

Module 1: One Page Reflection

David Bishop

Question 2: Define a trust, ethics, or bias issue within the context of an example (real or fictional) and propose design solutions to mitigate risk.

Response:

The rise of Artificial Intelligence (AI) systems in decision-making processes has brought significant issues of trust, ethics, and bias. One notable example that I have been noticing is the lack of transparency within AI models. Most models are proprietary from companies meaning that the public as well as many law enforcement agencies cannot determine how a model was trained or its predictions generated. This black box approach undermines any trust people may have in the system (if the person is skeptical).

Transparency is not just a trust issue; it is also a critical factor in identifying and addressing biases within AI models. Consider facial recognition as an example. “In 2015, Google Photos came under fire after its AI-based search tool pulled images of Black people when the term “gorilla” was searched.” [1] This incident likely did not result from intentional malice but rather from biases embedded within the training data or model architecture. Such outcomes might have been avoided with greater transparency in how the model made decisions and what data it was trained on. This is such a problem there are whole sites dedicated to different AI ethical issues [2].

This example highlights just one of the many challenges AI presents in terms of trust, ethics, and bias. If we are not careful with this emerging technology on how and where we utilize it, we will end up creating systems that have biases that can horribly affect the populace. If we focus to make AI more transparent, auditing training data, and possibly even embedding ethical

considerations into the design process we can develop AI systems that are both effective and non-problematic.

References:

1. Johnson, Arianna. "Racism and AI: Here's How It's Been Criticized for Amplifying Bias." Forbes, Forbes Magazine, 4 Oct. 2023, www.forbes.com/sites/ariannajohnson/2023/05/25/racism-and-ai-heres-how-its-been-criticized-for-amplifying-bias/.
2. "Artificial Intelligence: Examples of Ethical Dilemmas." UNESCO.Org, www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases. Accessed 3 Sept. 2024.