

UNIT 7

Ethical and Societal Implications of Artificial Intelligence

Artificial Intelligence (AI) is rapidly advancing and being applied in many fields, including agriculture, healthcare, weather forecasting, and computer vision, as we discussed in previous lectures. **While AI brings many benefits, we also need to consider the concerns about its possible negative effects on society.**

Princeton Review highlights **four major ethical and societal concerns with AI use¹:**

1. Bias and Discrimination

- 1.1. In Unit 2 (Machine Learning Concepts), we learned that model predictions depend on the training data. **If the data is biased, the model will replicate those biases, leading to unfair outcomes in areas like hiring, justice, healthcare, etc., with serious consequences.**
- 1.2. Examples of AI Bias

1.2.1. In Hiring

- 1.2.1.1. New America reports that recruiters increasingly use AI-driven algorithmic screening tools to classify job seekers².
- 1.2.1.2. A study highlighted in the article found that **candidates with distinctively Black names were less likely to receive responses compared to those with distinctively White names.**
- 1.2.1.3. Resume language also affects algorithmic assessments: **men often use assertive terms like "leader" and "competitive," while women use words such as "support" and "interpersonal," potentially leading algorithms to favor men as more qualified.**

1.2.2. In Education

- 1.2.2.1. Axios reported that the British government canceled A-level exams due to the pandemic and **relied on an AI algorithm to predict student results based on past school performance³.**
- 1.2.2.2. This decision led to significant grade disparities, with **students from less-advantaged schools receiving lower predicted grades compared to those from wealthier institutions, potentially impacting their future opportunities.**

2. Accountability and Regulation

- 2.1. As AI increasingly impacts decision-making, it's crucial to establish guidelines promoting **transparency, fairness, and accountability**
- 2.2. Efforts promoting responsible AI include:

2.2.1. University of the Philippines' Principles for Responsible AI⁴

2.2.1.1. Transparency

"People should be able to understand how AI systems work. Individuals should be informed if AI-enabled tools are used. To the extent possible, the methods should be explainable. Finally, users should be able to understand AI-based outcomes and identify ways to seek remedies to harms that they may cause." (Principle No. 7)

2.2.1.2. Fairness

"AI should be evaluated for gender bias, other forms of unfairness, and all forms of discrimination, especially in the data, models, and algorithms that are used. Collaborative procedures should be in place to actively identify, mitigate, and remedy these harms. AI developers should be mindful of its unintended consequences." (Principle No. 8)

2.2.1.3. Accountability

"Individuals, groups, departments, institutes, colleges, and constituent universities involved in the development, deployment, and use of AI **must take responsibility for the consequences of their actions.**" (Principle No. 5)

2.2.2. UNESCO Ethics of Artificial Intelligence⁵

2.2.2.1. Transparency and Explainability

"The ethical deployment of AI systems depends on their transparency & explainability (T&E). The level of T&E should be appropriate to the context, as there may be tensions between T&E and other principles such as privacy, safety and security."

2.2.2.2. Fairness and Non-Discrimination

"AI actors should promote social justice, fairness, and non-discrimination while taking an **inclusive approach to ensure AI's benefits are accessible to all.**"

2.2.2.3. Responsibility and Accountability

"AI systems **should be auditable and traceable.** There should be oversight, impact assessment, audit and due diligence mechanisms in place to avoid conflicts with human rights norms and threats to environmental wellbeing."

Note: Check the two guidelines on responsible AI for more details.

3. Privacy and Data Security

- 3.1. The reliance of AI on vast amounts of data raises significant concerns regarding data privacy and security, as **personal information may be used without consent.**
- 3.2. A Washington Post article highlights Sora, an AI by OpenAI that generates videos from text, using extensive datasets, including licensed and **publicly scraped content from the internet.** This raises legal concerns about potential copyright violations⁷.
- 3.2.1. **Publicly accessible data online could have contributed to the tool's development.**

Note: Consider using Facecheck.ID, a search engine that allows you to search the internet using a photo, to check if your image is in their database.

4. Economic Disruption and Job Displacement

- 4.1. While AI has simplified many tasks, it poses risks to economies and potential job losses.
- 4.2. Ipsos reports that **73% of Filipinos believe AI could lead to job losses, compared to a global average of 64%.** Additionally, **48% of Filipinos think AI is unlikely to create new jobs, above the global average of 44%**⁸.
- 4.3. GlobalData notes an **increase in AI-related job postings in Southeast Asia, indicating the need for Filipinos to adapt to an AI-driven job market.**
- 4.4. **House Bill No. 9488, titled the "Protection of Labor Against AI Automation Act,"** has been proposed by Quezon City Rep. Arjo Atayde to **empower rather than replace Filipino workers with AI.**
- 4.5. Forbes discusses AI's impact on coding, noting that **while AI can generate code, human expertise in logic, problem-solving, and programming is still essential to ensure code quality and prevent poorly structured "spaghetti code"**⁹.

References

- [1] <https://www.princetonreview.com/ai-education/ethical-and-social-implications-of-ai-use>
- [2] <https://www.newamerica.org/oti/blog/ai-discrimination-in-hiring-and-what-we-can-do-about-it>
- [3] <https://wwwaxios.com/2020/08/19/england-exams-algorithm-grading>
- [4] <https://up.edu.ph/up-principles-for-responsible-artificial-intelligence>
- [5] <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- [6] <https://www.washingtonpost.com/technology/interactive/2024/ai-video-sora-openai-flaws>
- [7] <https://www.bworldonline.com/technology/2024/01/04/566748/filipinos-expect-data-leaks-job-losses-from-ai-this-year>
- [8] <https://www.philstar.com/nation/2023/12/29/2322173/bill-aims-protect-filipino-jobs-ai>
- [9] <https://www.forbes.com/advisor/education/it-and-tech/artificial-intelligence-in-school>

Note: The sources for this chapter were compiled manually. However, an AI tool was used to summarize and paraphrase them for easier reading.