AI and Ethics

Group Presentation Assignment Brief

Introduction

This assignment requires students to work collaboratively to analyze and present on a specific ethical dilemma involving artificial intelligence. Through this presentation, you will develop critical thinking skills, engage with scholarly research, and articulate balanced perspectives on complex ethical issues emerging from AI technologies in society. This assignment connects theory to real-world applications and fosters dialogue on responsible AI development and implementation.

Case-Based Presentation Topics

Select **one** of the following scenarios as the focus of your group presentation:

Topic 1: Algorithmic Decision-Making in Healthcare Resource Allocation

Scenario: A large metropolitan hospital system has implemented an AI system to optimize patient triage and resource allocation in its emergency departments. The algorithm considers factors such as vital signs, medical history, geographic location, insurance status, and previous healthcare utilization patterns to assign priority scores. After six months of implementation, data analysis shows that patients from lower socioeconomic backgrounds consistently receive lower priority scores despite similar clinical presentations, resulting in longer wait times and potentially delayed critical care for these populations.

Guiding Questions:

- What ethical principles are in tension in this scenario, and how might they be balanced?
- How should responsibility for the algorithmic bias be distributed among system developers, healthcare administrators, and medical practitioners?
- What oversight mechanisms could be implemented to detect and mitigate such biases while maintaining the efficiency benefits of AI triage?
- How should transparency and explainability be balanced with the technical complexity of healthcare AI systems?

Recommended Sources:

- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
- Ibrahim, S. A., Charlson, M. E., & Neill, D. B. (2020). Big data analytics and the struggle for equity in health care: The promise and perils. *Health Affairs*, 39(2), 265-271.
- Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in artificial intelligence in health care. *Philosophy & Technology*, 34, 401-423.

- Cohen, I. G., Amarasingham, R., Shah, A., Xie, B., & Lo, B. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Affairs*, 33(7), 1139-1147.
- World Health Organization. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva: World Health Organization.

Topic 2: Facial Recognition Technology in Public Safety and Surveillance

Scenario: A mid-sized city has implemented a network of surveillance cameras equipped with facial recognition technology throughout its downtown area and transportation hubs. The system is connected to law enforcement databases and can identify individuals in real-time. City officials claim the system has reduced crime rates by 15% and expedited the apprehension of several suspects. However, civil liberties groups have documented multiple cases of misidentification, particularly among racial minorities, leading to wrongful stops and detentions. Meanwhile, local businesses report increased foot traffic as citizens perceive the area to be safer.

Guiding Questions:

- How should societies balance public safety benefits against individual privacy rights and civil liberties?
- What constitutes meaningful consent in public surveillance contexts, and is consent even possible?
- What technical and procedural safeguards should be required before implementing such technologies in public spaces?
- How should accountability be structured when automated systems and human decision-makers operate in tandem?
- What differential impacts might such systems have across diverse communities, and how should these be addressed?

Recommended Sources:

- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 81, 77-91.
- Selinger, E., & Hartzog, W. (2020). The inconsentability of facial surveillance. *Loyola Law Review*, 66(1), 33-77.
- Fussey, P., & Murray, D. (2019). *Independent report on the London Metropolitan Police Service's trial of live facial recognition technology*. University of Essex Human Rights Centre.
- Brayne, S. (2021). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.
- European Union Agency for Fundamental Rights. (2019). *Facial recognition technology: Fundamental rights considerations in the context of law enforcement*. Luxembourg: Publications Office of the European Union.

Topic 3: Autonomous Vehicles and Ethical Decision-Making

Scenario: An automotive manufacturer is finalizing the programming for its fleet of fully autonomous vehicles scheduled for commercial release next year. The engineering team must develop algorithms that govern the vehicle's responses in unavoidable accident scenarios—situations where harm cannot be eliminated, only distributed differently among various parties. The company's leadership is divided over whether these decision frameworks should prioritize passenger safety above all else, follow a utilitarian approach that minimizes total harm, adhere to established rules of the road regardless of consequences, or implement some hybrid approach. Meanwhile, regulatory bodies have not yet established clear legal standards for such scenarios.

Guiding Questions:

- What ethical frameworks are most appropriate for programming life-or-death decisions in autonomous systems?
- Who should determine the ethical principles embedded in autonomous vehicle algorithms—manufacturers, regulators, ethicists, or the general public?
- How transparent should manufacturers be about the decision-making frameworks their vehicles employ?
- How should responsibility and liability be distributed among manufacturers, software developers, owners, and passengers when autonomous systems cause harm?
- What implications might different decision frameworks have for consumer adoption, insurance models, and public trust?

Recommended Sources:

- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, 563(7729), 59-64.
- Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: An applied trolley problem? *Ethical Theory and Moral Practice*, 19(5), 1275-1289.
- Himmelreich, J. (2018). Never mind the trolley: The ethics of autonomous vehicles in mundane situations. *Ethical Theory and Moral Practice*, 21(3), 669-684.
- Sparrow, R., & Howard, M. (2017). When human beings are like drunk robots: Driverless vehicles, ethics, and the future of transport. *Transportation Research Part C: Emerging Technologies*, 80, 206-215.
- Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576.

Your Task

For your chosen scenario, your group must prepare and deliver a comprehensive presentation that:

1. **Analyzes and explains the scenario** in depth, providing relevant technical context and stakeholder perspectives
2. **Identifies key ethical dilemmas** present in the scenario, articulating the underlying values, principles, and rights at stake
3. **Compares stakeholder perspectives** by analyzing the positions, interests, and concerns of all major stakeholders involved
4. **Formulates a clear and justified position** on how the ethical dilemmas might be addressed, supported by reasoned arguments and scholarly sources
5. **Concludes with 1-2 thought-provoking questions** designed to stimulate class discussion following your presentation

Assignment Details

Group Size: 3-4 students per group

Duration: 8-10 minutes for oral presentation, followed by 5 minutes of class discussion

Format: PowerPoint or similar presentation slides with accompanying speaker notes

Submission Deadline: June 1, 2025, 11:59 PM

Submission Method: Upload presentation file (with embedded speaker notes) to the designated assignment folder on Moodle

Use of Generative AI:

This course acknowledges the increasing availability and utility of generative AI tools in academic work. The following guidelines apply to this assignment:

- **Permitted uses:** You may use generative AI tools for:
  - Brainstorming initial ideas and approaches
  - Generating or enhancing visual elements in your presentation
  - Refining language and structure of your content
  - Identifying potential sources (although you must independently verify their existence and relevance)
  - Checking grammar and formatting
- **Transparency requirements:** If you use generative AI in any capacity, you must:
  - Include a dedicated slide at the end of your presentation titled "AI Usage Statement"
  - Specify which tools were used (e.g., ChatGPT, DALL-E, Midjourney)
  - Clearly describe how each tool was used and for which specific components
  - Explain how you critically evaluated and refined any AI-generated content

- **Limitations:** The following aspects of your work must primarily reflect your group's original thinking:
  - Critical analysis of ethical dilemmas
  - Evaluation of stakeholder perspectives
  - Development of reasoned arguments
  - Formulation of your group's position

AI usage, when properly disclosed, will not negatively impact your grade. This policy aims to promote transparency and responsible AI use while maintaining academic integrity.

Evaluation Criteria

Criterion	Description	Weight
Analysis of Ethical Issues	Depth and accuracy of ethical analysis; identification of relevant ethical frameworks and principles; nuanced understanding of tensions and tradeoffs	25%
Research Quality	Integration of scholarly sources; evidence of engagement with current literature; appropriate citation of sources; critical evaluation of sources	20%
Balanced Perspective	Fair representation of diverse stakeholder viewpoints; consideration of technical, social, legal, and ethical dimensions	15%
Argumentation	Clarity and coherence of position; quality of supporting reasoning; anticipation of counterarguments	15%
Presentation Quality	Organization; visual design; time management; engagement with audience; quality of discussion questions	15%
Collaborative Process	Evidence of effective teamwork; equal distribution of speaking roles; cohesive integration of individual contributions	10%

Note: All group members will receive the same grade unless documented concerns about inequitable contribution are raised before the submission deadline.