

## Updates on AI bias and ethics

- Compiled by [Tracey Mollins](#) at [AlphaPlus](#)

Some people say that using an LLM - large language model - like ChatGPT is like using a calculator but calculators do not show us content that is racist, sexist or homophobic. We can work on our critical thinking skills to adapt to a AI world but what is the benefit of being exposed to this type of content?

[These Women Tried to Warn Us About AI](#) by Lorena O'Neil at Rolling Stone Magazine (August 2023)

*"Researchers — including many women of color — have been saying for years that these systems interact differently with people of color and that the societal effects could be disastrous: that they're a fun-house-style distorted mirror magnifying biases and stripping out the context from which their information comes; that they're tested on those without the choice to opt out; and will wipe out the jobs of some marginalized communities."*

[What ChatGPT Tells Us about Gender: A Cautionary Tale about Performativity and Gender Biases](#) in AI by Nicole Gross (June 2023)

*"This paper's central argument is that large language models work performatively, which means that they perpetuate and perhaps even amplify old and non-inclusive understandings of gender. Examples from ChatGPT are used here to illustrate some gender biases in AI. However, this paper also puts forward that AI can work to mitigate biases and act to 'undo gender'."*

[The Pear, You & AI](#) by Valentine Goddard

*"The Pear, You and AI is a women-led collaborative annotation initiative, designed as part of a larger project on Algorithmic Art to Counter Gender Bias in AI. In this initial phase, we are undergoing data collection based on your words and perceptions associated with words like women, beauty, imperfection."*

[A People's Guide to Artificial Intelligence](#) by Mimi Onuoha and Diana Nucera a.k.a. Mother Cyborg via Allied Media Projects ([PDF](#))

1. What does fairness look like when computers shape decision-making?
2. Who is creating the future, and how can we ensure that these creators reflect diverse communities and complex social dynamics?

*"This zine, published in August 2018, explores these questions through a series of explanatory text and whimsically illustrated pages that takes the reader on a journey that demystifies the often opaque world of artificial intelligence."*

## [5 Ethical Implications of AI in Education: A Guideline for Responsible Classroom Implementation](#)

by Luis Pardo (June 2023)

*"A responsible AI implementation in a school context begins with careful planning and consideration of all stakeholders' needs. This involves ensuring that AI tools are accessible and designed to accommodate diverse learning needs, including those of students with Special Educational Needs and Disabilities (SEND). Schools must ensure all students have access to the necessary technology to prevent the widening of the digital divide. The AI tools should be trained on diverse data sets to minimize algorithmic bias and should be designed to offer personalized learning experiences, considering each student's unique learning pace and style."*

[The Artificial Intelligence & Equality Initiative](#) from the Carnegie Council for Ethics in International Affairs

*"The **Artificial Intelligence & Equality Initiative (AIEI)** is an innovative impact-oriented community of practice seeking to understand the innumerable ways in which AI impacts equality for better or worse. We work to empower ethics in AI so that it is deployed in a just, responsible, and inclusive manner."*

[AI and education: guidance for policy-makers](#) from UNESCO (2021)

*"...while AI might have the potential to support the achievement of the Sustainable Development Goals (SDGs) of the United Nations, the rapid technological developments inevitably bring multiple risks and challenges, which have so far outpaced policy debates and regulatory frameworks. And, while the main worries might involve AI overpowering human agency, more imminent concerns involve AI's social and ethical implications – such as the misuse of personal data and the possibility that AI might actually exacerbate rather than reduce existing inequalities."*