Name: Danny Rovi

Fall 2025

**Instructions:** Watch the documentary “Coded Bias,” accessible here with your SU login: <https://libezproxy.syr.edu/login?url=https://video.alexanderstreet.com/p/EqZYn56RW>. Answer the questions below when you’re done watching. Be ready to discuss the documentary next week.

**Submission:** Submit on Blackboard by 11:59p ET Thursday 9/11/25. Double-check that you’ve uploaded it correctly before the deadline; if you upload something I cannot see, this assignment will be considered late. Generative AI use is not allowed for this assignment.

1. What were you surprised by in this documentary? Tell me about a scene or a quote that taught you something you didn’t know before.

One part of the movie that really surprised me was when Joy showed that some facial recognition systems couldn’t even detect her face until she put on a plain white mask. I always thought that this kind of technology was extremely advanced, efficient, and accurate, so it shocked me to see how badly it failed on women and people with darker skin. That made me realize that algorithms are in fact not neutral. They reflect the people who make them, which includes their biases. I learned that technology isn’t always fair, and it can make inequality worse if no one is paying attention.

2. How have you seen AI depicted in pop culture or mainstream media? What are some differences you can note in how algorithms are depicted there and how they were depicted in this documentary?

In modern day pop culture, I usually see AI shown as something really advanced or even scary, like robots in movies that take over the world or personal assistants that seem smarter than humans. Shows and movies often make AI seem almost perfect, like it can do anything without making mistakes. However, Coded Bias presented AI in a different sense. The documentary showed that algorithms are far from perfect; they can be biased and unfair, and sometimes even detrimental to society. The biggest difference is that mainstream media often makes AI look powerful and futuristic, while the documentary focused on the flaws and real human impact of these systems right now in the current moment.

3. Review the Constantaras reading. Then, take your time to examine and read through some of this database: <https://incidentdatabase.ai/apps/incidents/>. What are two questions you’d have if you were “interviewing” this dataset in order to find a story? What is one researched, provable/disprovable and specific hypothesis you have about this dataset?

Question One: What is the probability that chatbots produce sexist, racist, or homophobic language compared to finding alternate ways to respond to hateful prompts?

Question Two: What specific defenses are installed to try and prevent chatbots from producing this language? Are there any? If so, what are they?

Hypothesis: If it is possible for an AI chatbot to produce biased “hate speech”, should there be greater legal action and regulations to ensure that bots remain unbiased and only produce factual information?