluidament amb enginyers/es i professionals d'altres disciplines.

CG6. Identificar oportunitats per a aplicacions innovadores de la intel·ligència artificial i la robòtica en entorns tecnològics en contínua evolució.

CG7. Interpretar i aplicar la legislació vigent, així com especificacions, reglaments i normes en l'àmbit de la intel·ligència artificial.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algoritmes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CG9. Afrontar nous reptes amb una visió àmplia de les possibilitats de la carrera professional en l'àmbit de la Intel·ligència Artificial. Desenvolupar l'activitat aplicant criteris de qualitat i millora contínua, i actuar amb rigor en el desenvolupament professional. Adaptar-se als canvis organitzatius o tecnològics. Treballar en situacions de carència d'informació i/o amb restriccions temporals i/o de recursos.

Transversals:

CT2. Sostenibilitat i Compromís Social. Conèixer i comprendre la complexitat dels fenòmens econòmics i socials típics de la societat del benestar; tenir capacitat per relacionar el benestar amb la globalització i la sostenibilitat; obtenir habilitats per utilitzar de forma equilibrada i compatible la tècnica, la tecnologia, l'economia i la sostenibilitat.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

Bàsiques:

CB1. Que els estudiants hagin demostrat posseir i comprendre 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 recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

METODOLOGIES DOCENTS

Implicació activa dels alumnes. Això pot implicar convidar els estudiants a presentar les seves idees actuals sobre l'ètica o participar en activitats de presa de decisions ètiques abans que el professor els presenti el contingut per primera vegada.

El professor selecciona els textos segons les necessitats del curs.

Utilitzarem aquests mètodes per portar el pensament ELSEC a l'aula, les conferències, l'ús de casos pràctics, els jocs de rol i les discussions en grup.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

- 1.Explorar el paper de l'ètica en la pràctica i la investigació de la intel·ligència artificial
- 2.Poder desenvolupar un conjunt de criteris ètics, legals, socioeconòmics, culturals i de gener per al desenvolupament d'aplicacions d'IA, i avaluar cadascuna de les aplicacions identificades pel que fa a aquests criteris.
- 4.Determinar les tecnologies basades en la IA, eines, arquitectures i algorismes que serien els més adequats per a les aplicacions industrials.

HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	30,0	20.00
Hores grup petit	30,0	20.00
Hores aprenentatge autònom	90,0	60.00

Dedicació total: 150 h

CONTINGUTS

Introducció. principis ètics

Descripció:

En aquest mòdul, considerarem les lleis, polítiques i principis ètics per regular i gestionar l'ús de la IA per al bé comú.

Aplicacions de la Intel·ligència Artificial per a la Salut i altres dominis rellevants:

Descripció:

En aquest mòdul, explorarem les aplicacions actuals i emergents de la IA a la salut i altres dominis rellevants. La seva legislació i aspectes ètics. En particular, parlarem dels procediments de presa de decisions responsables.

Construir un enfocament ELSEC per a l'ús de la IA

Descripció:

En aquest mòdul, parlarem de com mesures diferents de la llei i la política poden garantir que la IA millori el benestar humà.

Aplicacions de la Intel·ligència Artificial

Descripció:

En aquest mòdul, explorarem les aplicacions actuals i emergents de la IA en diverses aplicacions de domini. Les sol·licituds seran la base per a un debat en profunditat i el material a avaluar

ACTIVITATS

Introducció. Principis ètics

Objectius específics:

1

Competències relacionades:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre 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 recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en ambients o entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CE17. Desenvolupar i avaluar sistemes interactius i de presentació d'informació complexa i la seva aplicació a la resolució de problemes de disseny d'interacció persona-ordinador i persona-robot.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CG5. Treballar en equips i projectes multidisciplinaris relacionats amb la intel·ligència artificial i la robòtica, interactuant fluidament amb enginyers/es i professionals d'altres disciplines.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT2. Sostenibilitat i Compromís Social. Conèixer i comprendre la complexitat dels fenòmens econòmics i socials típics de la societat del benestar; tenir capacitat per relacionar el benestar amb la globalització i la sostenibilitat; obtenir habilitats per utilitzar de forma equilibrada i compatible la tècnica, la tecnologia, l'economia i la sostenibilitat.

Dedicació: 40h

Aprenentatge autònom: 24h

Grup gran/Teoria: 8h

Grup mitjà/Pràctiques: 8h

Aplicacions de la intel·ligència artificial per a la salut:

Objectius específics:

1, 4

Competències relacionades:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre 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 recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en ambients o entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CE17. Desenvolupar i avaluar sistemes interactius i de presentació d'informació complexa i la seva aplicació a la resolució de problemes de disseny d'interacció persona-ordinador i persona-robot.

CG7. Interpretar i aplicar la legislació vigent, així com especificacions, reglaments i normes en l'àmbit de la intel·ligència artificial.

CG9. Afrontar nous reptes amb una visió àmplia de les possibilitats de la carrera professional en l'àmbit de la Intel·ligència Artificial. Desenvolupar l'activitat aplicant criteris de qualitat i millora contínua, i actuar amb rigor en el desenvolupament professional. Adaptar-se als canvis organitzatius o tecnològics. Treballar en situacions de carència d'informació i/o amb restriccions temporals i/o de recursos.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CG5. Treballar en equips i projectes multidisciplinaris relacionats amb la intel·ligència artificial i la robòtica, interactuant fluidament amb enginyers/es i professionals d'altres disciplines.

CG6. Identificar oportunitats per a aplicacions innovadores de la intel·ligència artificial i la robòtica en entorns tecnològics en contínua evolució.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT2. Sostenibilitat i Compromís Social. Conèixer i comprendre la complexitat dels fenòmens econòmics i socials típics de la societat del benestar; tenir capacitat per relacionar el benestar amb la globalització i la sostenibilitat; obtenir habilitats per utilitzar de forma equilibrada i compatible la tècnica, la tecnologia, l'economia i la sostenibilitat.

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup mitjà/Pràctiques: 2h

Construir un enfocament ELSEC per a l'ús de la IA

Objectius específics:

1, 2, 4

Competències relacionades:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre 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 recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en ambients o entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CE17. Desenvolupar i avaluar sistemes interactius i de presentació d'informació complexa i la seva aplicació a la resolució de problemes de disseny d'interacció persona-ordinador i persona-robot.

CG7. Interpretar i aplicar la legislació vigent, així com especificacions, reglaments i normes en l'àmbit de la intel·ligència artificial.

CG9. Afrontar nous reptes amb una visió àmplia de les possibilitats de la carrera professional en l'àmbit de la Intel·ligència Artificial. Desenvolupar l'activitat aplicant criteris de qualitat i millora contínua, i actuar amb rigor en el desenvolupament professional. Adaptar-se als canvis organitzatius o tecnològics. Treballar en situacions de carència d'informació i/o amb restriccions temporals i/o de recursos.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CG5. Treballar en equips i projectes multidisciplinaris relacionats amb la intel·ligència artificial i la robòtica, interactuant fluidament amb enginyers/es i professionals d'altres disciplines.

CG6. Identificar oportunitats per a aplicacions innovadores de la intel·ligència artificial i la robòtica en entorns tecnològics en contínua evolució.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT2. Sostenibilitat i Compromís Social. Conèixer i comprendre la complexitat dels fenòmens econòmics i socials típics de la societat del benestar; tenir capacitat per relacionar el benestar amb la globalització i la sostenibilitat; obtenir habilitats per utilitzar de forma equilibrada i compatible la tècnica, la tecnologia, l'economia i la sostenibilitat.

Dedicació: 30h

Aprenentatge autònom: 18h

Grup gran/Teoria: 6h

Grup mitjà/Pràctiques: 6h

Aplicacions de la Intel·ligència Artificial

Objectius específics:

1, 2, 4

Competències relacionades:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CB1. Que els estudiants hagin demostrat posseir i comprendre 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 recolza en llibres de text avançats, inclou també alguns aspectes que impliquen coneixements procedents de l'avantguarda del seu camp d'estudi.

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en ambients o entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CE17. Desenvolupar i avaluar sistemes interactius i de presentació d'informació complexa i la seva aplicació a la resolució de problemes de disseny d'interacció persona-ordinador i persona-robot.

CG7. Interpretar i aplicar la legislació vigent, així com especificacions, reglaments i normes en l'àmbit de la intel·ligència artificial.

CG9. Afrontar nous reptes amb una visió àmplia de les possibilitats de la carrera professional en l'àmbit de la Intel·ligència Artificial. Desenvolupar l'activitat aplicant criteris de qualitat i millora contínua, i actuar amb rigor en el desenvolupament professional. Adaptar-se als canvis organitzatius o tecnològics. Treballar en situacions de carència d'informació i/o amb restriccions temporals i/o de recursos.

CG8. Observar un exercici ètic de la professió en totes les seves facetes, aplicant criteris ètics en el disseny de sistemes, algorismes, experiments, utilització de dades, d'acord amb els sistemes ètics recomanats pels organismes nacionals i internacionals, amb especial èmfasi en seguretat, robustesa, privacitat, transparència, traçabilitat, prevenció de biaixos (de raça, gènere, religió, territori, etc.) i respecte als drets humans.

CG5. Treballar en equips i projectes multidisciplinaris relacionats amb la intel·ligència artificial i la robòtica, interactuant fluidament amb enginyers/es i professionals d'altres disciplines.

CG6. Identificar oportunitats per a aplicacions innovadores de la intel·ligència artificial i la robòtica en entorns tecnològics en contínua evolució.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

CT2. Sostenibilitat i Compromís Social. Conèixer i comprendre la complexitat dels fenòmens econòmics i socials típics de la societat del benestar; tenir capacitat per relacionar el benestar amb la globalització i la sostenibilitat; obtenir habilitats per utilitzar de forma equilibrada i compatible la tècnica, la tecnologia, l'economia i la sostenibilitat.

Dedicació: 60h

Aprenentatge autònom: 36h

Grup gran/Teoria: 12h

Grup mitjà/Pràctiques: 12h

Conclusions

Objectius específics:

2

Competències relacionades:

CB3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes rellevants d'índole social, científica o ètica.

CE15. Adquirir, formalitzar i representar el coneixement humà en una forma computable per a la resolució de problemes mitjançant un sistema informàtic en qualsevol àmbit d'aplicació, particularment els relacionats amb aspectes de computació, percepció i actuació en ambients o entorns intel·ligents.

CE16. Dissenyar i avaluar interfícies persona-màquina que garanteixin l'accessibilitat i usabilitat dels sistemes, serveis i aplicacions informàtiques.

CT8. Perspectiva de gènere. Conèixer i comprendre, des del propi àmbit de la titulació, les desigualtats per raó de sexe i gènere a la societat; Integrar les diferents necessitats i preferències per raó de sexe i de gènere en el disseny de solucions i resolució de problemes.

CT4. Treball en equip. Ser capaç de treballar com a membre d'un equip interdisciplinari, ja sigui com un membre més o realitzant tasques de direcció, amb la finalitat de contribuir a desenvolupar projectes amb pragmatisme i sentit de la responsabilitat, assumint compromisos tenint en compte els recursos disponibles.

Dedicació: 10h

Aprenentatge autònom: 6h

Grup gran/Teoria: 2h

Grup mitjà/Pràctiques: 2h

SISTEMA DE QUALIFICACIÓ

Anem a demanar a tots els participants a través de un questionari d'avaluació en quina mesura cadascun d'ells creu que el seminari va complir amb els objectius.

El curs s'avaluarà amb el lliurament de tres treballs, amb valors proporcionals. Algun serà en equip i almenys un serà individual. Els temes canviaran cada curs.

BIBLIOGRAFIA

Bàsica:

- Dignum, Virginia. Responsible artificial intelligence : how to develop and use AI in a responsible way [en línia]. Springer, 2019 [Consulta: 11/02/2025]. Disponible a:

<https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=5972861>. ISBN 9783030303716.

- Russell, Stuart J. Human compatible : artificial intelligence and the problem of control. New York: Viking, 2019. ISBN 9780525558613.

- Véliz, Carissa. Privacy is power : why and how you should take back control of your data. Brooklyn ; London: Melville House, [2021]. ISBN 9781612199153.

- Coeckelbergh, Mark. AI Ethics [en línia]. Cambridge: MIT Press, 2020 [Consulta: 12/02/2025]. Disponible a: <https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=6142275>. ISBN 9780262357067.