Rezwan Ahmed
CSC 448
November 15, 2023

Midterm Paper: Ethics in AI

Abstract:

The rapid integration of Artificial Intelligence (AI) into various aspects of daily life raises a profound ethical concern. Algorithmic bias, stemming from biased training data, has been a significant issue, leading to discriminatory outcomes in many areas like healthcare and recruitment. The lack of transparency in AI decision-making processes poses challenges to accountability, fairness, and ethical responsibility, particularly in complex fields such as healthcare and criminal justice. Privacy concerns are escalating as AI systems require vast amounts of personal data for learning, prompting a reevaluation of data ownership and privacy laws. The deployment of AI in surveillance systems and weapon technology further amplifies ethical questions, emphasizing the need for careful consideration of privacy rights and the potential consequences of autonomous decision-making in military applications.

In the present era, the rapid progress in Artificial Intelligence (AI) technology is reshaping our daily lives. AI has become omnipresent, integrated into various aspects in our society such as self-driving cars and virtual assistants. Our dependence on AI is growing more than ever. However, this increasing presence of AI in our daily lives also raises significant ethical questions. As AI becomes more pervasive, ethical concerns become more

apparent, leading us to consider the moral implications of its integration into various aspects of society.

One of the primary ethical concerns regarding AI is the presence of bias in AI algorithms. Algorithmic bias occurs when algorithms make decisions that systematically disadvantage certain groups of people. It can have disastrous consequences when applied to key areas such as healthcare, criminal justice, and credit scoring [1]. There have already been many such cases. For instance, scientists investigated a widely used healthcare algorithm and found that it severely underestimated the needs of black patients, leading them to have significantly less care [1]. Another example of algorithmic bias is Amazon's recruitment Algorithm. This algorithm was biased against applications of women and penalized resumes from women or any resumes using the word "women(s)" [2]. Often, algorithmic bias is unintentional, and biases can be present for many reasons. One of the main reasons for bias to be present is due to the data that was used to train the algorithm [3]. If the data that was used to train the AI was biased, then it is likely that the AI will become biased. Going back to Amazon's recruitment algorithm, the algorithm was trained using ten years of employment history data, but the data reflected a male-dominated workforce, and so the algorithm was biased toward males [2]. Efforts to eliminate bias in AI are underway, and while complete eradication is challenging, it can certainly be reduced by ensuring that the training data we use is carefully selected to represent diverse populations.

Another critical ethical concern is the lack of transparency and explainability in AI algorithms. Many AI systems, especially those based on complex machine learning models, operate as "black boxes," making it challenging to understand how they reach specific decisions [4]. This poses significant challenges for accountability and opens the door to potentially unjust or biased outcomes. In complex decision-making processes, such as those in healthcare and criminal justice, the inability to explain the reasoning behind AI-generated decisions may undermine public trust and confidence in the systems.

The lack of transparency not only hinders accountability but also raises fundamental questions about fairness and ethical responsibility. When individuals are subjected to decisions made by AI, such as medical diagnoses or legal judgments, they have the right to know how those decisions were reached. Without transparency, it becomes difficult to identify instances of bias or errors that may disproportionately impact certain demographic groups.

In healthcare, for example, where AI is increasingly utilized for diagnostic purposes and treatment recommendations, the opacity of algorithms can lead to disparities in patient care. If the decision-making process is not transparent, medical professionals may struggle to explain and validate AI-driven recommendations, potentially hindering collaboration between healthcare providers and their patients. This lack of clarity could also compromise the patient's right to be informed about their healthcare choices.

Privacy concerns related to AI have become increasingly prominent as technology continues to advance. AI systems rely on vast amounts of data to learn and make predictions, raising questions about the protection of personal information. When a user inputs their personal information into an AI model, there's the possibility that this data will become part of the model's future training dataset. When this happens, there is a possibility this data can show up as outputs to other users' queries, which is a particularly big issue if users have input sensitive data into the system [5].

Moreover, exploring the evolving landscape of data privacy laws and regulations globally is very important. For instance, the General Data Protection Regulation (GDPR) in Europe has significantly impacted how AI applications handle personal data [6]. Examining the concept of "data ownership" and how individuals can retain control over their personal information when interacting with AI systems adds another layer of ethical consideration. The ethical implications of data breaches and the potential misuse of personal data collected by AI systems are profound. Actual instances of these events and their consequences highlight the importance of dealing with privacy worries in the era of AI.

There are already AI systems that can accurately identify people. For example, in China, surveillance systems can accurately identify individuals and other objects. While this technology has practical applications, such as enhancing airport security by identifying individuals and their flight plans, it also raises significant privacy concerns [7]. Constant tracking and data collection by governments and other entities using this technology can

infringe upon individuals' privacy rights. The risk of these systems getting hacked or the data falling into the wrong hands adds another layer of concern, as the private information of many individuals could be leaked.

Furthermore, as AI technology advances, their integration into weapon systems raises important ethical questions. Autonomous weapons guided by AI for decision-making, bring up concerns beyond typical warfare standards. The use of AI in weapons means there might be less human oversight in crucial decisions, leading to worries about accountability and ethical guidelines [8]. Putting AI in military tools, like drones or self-governing combat systems, introduces the risk of unintended consequences and potential escalation of conflicts. For example, AI weapon systems use sensors to detect allies and targets. What happens if these systems mistakenly identify allies or civilians as enemy targets? Furthermore, in scenarios where these weapons struggle to identify targets due to adverse weather conditions or damaged sensors, their behavior remains unclear [9].

To sum it up, as AI becomes a bigger part of our lives, we need to think carefully about the ethical issues it brings. We must work to make sure AI doesn't favor certain groups and is clear about how it makes decisions. Protecting people's privacy is crucial, and we should keep an eye on changing laws. Using AI for surveillance and weaponry needs careful thought to avoid privacy breaches and unintended consequences. By talking about these concerns, making good rules, and developing AI responsibly, we can make sure AI benefits everyone and aligns with what's right for society.

References

[1] Friis, S., & Riley, J. (2023, September 29). *Eliminating algorithmic bias is just the beginning of equitable ai*. Harvard Business Review. https://hbr.org/2023/09/eliminating-algorithmic-bias-is-just-the-beginning-of-equitable-ai

[2] Needle, F. (2023, June 6). *Can artificial intelligence write better email subject lines than humans?*. HubSpot Blog. https://blog.hubspot.com/marketing/ai-email-subject-lines

[3] Wu, M. (2022, January 20). *Dealing with AI biases, part one*. CMSWire.com. https://www.cmswire.com/digital-experience/dealing-with-ai-biases-part-1-acknowledging-the-bias/

[4] Rudin, C., & Radin, J. (2019, November 22). *Why are we using black box models in AI when we don't need to? A lesson from an explainable AI competition*. Harvard Data Science Review. https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/8

[5] Hiter, S. (2023, September 20). *Ai and privacy issues: What you need to know*. eWEEK. https://www.eweek.com/artificial-intelligence/ai-privacy-issues/

[6] Wolford, B. (2023, September 14). *What is GDPR, the EU's new Data Protection Law?* What is GDPR, the EU's new data protection law? https://gdpr.eu/what-is-gdpr/#:~:text=The%20General%20Data%20Protection%20Regulation,to%20people%20in%20the%20EU.

[7] Grothaus, M. (2019, March 25). *China's airport facial recognition kiosks should make us fear for our privacy*. FastCompany. https://www.fastcompany.com/90324512/chinas-airport-facial-recognition-kiosks-should-make-us-fear-for-our-privacy

[8] Etzioni, A., & Etzioni, O. (2017, June). *Pros and cons of Autonomous Weapons Systems*. Army University Press. https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/

[9] Walsh, T. (2016, February 9). *Autonomous Weapons Open Letter: AI & Robotics researchers*. Future of Life Institute. https://futureoflife.org/open-letter/open-letter-autonomous-weapons-ai-robotics/