

Fall 2022

Jolly Ogbole

Stance on Article Titled: "These Robots were trained on AI. [They became racist and sexist.](#)"

While it is a fact that robots have become racist and sexist, I take the position that the adverse effects of racism and sexism exhibited by robots and Artificial Intelligence (AI) systems is a reflection of the biases and prejudices of their respective developers, and the quality of the data by which they were developed, designed and engineered. Robots are not inherently racists or sexists; the quality of output is determined by the quality of input (data); the severe lack of representation and diversity in the computing and artificial intelligence industry has contributed to the problems of so-called racists or sexists AI systems in society today. To make outstanding claims that robots are racists and sexists is to abdicate our responsibility and deny accountability for the issues that we create, even when done unconsciously. It's akin to calling out the symptoms as opposed to the root causes of a problem.

Robots and Artificial Intelligence systems are not inherently racists and sexists because unlike humans, they do not operate with preexisting notions of the world. They are human-made systems that operate according to the algorithms upon which they have been designed and programmed. Despite its seemingly unfathomable and sophisticated portrayal, the reality of AI and machine learning is that they are systems of equations. These systems take in data as inputs and outputs patterns and results that help to answer specific questions under specified constraints. Forms of artificial intelligence such as machine learning systems might be able to learn or grow, but that is within the limits set by the programmer and scientists. Since AI is developed by humans subject to bias, the not so surprising consequence of this operation is that the people who develop them can translate their own conscious and unconscious world views to these AI systems. Therefore AI cannot be said to be objective or operate at neutrality. By similar

effects, robots or AI do not have any intrinsic morals, consciousness or ethics in and of themselves, other than those imparted into them by developers.

The quality of output is determined by the quality of input (data). The principle and concept of garbage in, garbage out (GIGO) has always been the gold standard for building anything worthwhile in the computing industry. In plain terms, this is the concept that flawed raw materials (input data) will produce poor quality product output. Data is the building blocks for robotics and artificial intelligence systems Implementation. AI and machine learning is only as good as the data that it uses and the way that it learns. If the data sets are biased, skewed and full of errors, the results will also set us off in the wrong direction. Most data issues are not necessarily about the quantity of data but the quality of data that is fed into the AI model. Bias in data is the type of error that occurs when certain elements of a dataset are overweighted or overrepresented. More often than not, you would find the overrepresentation of minority groups in the datasets that are used to train AI models that predict crime or high level of recidivism in AI and machine learning models. These are factual events with real life consequences for humans in the real world. Why have AI developers ignored this core principle for so long or is it a deliberate act ? The answer might lie in my final argument. Nonetheless, given that data is at the core of AI systems and has far reaching effects on modern society, it behooves developers to prioritize data quality in the development of AI models.

The severe lack of representation and diversity in the computing and artificial intelligence industry has contributed to the problems of so-called racist or sexist AI systems in society today. It's a fact that the field of Computing, AI and Machine learning (ML) is overwhelmingly dominated by white and male experts globally. We cannot be surprised to find that the needs of minority demographics are often neglected when the people and experts who create these

systems that society runs on have little to no minority representation in their ranks. The flawed systems we have today is a natural outcome of this phenomenon. And we have ourselves to blame for not addressing it earlier. Humans have biases and that's not inherently negative. It's part of our natural evolution borne by our survival instincts and underpinned by life experiences. Hence, it should be foretelling that the AI and ML systems developed in a male white bubble would no doubt replicate and perpetuate historical gender and racial biases and power imbalances. This issue affects the data quality problems as well. When you think about confirmation bias; scientists and developers are more likely to seek out datasets that confirm their preexisting biases for use in their AI models. This could happen consciously and unconsciously. We can extrapolate from here to other dimensions of the same problem which have even greater ramifications for society. If the computing industry does not encourage more diversity to bring on people from all backgrounds who can play the "devil's advocate", these problems have no end in sight. When you think about my second and third arguments, it is obvious that the two are the perfect recipes for the creations of algorithms and AI systems that will continually perpetuate historical racial and gender biases. Perhaps policy initiatives could be implemented if the public and private sector are willing to come together to move the needle on this issue of monumental proportions for society.

In conclusion, the adverse effects of racism and sexism exhibited by robots and Artificial Intelligence (AI) systems is a reflection of the biases and prejudices of their respective developers and the quality of the data by which they were developed. Robots are not inherently racist or sexist; the quality of output is determined by the quality of input (data); the severe lack of representation and diversity in the computing and artificial intelligence industry has contributed to the problems of so-called racist or sexist AI systems today. Rather than

sensationalize the fact that robots exhibit racist and sexist tendencies, governments, industry and academia would do better to focus on the real elephants in the room as I've articulated in the preceding paragraphs.

Reference

Verma, Pranshu. "These Robots Were Trained on AI. They Became Racist and Sexist." The Washington Post, July 19, 2022.

<https://www.washingtonpost.com/technology/2022/07/16/racist-robots-ai/>.