



PARTNERSHIP ON AI

Retaining Diverse AI Talent: Why It Matters and What Teams Can Do

RECOMMENDATION
SUMMARY FROM
“AFTER THE OFFER:
THE ROLE OF
ATTRITION IN AI’S
‘DIVERSITY PROBLEM’”

As a field, AI struggles to retain team members from diverse backgrounds. Why is this such a widespread phenomenon and what can be done to close the gap? “After the Offer,” a new research paper from the Partnership on AI (PAI), seeks to answer these questions, providing four recommendations for how organizations can make the AI field more inclusive.

READ THE
RESEARCH
PAPER

AI systems have repeatedly shown troubling biases. The predominantly White and male makeup of AI teams could be one reason why.

AI systems are frequently criticized for propagating harmful biases against diverse populations. These criticisms are bolstered by the many examples of algorithmic bias in practice, such as:

- An [algorithm used to determine credit lines](#) being biased against female applicants
- The [exclusion or erasure of trans people](#) in health tech apps and other medical tech where outdated binary conceptions of gender are used
- Discrimination against home loan applicants by mortgage-approval algorithms, which [denied applications from people of color](#) at significantly higher rates than those of similar White people

Given the increasingly pervasive role of AI technologies in all of our lives, there has been increasing concern about who is designing this technology and whether they represent the people who are being affected by it.

People on AI teams, like those in the tech industry more generally, are predominantly White and male. Mitigating algorithmic bias may not be possible without workers who have a deep and comprehensive understanding of how this technology can negatively impact affected populations.

Diversity in the tech industry is weakened by employee attrition, a problem that disproportionately affects minoritized workers.

Although organizations often blame the “pipeline problem” (or the lack of available diverse talent) for the lack of diversity on teams, this does not address why AI workers with minoritized identities leave the field after entering it.

While no large-scale studies have been done on the reasons that workers leave tech organizations, the [most comprehensive account](#) of US tech industry attrition found that:

- Unfairness was the main driver of workers who voluntarily left tech jobs: 37% of respondents cited this as the main reason for leaving their previous employers.
- 30% of Black, Latinx, and Indigenous women reported being passed over for a promotion.
- About 25% of underrepresented people of color experienced stereotyping in their previous positions and 1 in 10 women reported receiving unwanted sexual attention in their previous jobs.
- At 20%, LGBTQ+ employees were also most likely to report bullying.

“After the Offer,” a new PAI research paper, provides four recommendations for promoting inclusion within AI organizations.

Through in-depth interviews with AI workers belonging to minoritized identities and others, PAI gathered rich information about how AI organizations can promote inclusion and positive team cultures. The paper recommends that organizations:

1 Systemically support Employee Resource Groups (ERGs)

- Employees specifically named ERGs as one of their main sources of support – even in work environments that were otherwise toxic.
- Institutions should seek a balance between supporting these groups and allowing them freedom to organize in pursuit of their own best interests.
- ERGs will not necessarily make an organization’s AI or tech more inclusive. Rather, systematically supporting ERGs will provide more support and community for minoritized workers, creating a more inclusive workplace in general.

2 Intentionally diversify at leadership and management levels

- Study participants repeatedly pointed to managers and upper-level leaders who belonged to minoritized identities (especially racial ones) as important influences.
- Bringing in folks from various academic, professional, and technical backgrounds to solve problems is especially crucial for AI teams.
- While intentional diversification at leadership and management levels is needed, it will not automatically create a pipeline for diversity at the leadership level, nor will it automatically override institutional culture or policies that ignore diversity, equity, and inclusion (DEI) best practices.

3 Make offered DEI trainings specific and more connected to the content of AI work

- Almost all participants reported that their organizations mandated some form of DEI training for staff. In general, trainings that they described as specific (e.g., trainings on anti-Black racism) received more praise than those that spoke to very general “DEI” topics such as race, culture, or implicit bias.
- However, even when they were specific, DEI trainings seemed to be disconnected from some aspects of the workplace climate or the content of what teams were working on.
- While being more specific and connected to the content of AI work could improve DEI trainings, several participants observed that such trainings alone will not overturn institutional stigmas against DEI. Meaningful change requires sustained effort and deliberate alignment of values.

4 Interrogate their values as practiced and fundamentally alter them to include the perspectives of people who are not White, cis, or male

- Participants frequently reported that a misalignment of values was a primary reason for them leaving (or wanting to leave) their organizations.
- Participants found it frustrating when organizations stated that they valued diversity and then failed to live up to this value with hiring, promotion, and day-to-day operations, ignoring the voices of minoritized individuals.
- If diversity is truly a value, organizations may have to investigate their systems of norms and expectations that are fundamentally male, Eurocentric, and do not make space for those from diverse backgrounds. They then must take additional steps to consider how such systems influence their work in AI.