

PSY-432 Ethics and Human Rights in AI and Technology (3CHs)

Pre-requisite: None

Course Description:

The rapid rise in artificial intelligence (AI) has created many opportunities globally, from facilitating healthcare diagnoses to enabling human connections through social media and creating labor efficiencies through automated tasks. However, these rapid changes also raise profound ethical concerns. These arise from the potential AI systems have to embed biases, contribute to climate degradation, threaten human rights and more. This course aims to introduce undergraduate students to the ethical considerations surrounding the use of technology and artificial intelligence (AI). The course will cover ethical theories, principles, and frameworks that are relevant to the development and use of technology and AI. The course will also help students develop critical-thinking skills and ethical decision-making, required in the development and use of AI.

Course Objectives:

This course will facilitate students to

- Develop an understanding of ethical theories, principles, and frameworks that are relevant to the use of technology and AI.
- Develop an understanding of human rights and their importance in the digital age.
- Identify and analyze specific ethical issues related to technology and AI.
- Develop critical thinking skills for evaluating the ethical implications of technology and AI.
- Promote ethical decision-making in the development and use of technology and AI.

Course Outcomes:

As an outcome, students will be able to:

- Identify and analyze ethical issues related to technology and AI.
- Apply ethical theories, principles, and frameworks to the development and use of technology and AI.
- Evaluate the ethical implications of technology and AI.
- Make ethical decisions in the development and use of technology and AI.

Course Contents:

- Introduction to Ethics in Use of Technology and AI
- Ethical and Legal Frameworks for development and use of Technology and AI
 - Digital inclusion and reducing the digital divide.

- Emerging Human rights issues
- Human Rights Advocacy and Activism: Strategies and Tactics for Promoting Human Rights in the Development and use of Technology and AI
- Data Protection and Privacy in Technology and AI
- Bias and Discrimination in Technology and AI
- Accountability and transparency in technology and AI
- Autonomy and responsibility in technology and AI
- Artificial Intelligence and Human Values
- AI and human values.
- Ethics in Research and Development of Technology and AI
- Corporate social responsibility: Ethical Design and Development of AI
- Leadership and Decision-Making in Technology and AI: Role of Ethics

Textbooks:

- Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. The Cambridge Handbook of Artificial Intelligence.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

Reference Books

- Wallach, W., & Allen, C. (2010). Moral machines: Teaching robots right from wrong. Oxford University Press.
- Taddeo, M., & Floridi, L. (2018). How AI can be a force for good. Science, 361(6404), 751-752
- Amnesty International (2018). Human rights in the age of artificial intelligence. Amnesty International.
- United Nations (2011). Guiding principles on business and human rights: Implementing the United Nations “protect, respect and remedy” framework. United Nations.

Fung, A., Graham, M., & Weil, D. (2016). Full disclosure: The perils and promise of transparency. Cambridge University Press.