

AI Ethics: A Challenge to the Ethical Foundations of Statistics

Keywords: AI, Ethics, History of Statistics, Explainability

Projects in AI ethics tend to teach a set of exogenous principles – fairness, transparency, trust - that statisticians and engineers are expected to apply. The message of such projects is that, whereas AI systems may not be morally neutral, those who build them *are* applying instrumental, calculative reasoning to build towards a design, and must be taught the ethical ramifications of their finished designs. The upshot of this framing of technical work within AI is that AI is an applied science, which allows little room for agency, and that the builders of AI systems have agency primarily in the decision *whether* to build something. Once that decision is made, however, the agency of the builder is narrowed significantly to the instrumental application of statistics and computer science.

This paper argues that the alarming direction of AI technology will only be reoriented once AI is questioned no longer from the standpoint of ethical principles exogenous to the discipline but from within, using ethical resources internal to statistics. Rather than accept the predominant self-interpretation of those who build major AI systems today as working within statistics that is grounded in axiomatic mathematics and thus purely instrumental, AI ethics research must dig deeper to uncover the competing moral self-interpretations that have animated probability and statistics from its beginnings to the present-day. Such an account would restore the sense of ethical agency that the notion of technology as applied science espoused by educators conceals from engineers and statisticians.

The ideal of universal validity, unknown in the origins of classical probability animated by the recognition of universal rationality as the foundation of civil life, grew from its 19th century eugenicist roots to create the current identity crisis of statistics – as being both objectively free of bias and of social value thanks to its many applications, applications which always require some degree of subjective judgment. The only resolution of this tension is an unmasking of the pretense of such claims to objectivity and an assertion of the former modern ideal of probability as an account and clarification of good sense in the face of chance, so that sensible reason may solidly ground social order and progress. The vision for AI that is illuminated by this moral ideal from within statistics is of an iterative relationship between statistical models and sensible domain experts, co-creating intelligence.