

Degree programme	Type	Course
Bachelor in Artificial Intelligence	OB	3

Contact lecturer

Name : Roger Deulofeu Batllori

Email : roger.deulofeu@uab.cat

Prerequisites

There are no prerequisites for this course.

Group languages

You can consult this information at the [end](#) of the document.

Objectives

The aim of the course is to introduce students to the main ethical issues related to the development and use of artificial intelligence. Drawing on basic concepts from ethical theory, such as consequentialism, deontology, autonomy, responsibility and justice, the course analyses specific cases related to algorithmic bias, privacy, explainability and opacity, autonomy, automated decision-making and AI governance. The course seeks to provide students with the conceptual and argumentative tools needed to identify ethical dilemmas, critically assess different positions and reasonably justify their own professional decisions. This dimension is especially relevant in the training of engineers, since the tools they design, develop and implement may have high-impact ethical and social consequences, and it is essential that they are able to anticipate, assess and responsibly address their implications.

Learning outcomes

- KM10 (Identify privacy and intellectual property protection mechanisms within the framework of artificial intelligence systems.) Identify privacy and intellectual property protection mechanisms within the framework of artificial intelligence systems.
- KM11 (Describe the rules and regulations of artificial intelligence at national, European and international levels.) Describe the rules and regulations of artificial intelligence at national, European and international levels.
- KM12 (Describe the cases of civil liability in the use of artificial intelligence.) Describe the cases of civil liability in the use of artificial intelligence.
- SM10 (Analyse the impact of artificial intelligence applications on privacy and the right to privacy of individuals.) Analyse the impact of artificial intelligence applications on privacy and the right to privacy of individuals.
- SM12 (Determine the rules and regulations applicable to an artificial intelligence use case.) Determine the rules and regulations applicable to an artificial intelligence use case.

Contents

1. Introduction to AI Ethics: Why Does It Matter?

The social, political and professional relevance of ethical reflection on artificial intelligence.

2. Normative Ethical Frameworks

Main approaches for evaluating actions, decisions and technologies: consequentialism, deontology and virtue ethics.

3. Freedom, Agency and Autonomy

How AI systems can support, limit or reshape human decision-making and individual autonomy.

4. Responsibility, Accountability and Justice

The distribution of responsibility among designers, developers, institutions and users, and the problem of fairness in AI systems.

5. Opacity, Transparency and Trust

The challenges posed by opaque algorithmic systems and the role of transparency in building trustworthy AI.

6. Explainability and Interpretability

The importance of understanding, justifying and communicating how AI systems produce their outputs.

7. Privacy, Data Protection and Consent

Ethical issues related to personal data, surveillance, consent and the governance of information.

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Active participation in class	15	0.6	KM10
In class debates	20	0.8	SM10
Practical exercises in class	20	0.8	KM11, SM12

Exceptionally, during the 2026-2027 academic year, this course will not offer teaching lessons due to the modification of the Degree in Artificial Intelligence.

The instructor will offer enrolled students the possibility of meeting during office hours to discuss the assessment process.

Annotation: within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous assessment activities

Title	Weight	Hours	ECTS	Learning outcomes
Essay	20%	10	0.4	KM10, SM12
Exam	80%	10	0.4	KM11, KM12, SM10

Exceptionally, the assessment for the 2026-2027 academic year will consist of an exam (80%) and an essay, the topic of which must be discussed in advance with the instructor.

The exam should be prepared through the study and critical analysis of a selection of readings from the bibliography.

For further information, please consult the instructor.

Bibliography

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 systmes. *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.

Software

No software is needed for this course.

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this [link](#). To access the information, you will need to enter the course CODE