

Fighting for Control over Algorithms and Datasets:

SOCIOTECHNICAL SOLUTIONS TO OVERCOME GENDER & RACIAL BIASES IN ARTIFICIAL INTELLIGENCE





Authors:

Bangkit Adhi Wiguna

Editor:

Achmed Faiz Yudha Siregar

Designer & Layouter:

Hanny Handayani

Introduction

Joy Buolamwini, a computer scientist and digital activist working at the MIT Media Lab, discovered a strange occurrence while interacting with some facial analysis software. While still a graduate student at MIT in 2015, Buolamwini tried out some facial analysis software and found that the artificial intelligence technology in the software had difficulty recognizing her dark-skinned face. However, when she wore a white mask, the software suddenly recognized her face easily. Buolamwini couldn't believe what she had just experienced for a moment. She took off her white mask again and the artificial intelligence in the software again did not recognize her face. She then asked her light-skinned colleague to try the same software. However, the same incident did not occur. The same software immediately recognized the face of her light-skinned colleague.¹ Buolamwini's experience made her realize that artificial intelligence technology, which has been held in high regard for its rationality and neutrality in thinking independently of humans,² actually still has the same bias as human thinking, in this case in the form of racial bias.

Apart from racial bias, artificial intelligence has also been proven to be trapped in gender bias. One popular case that shows that artificial intelligence can be trapped in gender bias is the case of Amazon's development of artificial intelligence-based recruitment tools. This tool was initially developed to simplify the recruitment process for Amazon employees.³ The artificial intelligence in this tool is designed to assess the Curriculum Vitae (CV) of prospective Amazon employees on a scale of 1 to 5 based on certain criteria. In 2014, Amazon piloted this software for the recruitment process. In practice, this tool turns out to exclude many potential female employees who apply for technical job positions, such as software developers.⁴ Amazon then carried out further investigations into their device and then officially stopped using it in 2015. They found that the software proved to give poor marks to CVs containing the word "women" submitted by prospective Amazon employees.⁵ Further use of this device will certainly be detrimental to prospective female employees who will apply to Amazon.

Table 1. Gender Shades Findings on Possible Errors of Artificial Intelligence in Detecting Faces

Face Type	Error Rate	Case/Problem
Light-skinned Male	1%	Artificial intelligence successfully identified 99% of the gender and race of light-skinned male users.
Light-skinned Female	7%	Artificial intelligence successfully identified 93% of the gender and race of light-skinned female users, leaving 7% of users misidentified by gender, race, or both at the same time.
Dark-skinned Male	12%	Artificial intelligence successfully identified 88% of the gender and race of dark-skinned male users, leaving 12% of users misidentified by gender, race, or both at once.
Dark-skinned Female	37%	Artificial intelligence only managed to identify 63% of the gender and race of dark-skinned female users, leaving 37% of users misidentified by gender, race, or both at once.

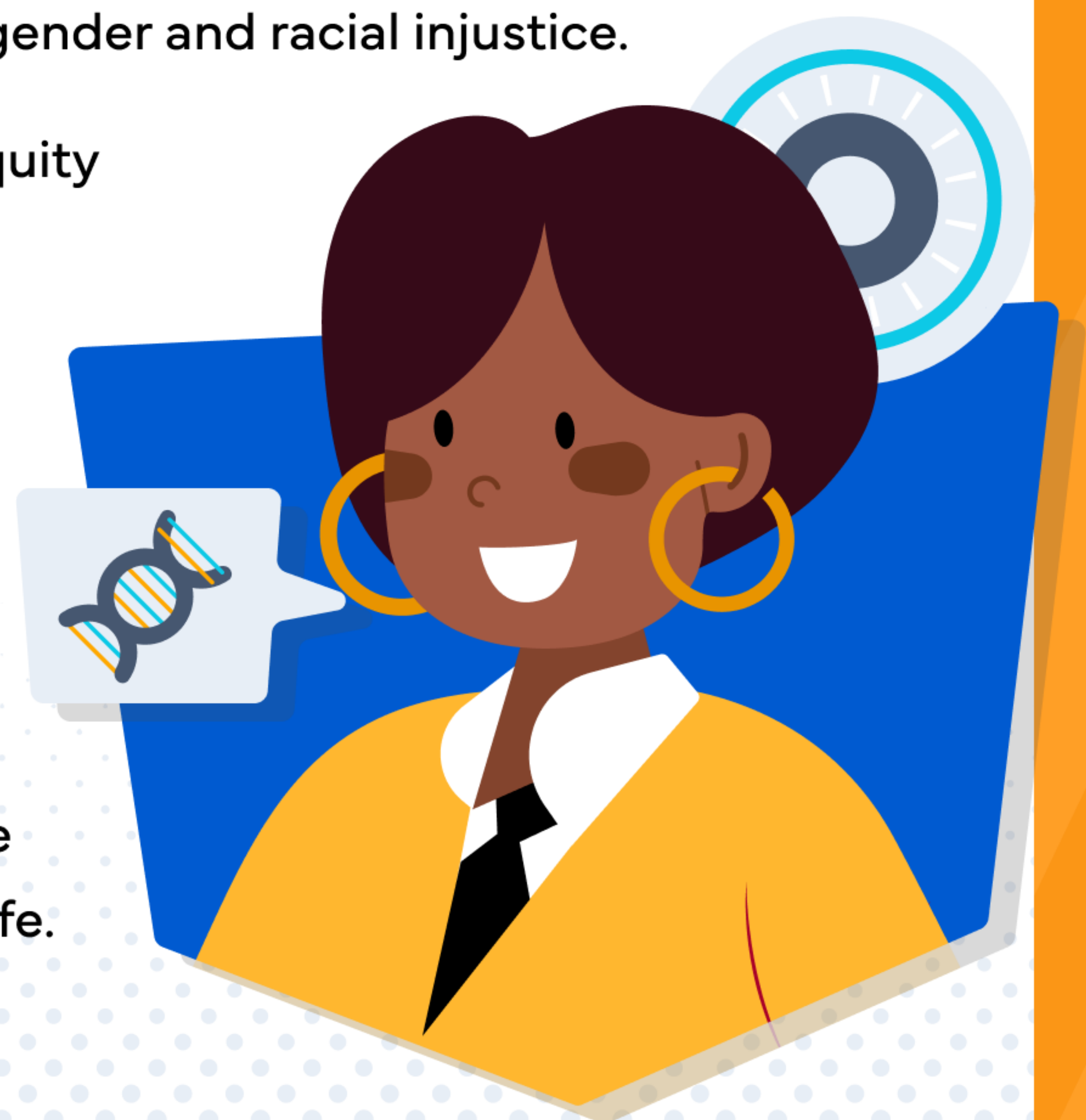
Source: Buolamwini⁶

Gender and racial bias in artificial intelligence technology is not nonsense. Gender and racial bias are also not isolated cases that only happen to one or two people or certain companies. The Berkeley Haas Center of Equity has conducted research on artificial intelligence technology used by many industries in the United States. They found that around 44% of the total artificial intelligence technology used detected gender bias and 25% of them had both gender bias and racial bias.⁷

Gender Shades, a project created by Buolamwini to combat gender and racial bias by artificial intelligence, also found bias in many widely used artificial intelligences. Gender Shades tests artificial intelligence technology produced and successfully marketed by IBM, Microsoft, and Face++ to analyze 1270 facial photos in a dataset created by Buolamwini. The findings of this research show that the digital intelligence technology being tested has a probability of error of 1% in detecting light-skinned male faces, 7% for light-skinned female faces, 12% for dark-skinned male faces, and 37% for dark-skinned female faces.⁸ These various findings show that artificial intelligence can be trapped in gender and racial bias.

The impact of gender and racial bias in artificial intelligence technology is no joke. The World Economic Forum predicts that by 2025 artificial intelligence technology could replace around 85 million jobs.⁹ By 2030, Goldman Sachs even predicts that artificial intelligence could replace around 300 million permanent jobs worldwide.¹⁰ Pew Research Center also predicts that artificial intelligence will greatly influence 77% of all types of jobs in the world.¹¹ In a smaller context, Amazon stated that artificial intelligence will become an important technology in its employee recruitment process in the next 5 years even though its artificial technology was proven to be trapped in gender and racial bias in 2015.¹² With the widespread influence of artificial intelligence on human work, gender bias and race in artificial intelligence can give rise to gender and racial injustice.

Furthermore, Berkeley Haas Center of Equity states that gender and racial inequities resulting from bias in artificial intelligence can be felt in three different contexts: (1) individuals, (2) society, and (3) markets.¹³ The impact of gender and racial bias by artificial intelligence in the individual context means that individuals with discriminated against gender or racial identities experience limited access to opportunities for a better life.





Gender and racial bias in artificial intelligence used in a company's recruitment process can make the access of people with certain gender and racial identities in the labor market more limited. In fact, a person's access or opportunity to work should be determined by the abilities he or she has and the need for those abilities, not determined by gender and racial identity. Gender and racial bias by artificial intelligence used in public service activities also presents injustice in the form of limiting access to public services for citizens with certain gender and racial identities. In fact, the state should provide public services equally to all its citizens.

At the societal level, gender and racial bias by artificial intelligence technologies can reproduce gender and race-based discrimination. The presence of technology such as artificial intelligence should be able to help humans solve problems that humans could not themselves before this technology arrived. However, gender and racial bias by artificial intelligence actually shows that this technology only reproduces, and even tends to exacerbate, the gender and race-based discrimination that has been a human problem throughout history. The use of artificial intelligence with gender and racial bias by certain individuals may be a means of transferring bias from one individual to another. At the market level, gender and racial bias in artificial intelligence technology can make a company's market reach, both in the commodity market and the labor market, increasingly narrow. Companies that use artificial intelligence with gender and racial bias become exclusive towards certain genders and races. This certainly hinders the accumulation of profits for the company so that its business development is disrupted. The fact that artificial intelligence, which is often praised for its rationality and superiority over other human-made technology, only reproduces a problem that humans cannot solve throughout their entire civilization leads to a crucial question: How can biases exist in artificial intelligence?

How Can Biases Emerge in Artificial Intelligence?

Several findings and cases previously described show that gender and racial bias are very likely to emerge in artificial intelligence. Technology, which has been assumed to be neutral and able to work more rationally than humans, can actually be trapped in the same biases experienced by humans. How can bias, especially gender and racial bias, arise in artificial intelligence?

Mendonça de Lima et al. found that gender and racial bias in artificial intelligence is caused by at least three factors.¹⁴ **First**, a biased algorithm. Algorithms are an important part of artificial intelligence, but they are not equipped with intelligence equivalent to human intelligence for their operation. As a result, the formation of the algorithm was carried out by a group of humans. In its formation, algorithms are very likely to be influenced by the human biases that shape them. **Second**, bias in the dataset that is the basis for artificial intelligence training. Several types of artificial intelligence that are currently widely used, such as machine learning and natural language processing (NLP), really require a set of datasets as a basis for learning. Through this dataset, artificial intelligence will learn to recognize patterns and work based on what it learns. If the dataset presented contains bias, artificial intelligence will also experience bias. **Third**, the structure of society where artificial intelligence is formed and used. Societal structures that perpetuate gender and racial discrimination will also shape the group mind-set that constructs artificial intelligence algorithms and forms the datasets that become the basis for their training.

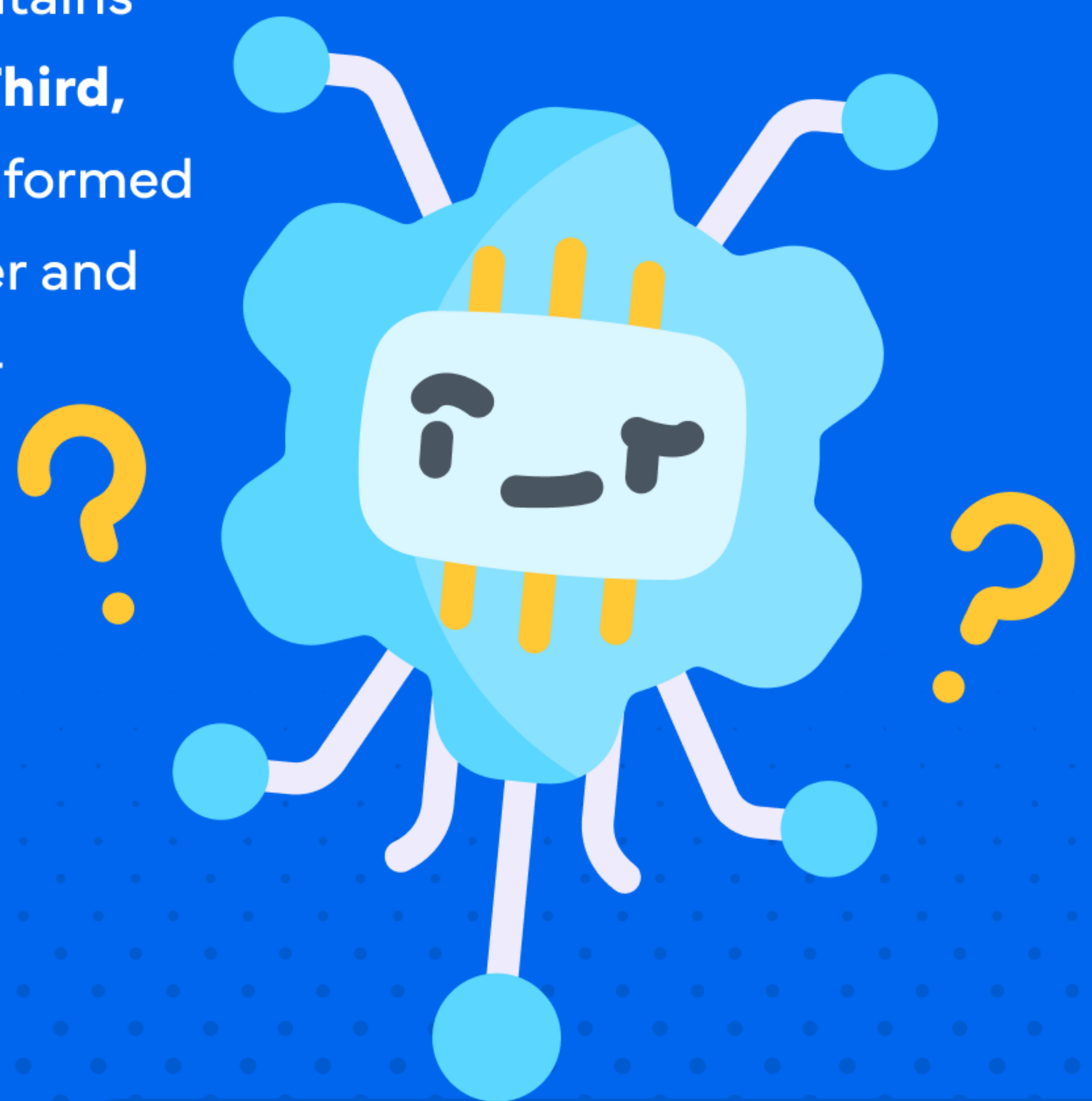


Figure 1. Artificial Intelligence Fails to Identify Oprah Winfrey's Face



Source: Buolamwini¹⁵

Several factors that cause the emergence of bias in artificial intelligence show that this technology cannot be completely separated from human intervention, which was previously full of bias. There are at least three types of artificial intelligence that might be created by humans, namely Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Superintelligence (ASI). ANI is a type of artificial intelligence that was specifically created to complete certain jobs with high capability and precision. This type still relies heavily on human intelligence, especially in its creation and use. Then, AGI is a type of artificial intelligence that is projected to reach a level of intelligence equivalent to humans. Lastly, ASI is a type of artificial intelligence that is projected to surpass human intelligence.¹⁶ Most of the artificial intelligence technologies used around the world today are dominated by ANI. Around 77% of all devices in the world use ANI technology, such as facial analysis devices, virtual assistants, generative artificial intelligence, and others. Meanwhile, AGI is currently still at a theoretical level, with no concrete examples of its use yet. Then, ASI is only a speculative model of future artificial intelligence.¹⁷

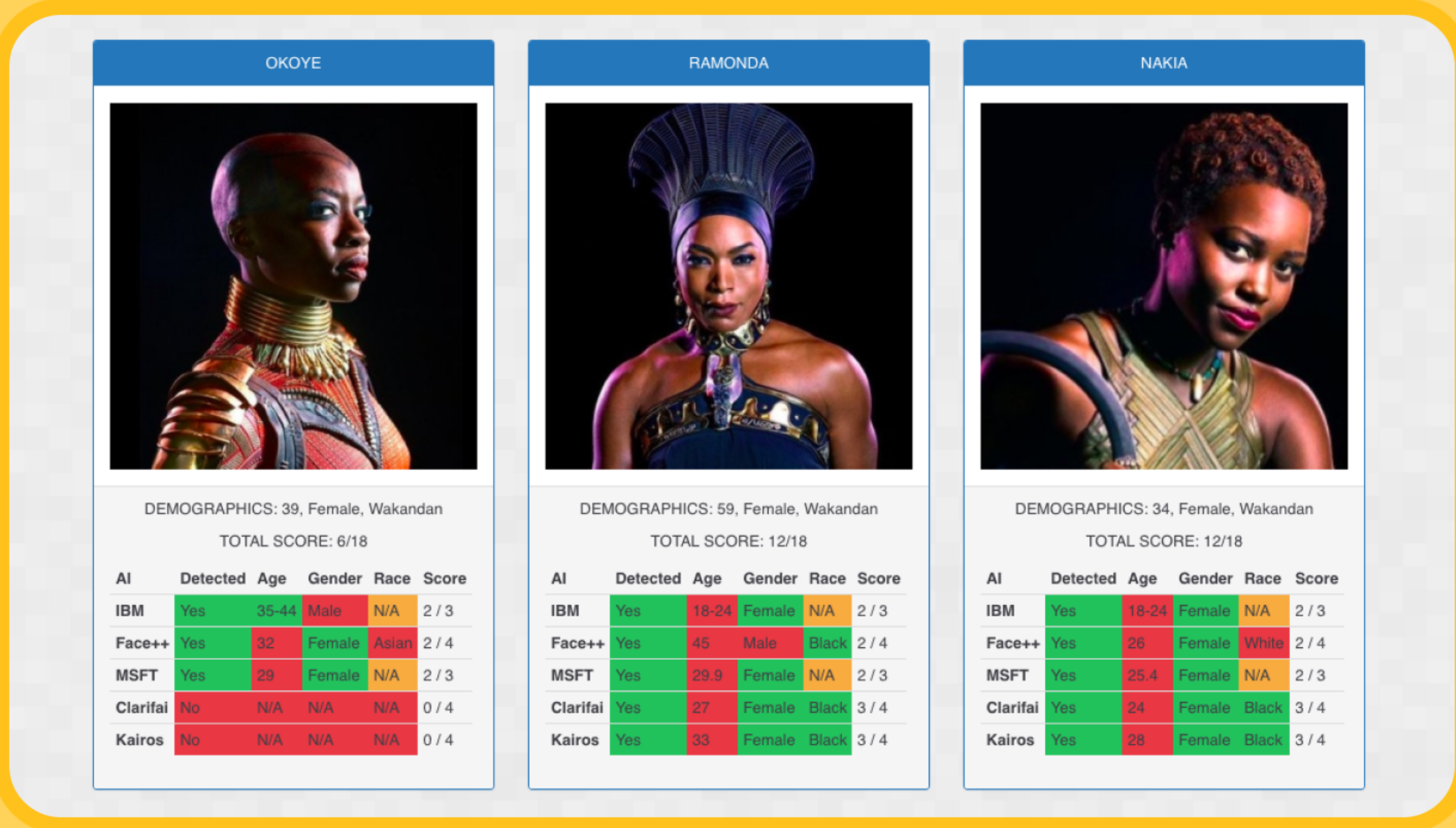


ANI is a type of artificial intelligence that still relies heavily on human intelligence, especially in its creation. Facial analysis and CV analysis technology for recruitment, for example, still relies heavily on algorithms and human-made datasets. If algorithms and datasets contain human biases that construct them, the artificial intelligence created is very likely to carry out the same biases. With the massive use of ANI, gender and racial bias by artificial intelligence is still very possible today.

With the discovery of gender and racial bias in several artificial intelligence technologies and the massive use of ANI which still relies on human intelligence, several efforts have been made to prevent gender and racial bias by artificial intelligence. Several researchers from Boston University who collaborated with Microsoft have developed a method to reduce bias in NLP by eliminating associations between words that are gender or race biased in the word embedding method. The word embedding method refers to associating one word with another word using certain mathematical computations so that artificial intelligence can learn human words more easily.¹⁸ NLP datasets and learning patterns are often biased. This bias usually manifests itself in the emergence of associations between words that are actually neutral and certain genders or races. Words that are gender neutral; such as beautiful, impulsive, emotional; can be associated with women. Meanwhile, words such as strong, responsive and agile are associated with men. Researchers try to remove these stereotyped associations so that gender and racial bias in NLP can be reduced. This method succeeded in reducing the representation of gender and racial stereotypes in NLP from 19% to 6%. However, this method will ultimately still leave bias in NLP so the focus of this problem is ultimately to eliminate gender and racial bias in society itself rather than in NLP datasets and word embeddings.¹⁹

Apart from text-based artificial intelligence such as NLP, gender and racial bias also surface in image-based artificial intelligence that uses machine learning. Face detection technology in smartphones and autonomous vehicles often does not recognize human faces of certain genders or races. Researchers from the University of Virginia, University of California Los Angeles, and the Allen Institute for Artificial Intelligence tried to solve this problem in 2019. They analyzed datasets from Microsoft Common Objects in Context (COCO) and imSitu containing images of humans in various contexts. They then compared ground truth labels, data entered by human workers into artificial intelligence technology, and prediction models, predictions about images made by artificial intelligence technology. This research found that the presence of gender/racial-related variables, such as clothing or certain roles, in images creates bias. Therefore, eliminating gender related variables in images could be a solution. This method is proven to reduce bias in image-based artificial intelligence by 53% for COCO-based and 67% for imSitu-based. However, this method still leaves bias because the data in the dataset will be continuously updated and the reproduction of gender and racial bias in society will be reproduced in the dataset and artificial intelligence technology.²⁰

Figure 2. Artificial intelligence belonging to several technology companies has difficulty identifying the face of a dark-skinned woman in the Black Panther movie.



Source: Gender Shades²¹

The technical solutions discussed previously have been implemented to prevent gender and racial bias in artificial intelligence. These solutions are proven to reduce gender and racial bias in artificial intelligence. However, these technical solutions have also proven to have limitations. Artificial intelligence that relies on human-made algorithms and datasets will continue to produce bias as long as the production spaces of the algorithms and datasets are not intervened. In other words, instead of intervening in artificial intelligence technology itself, the more appropriate thing to do is to carry out structural intervention by upholding gender and racial justice in the production spaces of algorithms and datasets that regulate the actions of artificial intelligence. This does not mean that the technical solutions that have been implemented can be put aside. Efforts to overcome gender and racial bias in artificial intelligence require comprehensive efforts, from the technical to the structural side. At this point, a sociotechnical solution is needed to address gender and racial bias in artificial intelligence.

The Sociotechnical Solution for AI Gender and Racial Biases

Sociotechnical solutions refer to the solutions proposed by Varsha P.S. to address the issue of gender and racial bias in artificial intelligence technology. Varsha stated that the problem of gender and racial bias in artificial intelligence is caused by gender and racial discrimination and vulnerabilities that occur in three important contexts in the creation of artificial intelligence: (1) domain experts, (2) end users, and (3) society.²² Domain experts refers to the context of producing algorithms and datasets that later determine the design and behavior of artificial intelligence technology. The discrimination and vulnerability that arises in this context is the minimal involvement of individuals who are experts in the construction of algorithms and datasets for certain gender and racial identities. As a result, the process of forming algorithms and datasets becomes biased so that the resulting artificial intelligence is biased. The lack of involvement of gender and racial minority groups in the formation of algorithms and datasets, for example, makes it difficult for the resulting artificial intelligence to identify such groups.²³

Then, end users refer to the context of using artificial intelligence technology which has been biased due to gender and racial discrimination and vulnerability in the context of domain experts. The use of biased artificial intelligence will cause users to be influenced by biases that have been internalized in artificial intelligence in the context of domain experts. The end user context is important because algorithm development and dataset updates to improve artificial intelligence capabilities will rely on input during the use process. Users who have been influenced by biased artificial intelligence are highly likely to provide biased feedback. Research by Sunny Shrestha and Sanchari Das found that input from users cannot be directly processed back into evaluation material for developing algorithms and datasets for artificial intelligence.²⁴ Because, this data is very likely to contain bias from users. Thus, discrimination and vulnerability in the end user context can reproduce gender and racial biases that have previously been internalized in the end user context.

Finally, society refers to the societal context that is the background for the formation of artificial intelligence. Societal structures that discriminate against certain genders and races have the potential to reproduce and strengthen gender and racial biases in artificial intelligence. Without structural efforts to combat discrimination and uphold gender and racial justice; like improving socioeconomic and status and opportunity,²⁵ promoting reproductive justice,²⁶ and reforming the criminal justice system in favor of people with marginalized gender and race;²⁷ discriminative bias in artificial intelligence technologies will continue to occur. The algorithms and datasets

that determine the form and behavior of artificial intelligence are nevertheless a reflection of society. With current artificial intelligence capabilities still limited to the ANI level, instead of avoiding or correcting human biases, artificial intelligence will actually reproduce them. Therefore, in this context, the struggle against gender and racial discrimination in artificial intelligence is actually the same as the struggle for gender and race in general.



Several efforts can be made to intervene in these three contexts which are sources of bias in artificial intelligence. Claude Draude, Goda Klumbyte, and Phillip Lücking proposed the concept of situated algorithms to solve problems in the context of domain experts.²³ This concept emphasizes that technology and the science that forms it cannot be separated from the humans who produce them. Therefore, to produce fair technology and science, the production process of both must be fair first, involving not only people with dominant gender and racial identities, but also minorities. In the context of the formation of artificial intelligence, the involvement of people with minority identities becomes very important. This involvement can be pursued from the level of discussion or decision making regarding algorithm and dataset design to the technical level by involving programmers with minority identities. Apart from algorithms, datasets that are used as learning material for artificial intelligence also need to be continuously modified so that they are not biased. Dataset creators must be able to ensure that the data that will later become learning material for artificial intelligence is not biased towards certain genders or races, but must try to include data that is as varied as possible and avoid associations between data that reflect bias and discrimination against certain identities.

Then, intervention against bias in the end user context can be carried out by strengthening collaboration between artificial intelligence producing companies and their users.²⁸ Apart from focusing on developing its artificial intelligence capabilities, companies must also be able to educate their users about the potential gender and racial biases that may be introduced by it. The state can also play a role in this context. When companies operate solely based on the logic of profit accumulation and are reluctant to fix problems of gender and racial bias in their technology, the state must take that role. Intervention by this state must of course pay attention to certain limitations so as not to limit the development of artificial intelligence technology itself. Apart from relying on the role of companies and the state, efforts to overcome the problem of gender and racial bias in the context of end users can also be carried out by artificial intelligence users themselves. Galit Wellner states that artificial intelligence users can form a kind of "ombudsman" institution to monitor biases that may appear in certain artificial intelligence technologies.²⁹ This institution can also become a forum for political power for users of artificial intelligence to pressure companies to take action against biases that appear in their technology.

Table 2. Three Important Contexts in Artificial Intelligence Production that Need Intervention to Combat Gender and Racial Bias

Context	Explanation	Possible Solution
Domain Experts	Domain experts refers to the production space of artificial intelligence technology which includes creating algorithms, datasets, and artificial intelligence training. Biases can arise in this context because artificial intelligence production spaces underrepresented people of certain gender and racial identities.	Situated algorithms and datasets: Involving more societal representation of various identities in the creation of algorithms and datasets that govern the form and behavior of artificial intelligence.
End Users	End users refer to the activities of using artificial intelligence by society, both commercially and through certain public services. Bias can arise because people as users are influenced by artificial intelligence which already contains bias due to its biased production space. Input from users to companies producing artificial intelligence can also be biased.	Forming cooperation between companies and artificial intelligence users. Users can also take the initiative to form a kind of "ombudsman" institution to control companies and their artificial intelligence technology so that they do not discriminate against certain genders and races.
Society	Society refers to the structural context of society which is the background for the production of artificial intelligence. If the structure of society is patriarchal and discriminatory against certain races, the production of artificial intelligence will very likely be affected by certain gender and racial biases.	Fighting gender and racial bias in AI is part of combating societal discrimination. Companies must ensure fair technology, the state must end discrimination, and society can work with academics to ensure algorithms and datasets promote justice.

Source: Analyzed and elaborated from the sociotechnical solution for gender and racial biases in AI.³⁰

Finally, intervention against bias in a societal context requires much more effort. Intervention in this context requires cooperation from companies, countries, as well as artificial intelligence users themselves. Companies must commit to addressing gender and racial bias in their technology. The state and society must be able to pressure and encourage companies to uphold these commitments. On the other hand, the state must also uphold gender and racial justice evenly so that discrimination does not become a common culture among society. The dominant role of companies and the state does not necessarily mean that society becomes passive actors. Society can play an active role by forming various initiatives to combat gender and racial bias in artificial intelligence.³¹ Academics can also play a role in combating gender and racial bias in this societal context. By intensively producing interdisciplinary research on gender and racial bias in artificial intelligence, academics can become important actors who educate the public while offering alternative solutions to overcome this problem. Academics can also collaborate with various community initiatives to gain control over algorithms and datasets so that gender and racial justice can be upheld.

Conclusion

Artificial intelligence technology has long been celebrated for its advantages in thinking creatively and independently of human thought. Artificial intelligence is even projected to be able to match or compete with human intelligence in the future through the formation of AGI and ASI. However, the current development of artificial intelligence technology has still not achieved the praise and hopes directed at it. This technology is proven to still be very dependent on human intelligence with the proliferation of ANIs that rely on human-made algorithms and datasets.

Because it still relies heavily on human intelligence, artificial intelligence is very likely to reproduce biases and errors made by humans. Gender and racial discrimination have been problems throughout human history.

The presence of artificial intelligence technologies, instead of solving these problems, reproduces them due to the current limitations of these technologies. The emergence of various cases of gender and racial bias by artificial intelligence technology belonging to various companies proves the current limitations of this technology. However, artificial intelligence cannot be blamed entirely for its gender and racial biases. It is a joint effort of all members of society.

There are at least three contexts in the production of artificial intelligence technology that must be intervened in to overcome gender and racial bias, namely domain experts, end users, and society. Intervention needs to be carried out in the production space, use and structural context of artificial intelligence society. These three layered and interrelated contexts demonstrate that technical solutions to the problem of gender and racial bias in artificial intelligence are insufficient. A sociotechnical solution that includes both technical and social aspects is needed to resolve gender and racial bias in artificial intelligence. Artificial intelligence technology manufacturers need to be committed to upholding gender justice, supported by various levels of society who are involved in the planning and evaluation process for making artificial intelligence. The government supervises this process through a firm bureaucratic framework that still provides freedom for innovation in the artificial intelligence industry. All of this needs to be done because the efforts to fight against gender and racial bias in artificial intelligence are actually the same as fighting against gender and racial discrimination that has occurred throughout human history. Artificial intelligence simply reproduces these problems because of its limitations.

References

- ¹ Buolamwini, J. (2019). "Artificial intelligence has a problem with gender and racial bias. Here's how to solve it". Media MIT. [online]. Available at: <https://www.media.mit.edu/articles/artificial-intelligence-has-a-problem-with-gender-and-racial-bias-here-s-how-to-solve-it/>
- ² Larson, E. (2022). The Myth of Artificial Intelligence. Harvard University Press.
- ³ Buolamwini, R. (2017). Gender Shades: Intersectional Phenotypic and Demographic Evaluation of Face Datasets and Gender Classifiers, published thesis. Massachusetts Institute of Technology.
- ⁴ Parsani, P. (2024). "Case Study: How Amazon's AI Recruiting Tool 'Learnt' Gender Bias". Cut the Saas. [online]. Available at: <https://www.cut-the-saas.com/ai/-case-study-how-amazons-ai-recruiting-tool-learnt-gender-bias>
- ⁵ BBC. (2018). "Amazon scrapped 'sexist AI' tool". BBC. [online]. Available at: <https://www.bbc.com/news/technology-45809919>
- ⁶ UN Women. (2024). "Artificial Intelligence and gender equality". UN Women. [online]. Available at: <https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality>
- ⁷ Buolamwini. (2017). Gender Shades: Intersectional Phenotypic and Demographic Evaluation of Face Datasets and Gender Classifiers. Massachusetts Institute of Technology.
- ⁸ OpenCertHub. (2024). "World Economic Forum: AI will take 85 million jobs in 5 years!". OpenCertHub. [online]. Available at: <https://www.opencertthub.com/en/wef-ai-replace-human-jobs/#::~text=WEF%20predicted%20that%20the%20rise,addition%20of%2012%20million%20jobs>.
- ⁹ Talmage-Rostron, Mark. (2024). "How Will Artificial Intelligence Affect Jobs 2024-2030". [online]. Available at: <https://www.nexford.edu/insights/how-will-ai-affect-jobs>
- ¹⁰ Howarth, J. (2024). "60+ Stats On AI Replacing Jobs (2024)". [online]. Available at: <https://explodingtopics.com/blog/ai-replacing-jobs>

- ¹¹ Dastin, J. (2018). "Insight - Amazon scraps secret AI recruiting tool that showed bias against women". Reuters. [online]. Available at: <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/>
- ¹² Smith, G., Rustagi, I. (2020). *Mitigating Bias in Artificial Intelligence: An Equity Fluent Leadership Playbook*. Berkeley Haas Center for Equity, Gender, and Leadership.
- ¹³ Mendonca, R., Pisker, B., Correa, V. (2023). "Gender Bias in Artificial Intelligence: A Systematic Review of the Literature". *TelSoc: Telecommunication & the Digital Economy* 11(2), pp. 8—30.
- ¹⁴ Buolamwini, J. (2018). "When AI Fails on Oprah, Serena Williams, and Michelle Obama, It's Time to Face Truth". Medium. [online]. Available at: <https://medium.com/@Joy.Buolamwini/when-ai-fails-on-oprah-serena-williams-and-michelle-obama-its-time-to-face-truth-bf7c2c8a4119>
- ¹⁵ George, S., George, H. (2023). *Beyond Human Intelligence: Exploring the Advancements and Implications of ANI, AGI, and ASI*. Book Rivers.
- ¹⁶ National University. (2024). "131 AI Statistics and Trends for 2024". National University Official Website. [online]. Available at: <https://www.nu.edu/blog/ai-statistics-trends/>
- ¹⁷ Bolukbasi, T., Chang, K., Zou, J., Saligrama, V., Kalai, A. (2016). "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings". 30th Conference on Neural Information Processing Systems.
- ¹⁸ Connor, S., Liu, H. (2024). "Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities". *AI & Society* 39, pp. 2045—2057.
- ¹⁹ Connor, Liu. (2024). "Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities". *AI & Society*, pp. 2051.
- ²⁰ Gender Shades. (2018). "Black Panther Face Scorecard". Gender Shades. [online]. Available at: <http://gendershades.org/bpscorecard.html>
- ²¹ Varsha. (2023). "How can we manage biases in artificial intelligence systems – A systematic literature review". *International Journal of Information Management Data Insight* 3(1), pp. 1—9.
- ²² Ananya. (2024). "AI image generators often give racist and sexist results: can they be fixed?". *Nature*. [online]. Available at: <https://www.nature.com/articles/d41586-024-00674-9>

- ²³ Shresta, S., Das, S. (2022). "Exploring gender biases in ML and AI academic research through systematic literature review". *Sec. Machine Learning and Artificial Intelligence* 5, pp. 1—17.
- ²⁴ Draude, C., Klumbyte, G., Lucking, P., Treusch, P. (2019). "Situated algorithms: a sociotechnical systemic approach to bias". *Online Information Review* 44(2), pp. 325—342.
- ²⁵ Clark, E., Cranston, E., Polin, T., Ndumbe-Eyoh, S., MacDonald, D., Betker, C., Dobbins, M. (2022) "Structural interventions that affect racial inequities and their impact on population health outcomes: a systematic review". *BMC Public Health* 22(2162).
- ²⁶ Otieno, A. (2022) "Eliminating Racial Discrimination: The Challenges of Prevention and Enforcement of Prohibition". *United Nations Chronicle*. [online]. Available at: <https://www.un.org/en/chronicle/article/eliminating-racial-discrimination-challenges-prevention-and-enforcement-prohibition>
- ²⁷ Shelton, R., Adsul, P., Oh, A. (2021). "Recommendations for Addressing Structural Racism in Implementation Science: A Call to the Field". *Ethnicity & Disease* 31(1), pp. 357—364.
- ²⁸ Bui, C. (2021). *Exploring Bias Against Women in Artificial Intelligence Practitioners' Views on Systems of Discrimination*, published thesis. Department of Informatics The Faculty of Mathematics and Natural Sciences University of Oslo.
- ²⁹ Wellner, G. (2022). "Some Policy Recommendations to Fight Gender and Racial Biases in AI". *International Review of Information Ethics* 32(1).
- ³⁰ Varsha. (2023). "How can we manage biases in artificial intelligence systems – A systematic literature review", pp. 1—9.
- ³¹ Hall, P., Ellis, D. (2023). "A systematic review of socio-technical gender bias in AI algorithms". *Online Information Review* 47(7), pp. 1264—1279.





Center for Digital Society

Faculty of Social and Political Sciences
Universitas Gadjah Mada
Room BC 201-202, BC Building 2nd Floor,
Jalan Socio Yustisia 1
Bulaksumur, Yogyakarta, 55281, Indonesia

Phone : (0274) 563362, Ext. 116
Email : cfds.fisipol@ugm.ac.id
Website : digitalsociety.id



facebook.com/cfds.ugm



[@cfds_ugm](https://line.me/@cfds_ugm)



[Center for Social Society \(CfDS\)](https://www.linkedin.com/company/Center%20for%20Social%20Society%20(CfDS)/)



[@cfds_ugm](https://twitter.com/cfds_ugm)



[Cfds_ugm](https://www.instagram.com/Cfds_ugm)



[CfDS UGM](https://www.youtube.com/CfDS%20UGM)