Committee: Human Rights Council (HRC)

Topic: Addressing algorithmic bias and discrimination in AI systems

Student Officer: Arina Ostapenko

Position: Deputy President

Dear Delegates,

My name is Arina Ostapenko, I am currently an IB1 student at St. Catherine's British School, and I am honoured and thrilled to be serving as one of the Deputy Presidents in the Human Rights Council in this year's CSMUN. I have attended MUN conferences for the last two years and have found the experience both amazing and enlightening, and I hope you will too. I got to socialise and meet people, create new connections and bonds with different people, but most importantly, I got to debate upon topics that currently affect our planet, and the societies we all thrive in.

In this year's conference, I will be in charge of the topic "Addressing algorithmic bias and discrimination in AI systems". This study guide is created to help you in your research and provide a thorough understanding of the topic for your preparation. I strongly encourage each delegate to take this help and use it with purpose, to enhance your comprehension of this issue, so you can confidently contribute as much as possible during the conference.

However, this study guide is for introductory purposes only, and it is crucial for every delegate to do their own research. If you have any further questions regarding the content covered in the study guide, do not hesitate to contact me via email at arinaa.ostapenkoo@gmail.com. Furthermore, if you have any other questions, please contact Aloi Gkania, the President of the Human Rights Council, at aloigkaniaa@gmail.com.

Best regards,

Arina

Topic Introduction

Algorithms in AI systems calculate patterns within data and predict output values based on input values already provided by humans. Therefore, biased algorithms have the potential to affect both the patterns in the original data, as well as the predicted output values.[1] This biased data can produce inequality and discrimination within these AI systems.

Reducing the trust in AI systems themselves or organisations that use them can be problematic. On the one hand, it can create several legal issues, including the possibility of a company being sued by a person or group due to discrimination that arises through biased algorithms.[2] On the other hand, possible economic issues could arise, as a loss of trust from specific groups that are affected by these biases towards organisations that use the biased AI systems, could cause these organisations to lose business.[3] It is further concerning because AI systems in areas such as healthcare, algorithmic bias could cause misdiagnoses in specific patient groups, for example, gender or ethnic minorities, as they are more likely to be targeted by algorithmic bias[4].

Definition of key concepts

Algorithm

An algorithm is "The set of rules a machine follows to achieve a particular goal".[5] Algorithms can exhibit algorithmic bias, which is "when systematic errors in machine learning algorithms produce unfair or discriminatory outcomes".[6]

---

[1] IBM Consulting. "Artificial Intelligence for Human Resources | IBM." *Ibm.com*, 17 July 2024, www.ibm.com/think/topics/ai-in-hr.

[2] IBM Consulting. "Artificial Intelligence for Human Resources | IBM." *Ibm.com*, 17 July 2024, www.ibm.com/think/topics/ai-in-hr.

[3] IBM Consulting. "Artificial Intelligence for Human Resources | IBM." *Ibm.com*, 17 July 2024, www.ibm.com/think/topics/ai-in-hr.

[4] Norori, Natalia, et al. "Addressing Bias in Big Data and AI for Health Care: A Call for Open Science." Patterns, vol. 2, no. 10, Oct. 2021, p. 100347, https://doi.org/10.1016/j.patter.2021.100347.

[5] Merriam-Webster. "Definition of ALGORITHM." *Merriam-Webster.com*, 2019, www.merriam-webster.com/dictionary/algorithm.

[6] "What Is Algorithmic Bias?" *IBM*, IBM, 20 Sept. 2024, www.ibm.com/think/topics/algorithmic-bias.

AI systems

"'AI systems are machine-based systems that are designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment".[7] In other words, they are the system within which an algorithm works.

Biased interpretation

"Biased interpretation happens when an AI's output is interpreted wrongly or is over-interpreted."[8] This is a consequence of algorithmic bias.

Proxy

"The authority given to a person to act as somebody else."[9] In the context of this topic, proxies are often used in algorithms.

Recidivism

"A person's relapse into criminal behaviour, often after the person receives sanctions or undergoes intervention for a previous crime."[10]

Filter bubble

"A situation in which a person only sees or hears news and information that supports what they already believe and like. This can be created on the internet as a result of algorithms."[11] Filter bubbles increase algorithmic bias, as they reinforce the bias that already exists within our AI systems.

---

[7]"Article 3: Definitions | EU Artificial Intelligence Act." *EU Artificial Intelligence Act*, artificialintelligenceact.eu/article/3/.
[8]"Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.
[9] "Proxy." *@CambridgeWords*, 12 Mar. 2025, dictionary.cambridge.org/dictionary/english/proxy.
[10]National Institute of Justice. "Recidivism." *National Institute of Justice*, 2023, nij.ojp.gov/topics/corrections/recidivism.
[11]Cambridge Dictionary. "FILTER BUBBLE | Meaning in the Cambridge English Dictionary." *Cambridge.org*, 11 Dec. 2019, dictionary.cambridge.org/dictionary/english/filter-bubble.

Feedback loops

"A feedback loop is a part of a system in which some portion of the system's output is used as input for future operations."[12]

Re-weighting data

"Reweighting is the process of adjusting the weights assigned to various elements in a dataset or model. "[13] This could be used as a solution for algorithmic bias.

Background Information

Possible Causes of Algorithmic Bias

Firstly, bias can occur in AI when it relies on proxies that stand in for protected characteristics like race or gender.[14] Even without directly using this sensitive information, the proxies can still cause discriminatory outcomes.[15] Proxies may be biased as they might have an accidental correlation with the sensitive groups that they were meant to replace. For instance, if an algorithm were to use postal codes as a proxy that represents economic status, the algorithm might disadvantage certain groups where postal codes are associated with a specific racial group.[16]

Secondly, algorithmic bias can be further magnified and increased through feedback loops. Due to the fact that algorithms learn and evolve from themselves in order to develop, existing biases in the input or output data of an algorithm can be magnified.[17]

---

[12] Fitzgibbons, Laura. "What Is Feedback Loop? - Definition from WhatIs.com." *SearchITChannel*, Apr. 2019, www.techtarget.com/searchitchannel/definition/feedback-loop.

[13] "Reweighting Meaning | Goong.com - New Generation Dictionary." *Goong.com*, 2024, goong.com/word/reweighting-meaning/.

[14] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

[15] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

[16] Jonker, Alexandra, and Julie Rogers. "What Is Algorithmic Bias?" *IBM*, IBM, 20 Sept. 2024, www.ibm.com/think/topics/algorithmic-bias.

[17] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

However, the main cause of algorithmic bias is existing bias in training data.[18] This happens because algorithms reflect the biases of their creators. For instance, if the creator of the algorithm carries their personal bias when creating the algorithm, their bias will transfer into the algorithm.[19] This prejudice can also be caused by biased interpretations of the data by AI.[20] If the training data used to create an algorithm isn't representative of all groups with the intent to use it, or the scenarios in which the AI will have to function, the algorithm created may be prejudiced.[21]

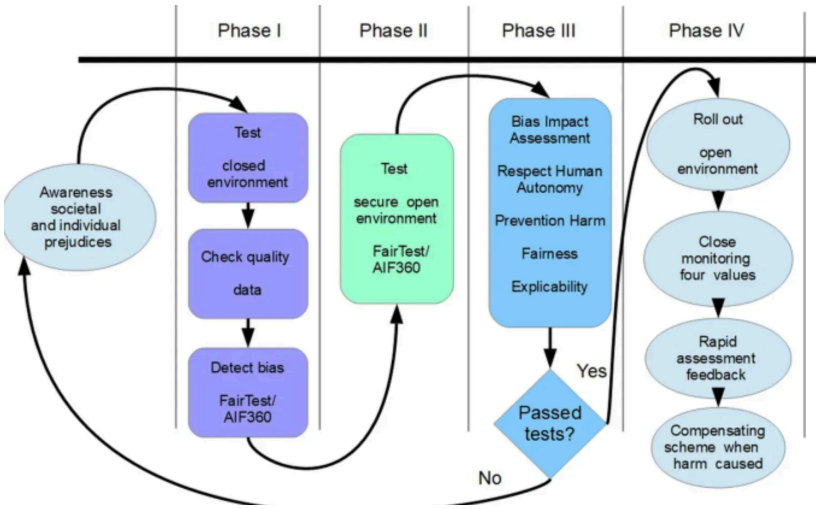

Figure 1: The phases through which an algorithm goes through to detect whether it is biased or not.[22]

Case studies

Algorithmic Bias in US Healthcare Exacerbating Racism

Racism in US healthcare algorithms can be found in multiple aspects. Due to the fact that some marginalised groups spend less money on healthcare because of the limited access to healthcare

[18] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.
[19] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.
[20] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.
[21] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.
[22] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

services, these groups might be classified by the algorithm as having a lower need for healthcare, which would exclude them from receiving necessary additional healthcare services.[23]

For instance, a widely used commercial algorithm used by US hospitals was created to predict healthcare costs, and resulted in underestimating the need of black patients for medical care.[24] The issue arose when Dr. Ziad Obermeyer of UC Berkeley reported that the algorithm often recommended healthier, white patients for healthcare risk management programmes compared to sicker, black patients.[25]

Racism within the Correctional Offender Management Profile for Alternative Sanctions (COMPAS)

Racism within COMPAS was seen through research by a non-profit investigative journalism newsroom named 'ProPublica'.[26] They led a study which analysed the risk scores assigned to 7000 people arrested in Broward county.[27] These risk scores are assigned by an algorithm which was created to assess the likelihood of defendants re-offending. Risk scores from 1-4 were low risk, 5-7 were medium risk, and 8-10 were high risk.[28] The algorithm wrongly predicted that more black defendants were likely to re-offend.[29]

[23] Lynch, Erin. "Tackling Racially Biased Health Care Algorithms - Petrie-Flom Center." *Petrie-Flom Center - the Blog of the Petrie-Flom Center at Harvard Law School*, 5 Oct. 2023, petrieflom.law.harvard.edu/2023/10/05/tackling-racially-biased-health-care-algorithms/.

[24] Gugu Ntsele. "Real-World Examples of Healthcare AI Bias." *Paubox.com*, Paubox, 11 May 2025, www.paubox.com/blog/real-world-examples-of-healthcare-ai-bias.

[25] Gugu Ntsele. "Real-World Examples of Healthcare AI Bias." *Paubox.com*, Paubox, 11 May 2025, www.paubox.com/blog/real-world-examples-of-healthcare-ai-bias.

[26] "Risk Assessments Biased against African Americans, Study Finds." *Equal Justice Initiative*, 2 June 2016, eji.org/news/risk-assessments-biased-against-african-americans/.

[27] "Risk Assessments Biased against African Americans, Study Finds." *Equal Justice Initiative*, 2 June 2016, eji.org/news/risk-assessments-biased-against-african-americans/.

[28] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.

[29] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.

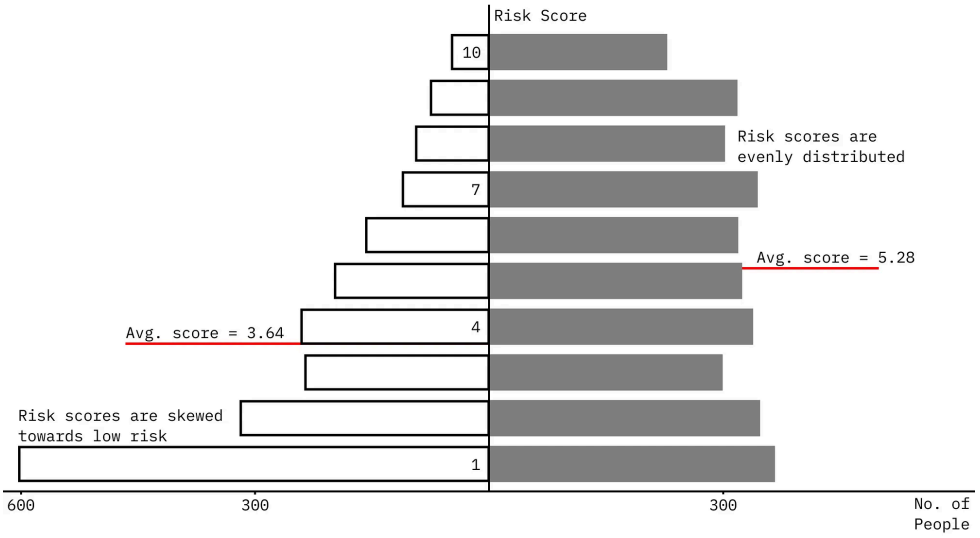

Figure 2: Black defendants have a higher average risk score, with a substantially larger number of high-risk predictions from the algorithm.[30]

ProPublica found that COMPAS was only correct in its predictions 61% of the time, and that the accuracy rate was similar between both groups.[31] This shows how a biased algorithm can happen as a result of previously biased data.[32]

[30] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.
[31] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.
[32] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.

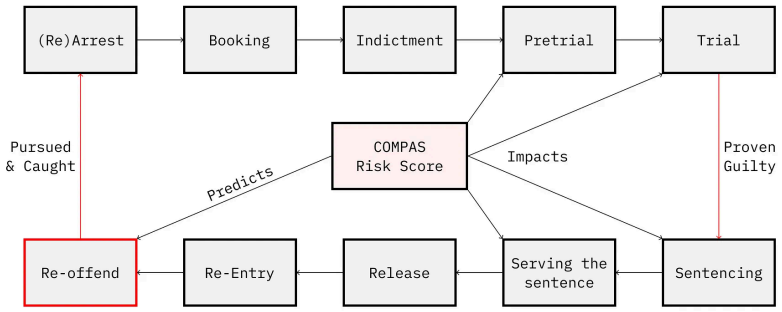

Figure 3: A simplified overview of steps in the criminal justice process and how COMPAS scores relate to them.[33]

Sexism in Amazon's Hiring Algorithm

In 2015, Amazon discovered that its hiring algorithm was heavily biased against women because it was used to accept more male applicants than female ones.[34] This is crucial as it portrays how a biased algorithm could be created, but also the impact that this can have. When considering such large corporations, it is important to understand that algorithmic bias can affect significant numbers of people.

Consequent Problems Associated with Algorithmic Bias

In our society, we receive a large number of recommendations based on the data we consume. In fact, most information we receive is curated specifically for us individually, using algorithms.[35] Algorithmic bias within curated data can create filter bubbles, which can further reinforce our personal biases, which can contribute to societal subdivisions being strengthened.[36]

[33] Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.

[34] Datatron. "Real-Life Examples of Discriminating Artificial Intelligence." *Datatron*, 25 Feb. 2021, datatron.com/real-life-examples-of-discriminating-artificial-intelligence/. 4rae

[35] "Why and How Media Curation by Algorithm Contributes | Eticas." *Eticas*, 20 Dec. 2023, eticas.ai/why-and-how-media-curation-by-algorithm-contributes/.

[36] "Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

Major countries, Organisations and Alliances

European Union (EU)

The EU considers algorithmic bias a very serious issue, which can be seen through the several legal frameworks that prevent discrimination within the EU law and the treaty of the European Union.[37] In fact, Article 2 outlines that the core values of the EU are respect for human dignity, freedom, democracy, equality, rule of law, and respect for human rights, including those of minorities.[38] Additionally, in Article 10 of the Treaty on the Functioning of the European Union, it is highlighted that "the Union shall aim to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation."[39] Biased algorithms completely trample this as they often cause minority groups to be further underrepresented. Thus, combating algorithmic bias would be a key step to ensuring that this core value is upheld. Because of this, in 2023, the EU released the AI act.[40] This act assigned applications of AI into three risk categories and banned systems that create an unacceptable risk.[41] High-risk applications are usually monitored, whereas systems that aren't considered high risk are usually left unmonitored.[42]

Further reading: https://ceur-ws.org/Vol-3442/paper-41.pdf

African Union

Algorithmic bias is a large issue in countries in Africa as it can greatly impact the trust of the public in AI systems. Reducing algorithmic bias will aid in the attainment of the African Union Agenda 2063,

[37] "EU Values Explained in One Minute." *Topics | European Parliament*, 29 Mar. 2021, www.europarl.europa.eu/topics/en/article/20210325STO00802/eu-values-explained-in-one-minute.
[38] "EU Values Explained in One Minute." *Topics | European Parliament*, 29 Mar. 2021, www.europarl.europa.eu/topics/en/article/20210325STO00802/eu-values-explained-in-one-minute.
[39] Aytekin, Ahmet. *Algorithmic Bias in the Context of European Union Anti-Discrimination Directives*. ceur-ws.org/Vol-3442/paper-41.pdf.
[40] European Union. "The EU Artificial Intelligence Act." *The Artificial Intelligence Act*, 2024, artificialintelligenceact.eu/.
[41] European Union. "The EU Artificial Intelligence Act." *The Artificial Intelligence Act*, 2024, artificialintelligenceact.eu/.
[42] European Union. "The EU Artificial Intelligence Act." *The Artificial Intelligence Act*, 2024, artificialintelligenceact.eu/.

because it would ensure that AI systems are fair and inclusive.[43] Tackling algorithmic bias could also improve the delivery of public services and advance areas like agriculture, education and health.

Historically, in Africa, there was a mistrust of foreign technologies due to past exploitation of the African community by creators of such technologies.[44] For instance, workers who developed and trained AI systems received wages that did not value their vital contribution to the creation of AI.[45] This consequently altered society's perceptions of AI in Africa. However, with enough community engagement and transparency, the trust in AI systems will likely grow, and eroding algorithmic bias could largely help in instilling this trust. Therefore, the African Union is aiming to reduce or completely remove algorithmic bias in order to try to gain higher levels of trust from the public.[46]

Further research on the Continental Artificial Intelligence Strategy within the African Union: https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf

Previous attempts to solve the issue

IBM's AI Fairness 360 Toolkit

The AI Fairness 360 Toolkit is a collection of precompiled, reusable files, functions, scripts, routines, and many other resources that computer scientists can access and reference[47], which in this case contains metrics that can be used to check for biases in datasets and machine learning models.[48]

This was an effective attempt at reducing bias within algorithms because it is designed in a way that can be used in varied fields such as finance, human capital management, education, and

[43] African Union. *CONTINENTAL ARTIFICIAL INTELLIGENCE STRATEGY Harnessing AI for Africa's Development and Prosperity*. July 2024, au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf.

[44] Pasipamire, Notice, and Abton Muroyiwa. "Navigating Algorithm Bias in AI: Ensuring Fairness and Trust in Africa." *Frontiers in Research Metrics and Analytics*, vol. 9, Frontiers Media, Oct. 2024, https://doi.org/10.3389/frma.2024.1486600.

[45] Writer, Guest. "African Digital Colonialism Is the New Face of Worker Exploitation - ICTworks." *ICTworks*, 17 Apr. 2025, www.ictworks.org/african-digital-colonialism/.

[46] Pasipamire, Notice, and Abton Muroyiwa. "Navigating Algorithm Bias in AI: Ensuring Fairness and Trust in Africa." *Frontiers in Research Metrics and Analytics*, vol. 9, Frontiers Media, Oct. 2024, https://doi.org/10.3389/frma.2024.1486600.

[47] "What Is an Open Source Library? Definition and FAQs | HEAVY.AI." *Www.heavy.ai*, www.heavy.ai/technical-glossary/open-source-library.

[48] "Addressing AI Bias: Real-World Challenges and How to Solve Them | DigitalOcean." *Digitalocean.com*, 2024, www.digitalocean.com/resources/articles/ai-bias.

healthcare.[49] This toolkit is often put to use in healthcare, because of situations like the biased algorithm in US hospitals in 2019, which predicted that black patients had lower costs than white patients for the same level of need.[50]

Microsoft's Fairlearn

Microsoft has developed an open-source toolkit, aiming to assess and improve the fairness of AI systems, and has achieved this aim to a high standard after being implemented.[51] Fairlearn provides different fairness criteria, based on the type of harm that can occur as a result of algorithmic bias.[52] Fairlearn has significantly contributed to the reduction of algorithmic bias as it has identified and mitigated it, which can be seen through a nearly equal balanced accuracy and false positive rate.[53]

Fairlearn has also been implemented in fields such as hiring and recruitment, ensuring fairness in candidate selection; in finance, ensuring fairness in credit scoring and loan approval; in healthcare, aiding in the creation of fair diagnostic or treatment recommendation models, and finally within education, guaranteeing fair admissions and scholarship recommendations.[54] The effective application of Fairlearn in these many different fields further proves the effectiveness and positive change brought by the implementation of Fairlearn.

Massachusetts Institute of Technology (MIT) Media Lab's Algorithmic Justice League (AJL)

The AJL conducts research and raises awareness about AI bias, through methods such as artistic expression[55], documentaries and campaigns.[56] By working with various people, the AJL has

---

[49] "Home." *AI Fairness 360*, ai-fairness-360.org/.
[50] "IBM Researchers Investigate Ways to Help Reduce Bias in Healthcare AI." *IBM Research Blog*, 9 Feb. 2021, research.ibm.com/blog/ibm-reduce-bias-in-healthcare-ai.
[51] "Fairness in Machine Learning — Fairlearn 0.13.0.Dev0 Documentation." *Fairlearn.org*, 2017, fairlearn.org/main/user_guide/fairness_in_machine_learning.html#.
[52] "Fairness in Machine Learning — Fairlearn 0.13.0.Dev0 Documentation." *Fairlearn.org*, 2017, fairlearn.org/main/user_guide/fairness_in_machine_learning.html#.
[53] "AI Fairness: A Deep Dive into Microsoft's Fairlearn Toolkit." *Opensenselabs.com*, 2025, opensenselabs.com/blog/ai-fairness.
[54] GeeksforGeeks. "Fairlearn: Assessing and Improving Fairness of AI Systems." *GeeksforGeeks*, 25 July 2025, www.geeksforgeeks.org/machine-learning/fairlearn-assessing-and-improving-fairness-of-ai-systems/
[55] "Person Overview ‹ Joy Buolamwini." *MIT Media Lab*, www.media.mit.edu/people/joyab/overview/.
[56] Buolamwini, Joy. "Algorithmic Justice League - Unmasking AI Harms and Biases." *Www.ajl.org*, 2023, www.ajl.org/.

underlined the need for reforming AI systems and developing algorithms that consider the needs and experiences of all demographic groups.[57]

Promoting fairness, it encourages other AI systems to implement guidelines or policies in order to reduce algorithmic bias within them, making it a less effective but also less costly method of reducing algorithmic bias.

Possible solutions

Bias Testing

This bias testing would involve evaluating different AI systems against known principles, such as upholding fairness, privacy, transparency, and accountability.[58] In order to detect inconsistencies in outcomes across different demographic groups, which in this case would be considered as distinct categories of people with shared characteristics, including age, race, and behaviour.[59]

This could be done in partnership with known guidelines of the National Institute of Standards and Technology (NIST).[60] These tests could be able to highlight areas where algorithms might unfairly favour or discriminate against certain groups.

Algorithmic Fairness Techniques

Implementing certain algorithm fairness techniques, such as re-weighting data, would allow us to have equal representation throughout all demographic groups used in the algorithm. This technique would include adjusting the weights of specific data points in order to fairly represent underrepresented groups in the training process of algorithms.

---

[57]Tackling Racial Bias in AI Systems: Applying the Bioethical Principle of Justice and Insights from Joy Buolamwini's "Coded Bias" and the "Algorithmic Justice League." Bangladesh Journal of Bioethics , 1 Mar. 2025, philarchive.org/archive/POLTRB-5.

[58] Grenawalt, Taylor. "Machine Learning Ethics: Understanding Bias and Fairness | Vation Ventures Research." *Www.vationventures.com*, 3 Oct. 2023, www.vationventures.com/research-article/machine-learning-ethics-understanding-bias-and-fairness.

[59] Golbeck, Jennifer. "Beyond the Individual." *Elsevier EBooks*, Elsevier BV, Jan. 2015, pp. 255–66, https://doi.org/10.1016/b978-0-12-801656-5.00023-8.

[60] Schwartz, Reva, et al. "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence." *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, vol. 1270, no. 1270, Mar. 2022, https://doi.org/10.6028/nist.sp.1270.

Transparency within AI Systems

Mandating transparency within AI systems, especially those used by the government, large businesses, or corporations with access to large amounts of personal data on individuals, could tackle algorithmic bias. This would work by mandating companies to disclose the purpose of the AI system and the types of data being collected and used. These AI systems, as well as mandatory disclosures, would be overseen by a UN-affiliated AI oversight agency, which would ensure that the human rights of individuals whose data is being used are protected.

Bibliography

"About Us — Fairlearn 0.12.0 Documentation." *Fairlearn.org*, 2021, fairlearn.org/v0.12/about/index.html.

"Admissible Heuristic: Causes & Solutions | BotPenguin." *Botpenguin.com*, 2024, botpenguin.com/glossary/algorithmic-bias.

African Union. *CONTINENTAL ARTIFICIAL INTELLIGENCE STRATEGY Harnessing AI for Africa's Development and Prosperity*. July 2024, au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf

"AI Fairness: A Deep Dive into Microsoft's Fairlearn Toolkit." *Opensenselabs.com*, 2025, opensenselabs.com/blog/ai-fairness.

"Algorithmic Justice or Bias: Legal Implications of Predictive Policing Algorithms in Criminal Justice." *JHULR*, 2025, jhulr.org/2025/01/01/algorithmic-justice-or-bias-legal-implications-of-predictive-policing-algorithms-in-criminal-justice/.

"Article 3: Definitions | EU Artificial Intelligence Act." *EU Artificial Intelligence Act*, artificialintelligenceact.eu/article/3/.

Aytekin, Ahmet. *Algorithmic Bias in the Context of European Union Anti-Discrimination Directives*. ceur-ws.org/Vol-3442/paper-41.pdf.

BBC. "Amazon Scrapped 'Sexist AI' Tool." *BBC News*, 10 Oct. 2018, www.bbc.com/news/technology-45809919.

Buolamwini, Joy. "Algorithmic Justice League - Unmasking AI Harms and Biases." *Www.ajl.org*, 2023, www.ajl.org/.

Cambridge Dictionary. "FILTER BUBBLE | Meaning in the Cambridge English Dictionary." *Cambridge.org*, 11 Dec. 2019, dictionary.cambridge.org/dictionary/english/filter-bubble.

"Proxy." *@CambridgeWords*, 12 Mar. 2025, dictionary.cambridge.org/dictionary/english/proxy.

Datatron. "Real-Life Examples of Discriminating Artificial Intelligence." *Datatron*, 25 Feb. 2021, datatron.com/real-life-examples-of-discriminating-artificial-intelligence/.

DigitalOcean. "Addressing AI Bias: Real-World Challenges and How to Solve Them | DigitalOcean." *Digitalocean.com*, 2024, www.digitalocean.com/resources/articles/ai-bias.

"EU Values Explained in One Minute." *Topics | European Parliament*, 29 Mar. 2021, www.europarl.europa.eu/topics/en/article/20210325STO00802/eu-values-explained-in-one-minute.

European Union. "The EU Artificial Intelligence Act." *The Artificial Intelligence Act*, 2024, artificialintelligenceact.eu/.

"Fairness in Machine Learning — Fairlearn 0.8.0.Dev0 Documentation." *Fairlearn.org*, fairlearn.org/main/user_guide/fairness_in_machine_learning.html.

"Fairness in Machine Learning — Fairlearn 0.13.0.Dev0 Documentation." *Fairlearn.org*, 2017, fairlearn.org/main/user_guide/fairness_in_machine_learning.html#.

Fitzgibbons, Laura. "What Is Feedback Loop? - Definition from WhatIs.com." *SearchITChannel*, Apr. 2019, www.techtarget.com/searchitchannel/definition/feedback-loop.

GeeksforGeeks. "Fairlearn: Assessing and Improving Fairness of AI Systems." *GeeksforGeeks*, 25 July 2025, www.geeksforgeeks.org/machine-learning/fairlearn-assessing-and-improving-fairness-of-ai-systems/.

Grenawalt, Taylor. "Machine Learning Ethics: Understanding Bias and Fairness | Vation Ventures Research." *Www.vationventures.com*, 3 Oct. 2023, www.vationventures.com/research-article/machine-learning-ethics-understanding-bias-and-fairness.

Gugu Ntsele. "Real-World Examples of Healthcare AI Bias." *Paubox.com*, Paubox, 11 May 2025, www.paubox.com/blog/real-world-examples-of-healthcare-ai-bias.

Hamilton, Melissa. "Justice Served? Discrimination in Algorithmic Risk Assessment." *Research Outreach*, 19 Sept. 2019, researchoutreach.org/articles/justice-served-discrimination-in-algorithmic-risk-assessment/.

"Home." *AI Fairness 360*, ai-fairness-360.org/.

"IBM Researchers Investigate Ways to Help Reduce Bias in Healthcare AI." *IBM Research Blog*, 9 Feb. 2021, research.ibm.com/blog/ibm-reduce-bias-in-healthcare-ai.

Jonker, Alexandra, and Julie Rogers. "What Is Algorithmic Bias?" *IBM*, IBM, 20 Sept. 2024, www.ibm.com/think/topics/algorithmic-bias.

Larson, Jeff, et al. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, 23 May 2016, www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

Lynch, Erin. "Tackling Racially Biased Health Care Algorithms - Petrie-Flom Center." *Petrie-Flom Center - the Blog of the Petrie-Flom Center at Harvard Law School*, 5 Oct. 2023, petrieflom.law.harvard.edu/2023/10/05/tackling-racially-biased-health-care-algorithms/.

Merriam-Webster. "Definition of ALGORITHM." *Merriam-Webster.com*, 2019, www.merriam-webster.com/dictionary/algorithm.

National Institute of Justice. "Recidivism." *National Institute of Justice*, 2023, nij.ojp.gov/topics/corrections/recidivism.

Norori, Natalia, et al. "Addressing Bias in Big Data and AI for Health Care: A Call for Open Science." *Patterns*, vol. 2, no. 10, Oct. 2021, p. 100347, https://doi.org/10.1016/j.patter.2021.100347.

Paschou, Vasiliki. "Bias in Artificial Intelligence: Risks and Solutions." *ActiveMind.legal*, 9 Apr. 2024, www.activemind.legal/guides/bias-ai/.

Pasipamire, Notice, and Abton Muroyiwa. "Navigating Algorithm Bias in AI: Ensuring Fairness and Trust in Africa." *Frontiers in Research Metrics and Analytics*, vol. 9, Frontiers Media, Oct. 2024, https://doi.org/10.3389/frma.2024.1486600.

Patalay, Prathamesh. "COMPAS : Unfair Algorithm ?" *Medium*, 22 Nov. 2023, medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a.

"Person Overview ‹ Joy Buolamwini." *MIT Media Lab*, www.media.mit.edu/people/joyab/overview/.

"Reweighting Meaning | Goong.com - New Generation Dictionary." *Goong.com*, 2024, goong.com/word/reweighting-meaning/.

"Risk Assessments Biased against African Americans, Study Finds." *Equal Justice Initiative*, 2 June 2016, eji.org/news/risk-assessments-biased-against-african-americans/.

Schwartz, Reva, et al. "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence." *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*, vol. 1270, no. 1270, Mar. 2022, https://doi.org/10.6028/nist.sp.1270.

Tackling Racial Bias in AI Systems: Applying the Bioethical Principle of Justice and Insights from Joy Buolamwini's "Coded Bias" and the "Algorithmic Justice League." Bangladesh Journal of Bioethics , 1 Mar. 2025, philarchive.org/archive/POLTRB-5.

"What Is Data Utility and Why It Matters in 2025." *Fanruan.com*, 2025, www.fanruan.com/en/glossary/what-is-total-sales-and-why-is-it-important/data-utility.

"Why and How Media Curation by Algorithm Contributes | Eticas." *Eticas*, 20 Dec. 2023, eticas.ai/why-and-how-media-curation-by-algorithm-contributes/.

Writer, Guest. "African Digital Colonialism Is the New Face of Worker Exploitation - ICTworks." *ICTworks*, 17 Apr. 2025, www.ictworks.org/african-digital-colonialism/.