This week, you have spent some time exploring different ethical and privacy implications of different AI technologies. In this discussion, you will be asked to explore hidden biases in artificial intelligence, which have been a hot topic in the news. Before beginning your discussion, read through *all* of the following articles. You are also encouraged to do additional research on your own.

* [Why Algorithms Can Be Racist and Sexist](https://www.vox.com/recode/2020/2/18/21121286/algorithms-bias-discrimination-facial-recognition-transparency)
* [Overcome and Prevent Bias in AI](https://web.archive.org/web/20240710160510/https:/f8federal.com/overcome-and-prevent-ai-bias/)
* [How to Keep Bias Out of Your AI Models](https://customerthink.com/how-to-keep-bias-out-of-your-ai-models/)

For your initial post, write a response of 2–3 paragraphs that addresses *each* of the following prompts. Your answer must include references to resources used, properly cited in APA format.

* Provide 1–2 current examples of hidden bias encountered in popular applications.
* Define some of the industry efforts to reduce bias (such as responsible research and innovation [RRI], ethics by design, and right to be forgotten).
* How can you apply some of these principles to reduce hidden bias in the applications you described? How do these relate to the efforts currently underway?

In your responses to your peers, evaluate their proposed solutions to reduce bias and propose alternatives. Remember to respond to *at least two* of your peers. Be sure to include cited sources to support your points.

I believe it’s very possible to have hidden biases in A.I. Definitely a major concern, as it can lead to discrimination or reinforcement of societal inequities. The articles in our prop gave us a bases for this, but I wanted to look at another article.

An example of this is facial recognition software. From your phone, windows login, NBA 2k face scan, all are known to perform quite poorly for those of darker skin tones. Research by the MIT Media Lab revealed that error rates for these groups are significantly higher due to insufficient diversity in training data (Buolamwini & Gebru, 2018). Another example is biased recruitment tools. Another issue I found while browsing was Amazon’s experimental hiring system that used an A.I. to sort through applications. Nothing wrong in that, as most jobs used some sort of A.I. to sort through thousands of applicants. But Amazon’s system had shown discriminatory habits towards women, by placing lower priority to their applications. This was expanded to include any reference of the word “woman” in the resume (Dastin, 2018).

As a way to combat these issues that have occurred & potentially prevent future issues, industries that use A.I. are adapting strategies to reduce bias. Responsible Research & Innovation is focusing on transparency, inclusivity, & accountability in their A.I. training programs. "Ethics by design" ensures ethical considerations are integrated into the design phase, promoting diverse teams, inclusive datasets, and rigorous testing. The "right to be forgotten," part of laws like the EU’s GDPR, allows individuals to remove their personal data from systems, reducing reliance on outdated or irrelevant information.

Application of those principles can directly address the biases from above. Facial recognition, ethics by design would involve using diverse datasets. Regular audits could further ensure fairness across different racial groups. Another example is for recruitment tools encourages external reviews of algorithms. This way, a wide spectrum of ethnicities could have input of opinion.

Silipo, R. (2020, March 3). How to keep bias out of your AI models | CustomerThink. Customerthink.com. <https://customerthink.com/how-to-keep-bias-out-of-your-ai-models/>

Janet. (2021, January 15). Overcome and Prevent Bias in AI - Figure Eight Federal. Figure Eight Federal. [https://web.archive.org/web/20240710160510/https://f8federal.com/overcome-and-prevent-ai-bias/](https://web.archive.org/web/20240710160510/https:/f8federal.com/overcome-and-prevent-ai-bias/)

Heilweil, R. (2020, February 18). Why algorithms can be racist and sexist. Vox. <https://www.vox.com/recode/2020/2/18/21121286/algorithms-bias-discrimination-facial-recognition-transparency>

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1-15. Retrieved from <https://proceedings.mlr.press/v81/buolamwini18a.html>

Dastin, J. (2018, October 10). *Insight - Amazon scraps secret AI recruiting tool that showed bias against women | reuters*. Insight - Amazon scraps secret AI recruiting tool that showed bias against women. <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/>

Evening Daveon,

Some examples of biases in facial recognition & how hiring algorithms show how pervasive hidden biases can be programmed in A.I. I believe that focusing on diversive training data & ensuring audits are conducted in a fair manner will be essential to addressing these issues. Regularly auditing hiring tools to detect and correct imbalances is another vital practice, exemplified by the industry's use of fairness metrics to mitigate bias. Historical data often reinforces systemic inequities, as seen in certain hiring algorithms, underscoring the need for consistent bias checks. Coupling these efforts with a strong focus on transparency paves the way for more equitable and accountable AI systems.

Hi there Raheem,

From our readings, I’ve determined that biases can be found in both facial recognition & different recruitment processes. Shining light on these issues allows us to highlight the biases embedded in A.I. systems. In the hiring process, biases are problematic regardless of whether AI is used. While automating resume evaluations can improve efficiency for companies handling large volumes of applications, it raises concerns about fairness. One example, A.I. tools have been known for filtering out candidates based on factors not concerning their jobs, such as gender rather than qualifications && ability to perform the job.

To address these issues, it is essential to focus on diverse training data and ensure fairness through regular audits. Auditing hiring tools to detect and correct imbalances, as seen in the industry’s use of fairness metrics, is a vital step in mitigating bias. Coupled with consistent bias checks and a strong commitment to transparency, these practices can help pave the way for more equitable and accountable AI systems.