

PHIL 15: ETHICS IN ARTIFICIAL INTELLIGENCE

Foothill College Course Outline of Record

Heading	Value
Effective Term:	Summer 2025
Units:	4
Hours:	4 lecture per week (48 total per quarter)
Advisory:	Not open to students with credit in HUMN 15.
Degree & Credit Status:	Degree-Applicable Credit Course
Foothill GE:	Area 3: Arts & Humanities
Transferable:	CSU/UC
Grade Type:	Letter Grade (Request for Pass/No Pass)
Repeatability:	Not Repeatable
Cross-Listed:	HUMN 15

Student Learning Outcomes

- Students will critically analyze and evaluate ethical dilemmas in artificial intelligence, applying ethical theories and frameworks to real-world AI applications.
- Students will demonstrate an understanding of fairness, accountability, transparency, and bias in AI systems and propose ethical solutions to mitigate potential harms in AI development

Description

Embark on a philosophical exploration of the ethical dimensions inherent in artificial intelligence (AI) in this engaging course offered through the philosophy department. The course delves deeply into the moral quandaries posed by AI technologies, examining issues such as algorithmic bias, the nature of consciousness in AI, ethical decision-making in machine learning, and the societal impact of automation from a philosophical standpoint. Through the lens of philosophical theories, students will critically analyze real-world AI applications and ethical frameworks and engage in stimulating debates to foster a nuanced understanding of the ethical implications of AI.

Course Objectives

The student will be able to:

- Learn philosophical foundations in moral theories and moral decision-making
- Build the practice of ethical inquiry into artificial intelligence (AI) emerging technologies
- Develop an understanding about the landscape and the scope of moral responsibility
- Explore the ethics of emerging technologies
- Build a connection between ethics and society
- Build an understanding about concepts of fairness, transparency, accountability, and equity in evaluating AI

Course Content

- The emergence of artificial intelligence (AI)
 - Definition of AI
 - The digital revolution and AI
 - The Turing Test and measuring AI intelligence
 - History of AI
- Introduction to ethics in AI
 - Definition of ethics and its significance in human decision-making
 - Introduction to major ethical theories (utilitarianism, deontology, virtue ethics, existentialism, ethical pluralism, ethical egoism)
 - Ethical frameworks for AI: consequentialist vs. deontological approaches
 - Ethical dilemmas
 - Overview of AI and its ethical implications
- Ethical considerations in AI research and development
 - Ethical guidelines and principles for AI research
 - Bias, equity, and fairness in AI algorithms, models, and datasets
 - Privacy and data protection in AI systems
 - Transparency and accountability in AI decision-making
 - Responsible AI
- AI and moral decision-making
 - Moral agency and responsibility in AI systems
 - Autonomous vehicles and the trolley problem
 - Moral dilemmas in AI healthcare applications
 - Ethical considerations in AI-driven decision support systems
- AI and social justice
 - Equity and fairness in AI applications
 - Algorithmic discrimination and stereotyping
 - Ethical considerations in protecting human rights in the age of AI
 - Societal implications of AI technologies on employment, inequality, and democracy
 - Ethical implications of AI for marginalized communities
 - Ethical design principles for promoting social justice in AI systems
- AI and creativity
 - Artificial intelligence and its applications in creative industries
 - The role of AI in generating, enhancing, and distributing creative content
 - Ethical implications of data collection and usage in AI-driven creative projects
 - Intellectual property rights in AI-generated works
 - Ethical considerations on ownership, attribution, and licensing of AI-created content
 - Legal and ethical challenges in determining authorship and copyright in collaborative AI projects
- Ethical governance of AI
 - Regulation and policy frameworks for AI ethics
 - Diversity and inclusion in AI development and deployment
 - International perspectives on AI ethics and governance
 - The role of industry, academia, and government in shaping ethical AI practices
 - Ethical considerations in AI policy making and implementation
- Ethical reflection and future directions
 - The role of humans in defining ethical boundaries for AI technologies

- b. The role of emotional intelligence and empathy in human interactions
- c. Fostering a relationship between AI and human flourishing
- d. Reflection on the evolving landscape of AI ethics and human values
- e. Ethical considerations in emerging AI technologies (e.g., AGI, neuro-technology)
- f. Ethical responsibilities of AI developers, researchers, and users
- g. Ethical activism and advocacy in the field of AI

- 3. Industry experience with guest lectures from ethicists that work for the AI industry.

Discipline(s)

Humanities or Philosophy

Lab Content

Not applicable.

Special Facilities and/or Equipment

When taught as an online section, students and faculty need ongoing and continuous internet and email access.

Method(s) of Evaluation

Methods of Evaluation may include but are not limited to the following:

Discussion
 Essay and short text
 Term paper
 Midterms (2)
 Final exam

Method(s) of Instruction

Methods of Instruction may include but are not limited to the following:

Discussion
 Lecture
 Project based learning
 Group projects

Representative Text(s) and Other Materials

Valor, Shanon. The Mirror AI: How to Reclaim Our Humanity in an Age of Machine Thinking. 2024.

O'Neil, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. 2017.

Christian, Brian. The Alignment Problem: Machine Learning and Human Values. 2020.

Though it is over 5 years old, the O'Neil work is the seminal text in the field.

Types and/or Examples of Required Reading, Writing, and Outside of Class Assignments

1. Midterm and final exams will be short essay questions and evaluation of case studies in the area of AI and ethics.
2. Projects with makerspace for design and prototype experience.