

WM9QD-15 Ethical Artificial Intelligence Implementation

25/26

Department

WMG

Level

Taught Postgraduate Level

Module leader

Awinder Kaur

Credit value

15

Module duration

4 weeks

Assessment

100% coursework

Study location

University of Warwick main campus, Coventry

Description

Introductory description

This module explores the various ethical issues involved in creating and using the Artificial Intelligence technologies. Students will examine how technology, ethics, law, and society interact, aiming to develop long-lasting Artificial Intelligence practices. The course includes theories, real-world examples, and hands-on approaches to make sure Artificial Intelligence systems are built and operated ethically. Students will actively participate in ongoing discussions and learn how to make well-informed and considerate decisions in the ever-evolving ethical field of Artificial Intelligence.

Module aims

The Ethical Artificial Intelligence Implementation module aims to:

1. Equip students with a thorough understanding of the ethical principles and dilemmas inherent in Artificial Intelligence technology.
2. Foster critical thinking about the broader implications of Artificial Intelligence on society, law, and industry.

3. Develop the ability to analyze and construct ethical frameworks and guidelines for Artificial Intelligence development and use.
4. Encourage practical application of ethical principles in the design, implementation, and assessment of Artificial Intelligence systems.
5. Promote collaboration and dialogue among students to explore diverse perspectives on ethical Artificial Intelligence.

Outline syllabus

This is an indicative module outline only to give an indication of the sort of topics that may be covered. Actual sessions held may differ.

1. Introduction to Ethics in Artificial Intelligence:
Overview of ethical theories and principles.
History and evolution of ethical considerations in Artificial Intelligence.
2. Artificial Intelligence and Societal Impact:
Case studies of Artificial Intelligence's impact on society, economy, and culture.
Discussion on Artificial Intelligence and job displacement, surveillance, and social justice.
3. Legal and Regulatory Frameworks:
Overview of current laws and regulations governing Artificial Intelligence.
International perspectives on Artificial Intelligence governance.
4. Bias, Fairness, and Transparency in Artificial Intelligence:
Understanding algorithmic bias and its consequences.
Strategies for ensuring fairness and transparency in Artificial Intelligence systems.
5. Privacy and Data Protection:
Exploration of privacy concerns in Artificial Intelligence.
Techniques for protecting data in Artificial Intelligence systems.
6. Developing Ethical Artificial Intelligence Solutions:
Tools and methodologies for ethical Artificial Intelligence development.
Case studies on successful ethical Artificial Intelligence implementation.
7. Collaborative Ethical Deliberation:
Group activities focusing on ethical decision-making in Artificial Intelligence development.
Debates and discussions on controversial Artificial Intelligence topics.

Learning outcomes

By the end of the module, students should be able to:

- Apply key ethical theories to evaluate and inform the development and implementation of Artificial Intelligence technology solutions.
- Critically assess the broader societal implications of Artificial Intelligence including cultural,

economic, and social dimensions.

- Construct and utilize ethical guidelines and frameworks to guide the development and the use of Artificial Intelligence systems.
- Devise comprehensive strategies for compliance with legal standards and regulatory requirements related to Artificial Intelligence in various jurisdictions.
- Collaborate effectively in teams to address complex ethical dilemmas in Artificial Intelligence, integrating diverse perspectives and fostering an inclusive environment for ethical deliberation.

Indicative reading list

Dubber, M.D., Pasquale, F. & Das, S. 2021;2020;, The Oxford handbook of ethics of AI, 1st edn, Oxford University Press, New York, NY, United States of America.

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E. & Herrera, F. 2023, "Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation", Information fusion, vol. 99, pp. 101896.

Russell, S.J. 2019, Human compatible: artificial intelligence and the problem of control, Allen Lane, UK.

International

Topics are of high international demand

Subject specific skills

Ethical Reasoning

Regulatory Knowledge

Bias Detection and Mitigation

Privacy Protection

Transferable skills

Critical Thinking

Communication

Collaboration

Problem-Solving

Study

Study time

Type	Required
Lectures	10 sessions of 1 hour (7%)
Seminars	20 sessions of 1 hour (13%)
Online learning (independent)	60 sessions of 1 hour (40%)
Assessment	60 hours (40%)
Total	150 hours

Private study description

No private study requirements defined for this module.

Costs

No further costs have been identified for this module.

Assessment

You must pass all assessment components to pass the module.

Assessment group A

	Weighting	Study time	Eligible for self-certification
Group Assessment	30%	18 hours	No
This assessment is a group project where students will work in teams to analyze a real-world case study involving ethical considerations in AI development and deployment.			
Peer Marking Process will be adopted in this assessment			
Assignment	70%	42 hours	Yes (extension)
This assessment is an individual essay where students are required to critically evaluate existing ethical frameworks for AI and discuss practical strategies for implementing these frameworks in AI development and deployment.			

Feedback on assessment

Verbal feedback for group assessment. Written feedback for assignment.

Availability

There is currently no information about the courses for which this module is core or optional.