

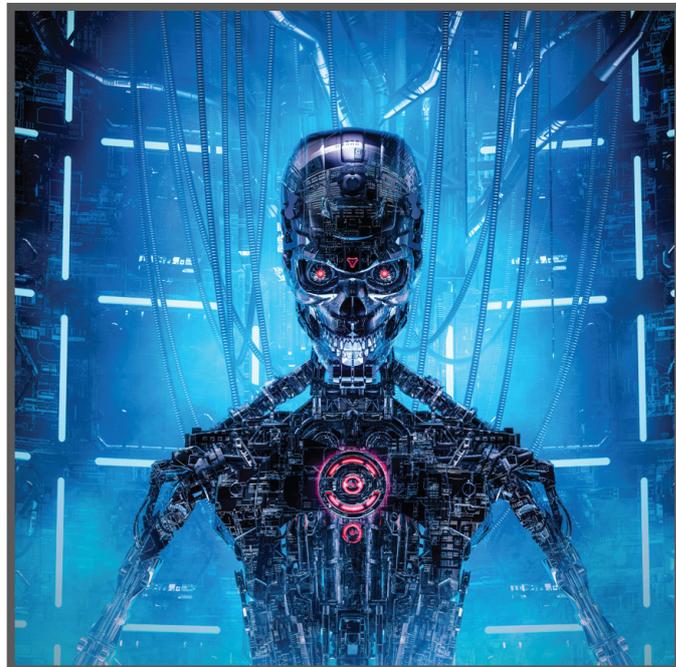


THE ROLE OF ETHICS IN AI



INTRODUCTION

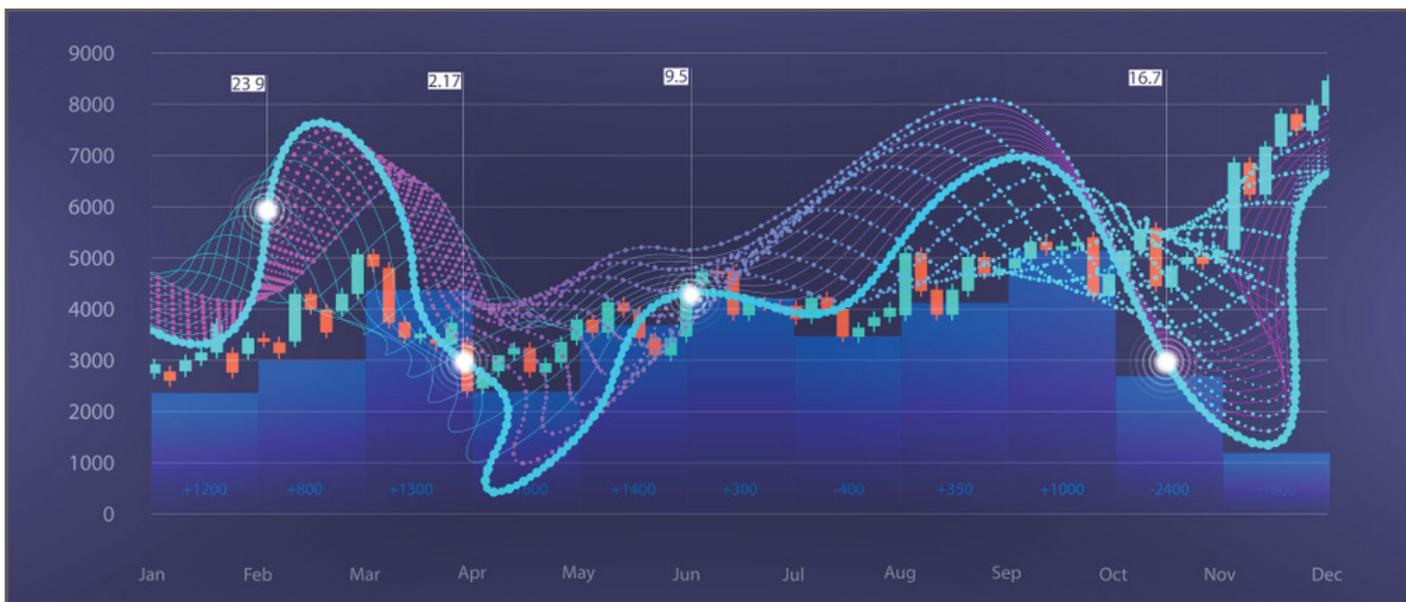
The ethical risks of artificial intelligence are extensively documented, ranging from biased recruiting systems to racist chat bots and mistakes in auditing algorithms. Organizations are scrambling to grasp the ethical significance of programs as the threat of AI legislation looms. The embryonic AI auditing sector is emerging to accommodate this need. However, the absence of a common AI ethical system has resulted in uneven methods – and concerns that AI auditing may not only fail to recognize unfairness, but it could even legitimize dangerous technology.



BROAD DIVISIONS

According to Emre Hagendorff, a UCL scholar and the founder of Integrative AI (an auditing and accounting start-up), a divided methodology to AI audits has emerged. Around one end of the spectrum, there seems to be light-touch consulting, in which businesses want high advice, an ethical plan, or a brand enhancement. A forensic investigation, on the other hand, involves an audit investigating an industry's sophisticated algorithms.

Nevertheless, major consultants such as PwC and Deloitte are establishing themselves as prominent competitors. According to McKee & Porter (2020), this is driving smaller competitors to distinguish themselves with specific specializations such as algorithm transparency and computational resilience. Some deliberately concentrate on 'de-biasing' since discrimination is a hot topic of focus. Many organizations also operate in the field of 'interpretability,' i.e., being able to transparently and accurately describe why an AI algorithm made the choice it did.



STANDARDS

A simple reality to grapple with is the fact that there is no true consensus on what constitutes as 'ethical AI.' A poorly conceptualized notion of 'ethical AI' may lead unscrupulous corporations to conduct farcical AI audits to legitimize harmful technologies, instead of analyzing relevant areas such as underpinning values, fundamental assumptions etc.

An example of such a misuse was the case of recruitment company, which used the facial features and actions to assess prospective applicants. One clearly worrying underpinning value or fundamental assumption in this use case was that character may be determined based on face traits, as well as evolutionary psychology, which links the form of the skulls to intellectual capabilities. As can often be the case, pseudoscience is deeply entwined with racism and hatemonger ideas.

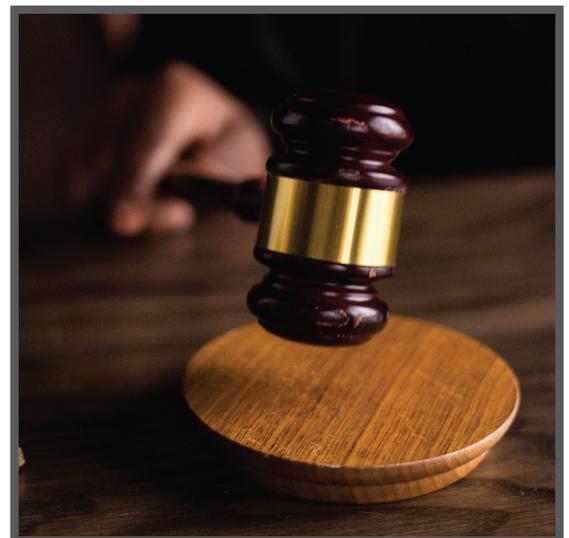


BUILDING CONSENSUS

The first step in implementing ethical AI would be to build consensus and converge on a formal framework. One difficulty towards achieving this is the complication that scientists and engineers, the components of the framework need to necessarily be defined in a precise manner. Bias, for example, has an extremely specific (and thus, possibly narrow), measurable meaning for AI practitioners. However, when the term 'bias' is thrown around in ethical AI conversations, it usually has a connotation of social justice, which is far less precise, almost to the point of being nebulous. In fact, it may not even be possible to solve completely with the aid of technology. This is further backed up by research conducted by academics at the Oxford University and the Alan Turing Institute. They discovered that technological work on AI ethics seldom corresponds to legal and philosophical concepts of ethics.

EXTERNAL REGULATION

One school of thought believes progress can be achieved if the burden is shifted to external regulatory agencies. Since corporations have little incentive to effectively conduct or publicize algorithm audits, much alone reveal that they utilize biased algorithms for profit and gain, having an external agency regulate by establishing transparent and objective standards might be a starting point. These could even be sector specific to have the appropriate bite. A sector specific example are the Fair Lending requirements in the US.



CONCLUSION

The construction of a framework to embed ethics in the applications of AI is a very lively topic now. While the consensus building, for all its challenges, is still an exercise in progress, Emre Hagendorff offers a few dimensions to consider democracy, privacy, durability, interpretability, and justice.

GET IN TOUCH

USA:

7320 East Fletcher
Avenue, Temple Terrace,
Tampa 33637

INDIA: Bangalore

IndiQube Penta 32,
Shanthala Nagar,
Richmond Town, Bengaluru,
Karnataka 560025

Email

info@vrize.com

